

# 15

## Expressivism by Force

*Seth Yalcin*

### 15.1 Introduction

Here is what happens in this chapter.\* I make a certain distinction. I use the distinction to frame two directions that an expressivist view of normative language might take. I then plump for one of these directions.

Here is what happens in a bit more detail. I make a distinction: there is on the one hand the traditional speech act-theoretic notion of illocutionary force, and there is on the other hand the kind of notion of force we have in mind when we are theorizing in formal pragmatics about conversational states and their characteristic modes of update. I say these notions are different, and occur at different levels of abstraction. They are not helpfully viewed as in competition. I then say that the expressivist idea that normative language is distinctive in force can be developed in two sorts of directions, depending on which of the two senses of ‘force’ just distinguished is emphasized. One familiar tradition tries to develop expressivism as the thesis that the meaning of normative language is somehow to be explained via its putative connections to non-assertoric illocutionary forces. But that path is prone to Frege-Geach-style worries. Expressivists do better to take the other path, and start with the idea that normative discourse is distinctive in respect of its dynamic effect on the state of the conversation (Yalcin 2012a,b; cf. Lewis 1979a,b; Veltman 1996; Ninan 2005; Stalnaker 2014; Pérez-Carballo and Santorio 2016; Starr 2016; Willer forthcoming). This approach is not in principle subject to special worries about compositionality. It can be developed further using static semantic or dynamic semantic tools. It coheres with familiar lines of thinking about the metaphysics of content. It goes far, I suggest, in accommodating the core ideas expressivists have traditionally wanted to capture.

That summarizes the first seven sections of the paper. You might like to stop there. From section 15.8 forward, I go on a bit more about how one might develop an expressivism about normative language in this style, building on Gibbard (2003)’s notion of a plan-laden state of mind, and his technical notion of a ‘hyperplan’. I find much that is attractive in Gibbard’s formal model of normative states of mind, but as for the question how best to philosophically gloss that model, I take a different

\* Forthcoming for Daniel Fogal, Daniel Harris, and Matt Moss (eds.) *New Work on Speech Acts*. I am greatly indebted to Daniel Harris for illuminating comments on an earlier draft. Thanks also to Selim Berker and Edward Schwartz.

approach. I pursue a way for the expressivist to approach answering the question: “In virtue of what does a state of mind have the plan-laden content it has?” The sort of answer I recommend is broadly ‘functionalist’ and ‘representationalist’ in character. I suggest that the expressivist can approach this question in the same general way that, e.g. Lewis (1979c, 1994) approaches the analogous question as it arises for his modeling proposal about content. I deny that an expressivist in this vein has any special problem about explaining what it is for two normative states of mind to be inconsistent. Ultimately the sort of expressivism I envisage is perhaps distinctive in that it does not call for a radical rethinking of semantics, its foundations, or the theory of content; on the contrary, it presupposes and conservatively extends a broadly non-deflationary and representationalist conception of the mental, and can be made to mesh with compositional semantic theories of familiar varieties.

## 15.2 Separating illocutionary force and dynamic force

There are many things one could mean (and that have been meant) by ‘force’ in theorizing about linguistic communication. As advertised, I want to start by separating two broad ideas. The first idea I will call *illocutionary* force. The second idea I will call *dynamic* force.

Begin with illocutionary force. The terminology owes of course to Austin, but as I want to frame it, this general way of approaching force corresponds to a big tent, and goes back at least to Frege on assertion. There are two key ways that Frege gives us a handle on the notion of assertoric force. First, force is contrasted with an intuitive notion of content. Frege would say that the assertion that *p* and the query whether *p* both “contain the same thought” or have the same content, namely, that *p*. ‘Force’ is the name for the remainder, for the dimension that varies across these speech acts. This way of introducing the notion of force remains common in contemporary work.<sup>1</sup> Second, drawing on aspects of Kant’s theory of judgment, Frege described assertion as the outward manifestation of the inner act of judgment, where an act of judgment is something like an event of coming to believe, an event of taking-to-be-true. On such a view, assertoric force has a constitutive tie to the mental state of believing or judging, and (thereby) to the normative requirements governing these states of mind. You shouldn’t believe (judge) that *p* unless you have the appropriate epistemic relation to *p* (whatever that epistemic relation may be exactly—sufficient evidence, justification, knowledge, etc.). Since on this view assertion just is a way of manifesting a mental state of belief or judgment, it is naturally taken to be subject to normative requirements of a similar character.<sup>2</sup>

<sup>1</sup> An example from the first page of Searle and Vanderveken (1985): “In general an illocutionary act consists of an illocutionary force *F* and a propositional content *P*. For example, the two utterances, “You will leave the room” and “Leave the room!” have the same propositional content, namely that you will leave the room; but characteristically the first of these has the illocutionary force of a prediction and the second has the illocutionary force of an order.” See also Green (2015).

<sup>2</sup> A related approach is that of Williamson (1996, 2000), who argues that assertion is distinguished by a certain constitutive norm, one which makes reference to the knowledge state of the speaker. (Assertion could be said to ‘manifest’ a state of knowledge on this approach partly in virtue of the fact that the speech act itself is partly constituted by the rule one asserts *p* only if one knows *p*.) Williamson’s general

Austin (1961, 1962) followed Frege in conceiving of the force of a speech act as the sort of thing that (*inter alia*) situates the content expressed with respect to extra-conversational aspects of the speaker's state of mind, but his work placed a special emphasis on the broader rational objectives that animate speakers. The answer to the question of what the illocutionary force of a speech act is, is approached by asking what the agent is trying to do in speaking. Speech acts are individuated in part by appeal to the kinds of state of mind they are normally associated with, but also in part by appeal to the typical sorts of rational objectives associated with performing them. The illocutionary force of a speech act locates it at a level of description incorporating a broad sphere of human activity and social interaction. Searle (1969) is an extended development of Austin's approach. Some in this tradition, drawing on Grice (1957), place a special emphasis on a particular subclass of objectives, namely those having to do with getting the addressee to recognize the speaker's intentions (see, e.g. Strawson 1964; Bach and Harnish 1979). Within this broad approach to force, there is of course considerable room for debate about how exactly to divide the space of forces, and about how to analyze the force of any given speech act.

Since there is a seemingly boundless array of things speakers might do by using language, this way of thinking about force tends to lead to a rich diversity of forces. Searle and Vanderveken (1985), for instance, suggest that the following kinds of speech acts all correspond to characteristically different forces: assertions, predictions, reminders, objections, conjectures, complaints, orders, requests, declarations, promises, vows, pledges, apologies, admissions, boasts, laments, and bets. (That is a selection; their official list is longer.) To perform any of these acts is in part to manifest that one's beliefs, desires, intentions, and/or actions outside of the context of the discourse are subject to certain norms. (Perhaps, as with performatives, in virtue of the very performance of the speech act.)

A second, very different thing one could have in mind by the 'force' of an utterance is (something like) the characteristic kind of change to the state of the conversation the utterance is apt to produce. This sort of idea—what I want to call *dynamic* force—began to come into focus in the seventies, when theorists like Karttunen, Stalnaker, Lewis, Kamp, and Heim began to formalize pragmatics and explore various ways of making the whole conversation or discourse itself the object of systematic formal investigation (Karttunen 1969, 1974; Stalnaker 1974, 1978; Lewis 1979b; Kamp 1981; Heim 1982; see also Hamblin 1971; Gazdar 1979).<sup>3</sup> The account of assertion in Stalnaker (1978) provides a paradigm case of this way of thinking about force. On this view, the characteristic conversational effect of successful assertion is to change the conversational state by adding information to it. Stalnaker models the state of a conversation by a set of possible worlds, the possible worlds left open by what is jointly presupposed by the interlocutors in the discourse—what he calls a *context set*. To gain information is, in the possible worlds setting, for there to be fewer possibilities

approach suggests that speech acts forces are individuated by their "constitutive rules"—rules which (like the knowledge rule for assertion) are typically articulated via appeal to the intentional mental states of the interlocutors—their knowledge, beliefs, preferences, etc.

<sup>3</sup> While the present paper is greatly indebted to these works, the particular way of framing the notion of dynamic force that will emerge here may not exactly align with any of them.

compatible with your information, so the effect of assertion is modeled as an operation that eliminates possibilities from the context set. When one says, ‘It’s raining’ and all goes well,<sup>4</sup> possibilities in the context set where it is not raining are eliminated. This reflects the fact that our shared conversational information now ceases to be compatible with the possibility that it isn’t raining.

The array of possible dynamic forces depends on the variety of interesting dynamical changes that conversational states are capable of—a matter which of course interacts with the question what kind of structure conversational states are best modeled as having. Theorists working with a dynamic conception of force often use richer objects than context sets in their models of conversational states, recognizing additional structure as necessary to model whatever language fragment is of target concern.

To give an illustration of this approach beyond ordinary assertion, consider questions. Here it is obvious that without elaboration, the simple context set picture is inadequate: a question like ‘Is it raining?’ does not characteristically add information (eliminate possibilities) to the conversation; but neither does it remove information (add possibilities). Abstractly, a question seems to “frame an issue” in a way that serves to steer the discourse in a particular direction. As Hamblin suggested in classic work, “Pragmatically speaking a question sets up a choice situation between a set of propositions, namely those propositions that count as answers to it” (Hamblin 1973, 254). Since the work of Hamblin, and Karttunen (1977), it has been usual to understand the semantic value of an interrogative sentence as the kind of thing that determines a set of propositions, the propositions that could serve to answer the question. Where the complete possible answers to a question are mutually exclusive, we can think of a question as determining a partition of logical space (Hamblin 1958; Belnap and Steel 1976; Groenendijk and Stokhof 1984). Now we can bring this idea into our model of conversational states in various ways, in order to clarify and model the dynamic force of questions of this kind. One simple possibility would be as follows (cf. Roberts 1996, 2012; Hulstijn 1997; Yalcin 2011):<sup>5</sup> we suppose that a conversational state includes, in addition to a set of open possibilities (a context set), a set of ways of partitioning logical space. We use the latter element to model the question(s) that are in focus in the discourse. The characteristic dynamic effect of a question would be to eliminate partitions from this set. Take for example the question, ‘Is it raining?’ This will semantically determine a simple bipartite partition of logical space into

<sup>4</sup> By “All goes well” I mean that speaker was heard by the addressee, that both speak English, that each takes herself to be in conversation with the other, etc.

Stalnaker glosses the dynamic update effect of assertion as a “proposal” to change the context set: my assertion of ‘It’s raining’ changes the context set in the way described only if my interlocutor does not object. I think it is best to take the “proposal” talk metaphorically, and not view assertion as literally explained as an illocutionary act of proposing. (That would just pass the buck to the question of what proposing is.) I myself would favor dropping the “proposal” talk entirely, holding instead that assertions simply always change the state of the conversation in their characteristic fashion. That is to say, the update does not “wait” for the addressee’s permission, implicitly or explicitly. Rejections of assertions do not stop the relevant changes to the conversational state from happening; rather, they undo a change that has taken place.

<sup>5</sup> It should be noted that the idea that the context set is interestingly partitioned in a manner that is sensitive to the distinctions of interest to the interlocutors is an idea Stalnaker has raised in various places; see, e.g. Stalnaker (1981, 1986, 2014).

rain possibilities and no-rain possibilities. The dynamic effect of the question would be to remove from the conversational state those partitions of logical space that fail to incorporate at least this distinction. That is, it serves to rule out partitions that fail to cut logical space in a manner that is sensitive to the question of rain.<sup>6</sup> By eliminating the partitions that fail to include the question of rain, the question of rain becomes, as it were, ‘visible’ or ‘in focus’ in the discourse.

Imperatives make for another illustration. Perhaps the best known recent account in the dynamic force style is due to Portner (2004, 2007, 2017). He suggests that imperatives semantically express individual (address-relativized) properties. He proposes (roughly) that a conversational state includes, for each interlocutor, a “To-Do List” for that interlocutor—a sequence of properties that the agent is mutually understood in the conversation to be under some kind of requirement to realize. He then proposes that the dynamic force of an imperative is to update the To-Do List of the addressee, adding the property expressed by the imperative to the addressee’s To-Do List.

For at least a half-dozen more ways of enriching conversational states in order to associate certain fragments of language with distinctive dynamic forces, see Lewis (1979b). He postulated a rich “conversational scoreboard” including a number of dimensions beyond the context set, among them: (i) a ranking of comparative salience of objects (for, e.g., tracking the way that the referent of a definite description may be sensitive to the preceding discourse); (ii) a parameter tracking the prevailing standards of precision (for modeling vagueness; see also Lewis 1980, King 2003); (iii) a component registering the admissible modal accessibility relations (for modeling the dynamic effect of unembedded modal utterances); (iv) a parameter mapping names to their bearers (for modeling the dynamic effect of performative speech acts of dubbing); (vi) a component representing the possible plans of the interlocutors (for modeling talk of what to do).

I take it that dynamic force is basically the notion in play in what has lately come to be called ‘dynamic pragmatics’ (see, e.g. Stalnaker (this volume) and Portner (this volume)). But while typical proponents of dynamic pragmatics frame that view as packaged with a rejection of dynamic semantics, the notion of dynamic force operative here is meant to be neutral on that issue. (The notion of dynamic force occurs at what Rothschild and Yalcin (2015, 2016) call the ‘conversation systems’ level of description.) Whether or not the compositional semantic value of a sentence is identical to its way of updating the conversational state (its *context change potential*, or CCP), the dynamic semanticist and the dynamic pragmaticist agree that (unembedded) sentences *have* CCPs. Both can therefore ask, for various fragments of language, what interesting distinctions there are to be made amongst the CCPs they

<sup>6</sup> More exactly: say a partition of logical space (question)  $\Pi_1$  *includes* partition  $\Pi_2$  just in case every element of  $\Pi_2$  is equal to the union of some set of elements of  $\Pi_1$ . Then the dynamic effect of a question is to eliminate from the conversational state those partitions that do not include the question. This is basically the notion of inclusion defined by Groenendijk and Stokhof (1984) and by Lewis (1988a,b).

(I don’t mean to defend this particular theory of the dynamic force of questions here against relevant competitors, the most obvious being perhaps Roberts (1996, 2012), who postulates *inter alia* a component of the conversational state that tracks a sequence of questions ordered by priority. I just want a simple example of a dynamic force picture of (partition-like) questions on the table.)

recognize. And as I am understanding it, that is just the question: what the interesting distinctions are between dynamic forces.

### 15.3 That illocutionary force and dynamic force do not line up nicely

With illocutionary force and dynamic force distinguished, the next thing to emphasize is that they need not line up in any particularly neat way. The things we are apt to call assertions in the illocutionary sense may be diverse in respect of their dynamic forces; and in the other direction, it may be that a single underlying dynamic force is what is in play across speech acts with diverse illocutionary forces. I do not take this to be a new point. Stalnaker (1978) already noted that his idea about dynamic force seemed appropriate to cases of what we would naturally call, in the illocutionary sense, ‘assertion’ and also to what we would naturally call in the illocutionary sense ‘supposition’, and that therefore the account was not helpfully understood as an analysis of the traditional (illocutionary) notion of assertion. Indeed, it is rather misleading to use the same words—‘assertion’ and ‘force’—in both the illocutionary way and the dynamic way. That can suggest competing analyses of the same phenomenon. But there is no conflict here. These concepts apply at different levels of description.

It might be right to suppose that Stalnakerian assertoric dynamic force is put to an especially common and important purpose in normal linguistic interaction, namely the purpose of transmitting belief or knowledge. That is why it made some sense for Stalnaker to describe his dynamic model of conversational update as a model of what assertion typically does, taking ‘assertion’ in tradition illocutionary sense. But this use to which the kind of dynamic force described by Stalnaker can be put—the use of transmitting belief or knowledge—should not be wrapped into its identity, or its conditions for individuation. To do that would be to lose some of the power of this way of theorizing about conversation.

Stalnaker’s dynamic force account captures an abstract idea—basically, the idea of adding some more information to a certain existing body of information (viz., the conversational state). As I want to recommend we understand it, the story prescind from the question what exactly the conversational state is taken by the interlocutors to be characterizing, and from the question what the speaker might be aiming to do, extra-linguistically speaking, by adding certain information to that state. In this way, it prescind from exactly the kind of facts that are thought to be essential for individuating speech acts on illocutionary conceptions of force. This level of abstraction for theorizing about conversation is high—or if you prefer, narrow—but it is a fruitful one for modeling core features of discourse and of our linguistic competence. It makes sense to distinguish the game of updating conversational states from the diverse uses to which this game could be put—even when certain uses seem particularly canonical or salient or important. In making a series of declarative statements, a person may be recounting events that transpired yesterday, or they may be telling a story everyone mutually recognizes to be a fiction. In both cases, we should like to say that they are exploiting the meaning of their sentences and certain conventions about the dynamics of conversational update in their language

to add information to something like a store of information already mutually held in common. In both cases we might naturally model the impact of these speech acts on the conversational state in Stalnaker's way, in terms of the elimination of possibilities from a context set. The concept of dynamic force enables us to capture key similarities about the dynamics of discourse across diverse uses of language. Thus if we use 'assertion' in the dynamic force sense, we should not be taken to be assuming anything very substantive about what the illocutionary force of the speech act in question was.

There is a certain point that is apt to get lost in the preceding, so let me pause to draw it out and emphasize it. One should not assume that the information captured by the conversational state, in the target technical sense of 'conversational state', must reflect what is common belief among the interlocutors. The information incorporated into the conversational state needn't be common belief or common knowledge.<sup>7</sup> Sometimes more is conversationally common ground than what is common belief—as when we converse under (explicit or tacit) hypothetical suppositions, or when we make polite conversation, allowing presuppositions into the conversation that we don't plan to take home. Other times, the state of the conversation does not include propositions that are common belief, as when we reason under counterfactual suppositions, or tell stories. Moreover, what exactly one comes away from a conversation actually believing or knowing always depends on subtleties about trust and authority (real and perceived). Such factors may, but needn't, influence what one lets into the conversational state. Whether they will or not in any given case depends on the goals and interests of the interlocutors, and on the mutually understood point of the conversation. A model of the abstract dynamics of conversation, and of the general way in which the state of the conversation is changed by speech acts, can to a great degree abstract from these factors.<sup>8</sup> We can postulate a basic mental state—call it *presupposition*, or the *conversational state*—to play the desired role in theorizing about linguistic communication. It is this state, in the first instance, that we are coordinating on in conversation. (Thus we can say, for example, that  $\phi$  is *common ground* in a conversation just in case it is common knowledge, or common belief, among the interlocutors that  $\phi$  is being presupposed.) We need not show how to reduce this state of mind to other, more familiar states in advance of theorizing. If it helps us to explain things, it will earn its own keep.<sup>9</sup>

<sup>7</sup> In this paragraph I repeat some points made in Yalcin (2012a).

<sup>8</sup> Kölbel (2011) puts the gist of it well:

the exchange of information is only one among many ultimate purposes that linguistic exchanges can have. When we converse in pursuit of the aim of information exchange, we do so by pursuing the language-internal objective of changing the conversational score, an objective that can serve many other aims too. We will gain a better understanding both of conversation and of information exchange if we keep this in mind. (51)

The basic point carries over to dynamic forces which are not (or not merely) information-adding moves.

<sup>9</sup> To be clear, I am not denying in advance of inquiry that the target notion of a conversational state might somehow be reduced to other mental states. What I am rejecting is the presupposition that some such reduction must be carried out in order for theorizing to proceed. One needs to moor the technical notion in a sufficient body of explanatory theory before clear questions of reduction can be framed and profitably pursued.

So again, there is no tension between the illocutionary and dynamic approaches to force. They are not competing analyses of the same phenomena, but are rather concerned with different explananda. Readers who enjoy Austinian taxa should perhaps situate dynamic force as a feature of what he called the “locutionary act”, inasmuch as fixing the dynamic force of a speech act still generally leaves it substantially underdetermined what the speaker was up to, or was trying to do, in performing a speech act with that dynamic force.

In separating illocutionary and dynamic notions of force, I am not trying to attack those theorists who hope to theorize about illocutionary force by appeal to a notion of dynamic force—who want to offer a theory of illocutionary forces which appeals partly to an independently understood notion of dynamic force. On the contrary, such theorists should welcome the distinction I press, since in making this separation, I am drawing out a sense in which they can claim to be explaining features of illocutionary acts using independently understood materials.<sup>10</sup>

One kind of theorist in this vein holds that dynamic forces can be used to group illocutionary forces. Stalnaker’s abstract treatment of the dynamic effect of assertion, for instance, might be argued to group together what ordinary speakers call “assertion”, but also other speech acts like supposing or hypothesizing. I myself am not especially interested in the prospects for constructing a botany of illocutionary acts, so I leave it to others to show that such an approach might yield some explanatory insight.<sup>11</sup> In any case, whether such a theory can be worked out is orthogonal to the point I’m making, which is just that there is an interesting notion of dynamic force which does not reduce to, and is not a species of, illocutionary force.

It could well be that the specific array of dynamic forces we in fact find in natural language is explained, or partly explained, by the fact that we have an interest in using such forces to perform various illocutionary acts. That idea is entirely compatible with everything said so far.

On the view I favor, knowing the dynamic forces associated with the sentences of a language is part of linguistic competence with the language. If we think of language as complex tool that we do things with, I am situating dynamic force as a component of

<sup>10</sup> Am I one of these theorists? I see no problem with informal elucidations of familiar sorts of illocutionary acts in part by appeal to their dynamic conversational effects; on the contrary, I employ such elucidations on occasion below, and have done so elsewhere (e.g. Yalcin 2007, Yalcin 2012a). What is less clear to me is whether instructive, theoretically fruitful general analyses of illocutionary acts/forces are possible. It does not appear that the literature licenses great optimism about the prospects for a science of illocutionary acts, where such a theory is understood as taking us beyond a relatively shallow botanization of human speech behaviors, the later stated mostly in terms of common sense categories. Theories in this vein seem at risk of devolving into conceptual analyses of common sense speech act notions, with concomitant loss of grip on what was supposed to be getting explained. There is, relatedly, a general worry, emphasized by Chomsky (2000), about the (un)fitness of ordinary common sense notions for use in scientific inquiry (especially when what is to be explained is human language and behavior). Common sense notions rarely perform well when pressed into service as theoretical notions; they have their own lives. (I try to expand on this latter worry in Yalcin (forthcoming).)

<sup>11</sup> The question here is whether a speech act’s having an illocutionary force of a certain kind implies it has a dynamic force of a certain kind. On the face of it, it seems not: it seems one can know that, for instance, a speech act had the illocutionary force of a command without knowing whether it had the dynamic force characteristic of an imperative, declarative, or interrogative—plausibly one can issue commands in the illocutionary sense via various kinds of context-change potential.



the tool, not as a component of actions of using the tool. Once we start talking about actions of using the tool—once we are talking about interlocutors qua rational agents, using language to achieve various objectives, communicative and otherwise—we are at the speech act level of description.

## 15.4 Normative language as distinctive in force

Having now separated two very different sorts of thing one could mean by ‘assertion’, let us consider the idea that normative sentences are not assertion-like. This is an idea that has permeated expressivist approaches to normative discourse since those views began to take shape in the first half of the twentieth century. Some representative early statements of this idea, beginning with W. H. F. Barnes:

Value judgements in their origin are not strictly judgements at all. They are exclamations expressive of approval. This is to be distinguished from the theory that the value judgement, “A is good,” states that I approve A. The theory that I am now putting forward maintains that “A is good,” is a form of words expressive of my approval. To take an illustration :— When I say “I have a pain,” that sentence states the occurrence of a certain feeling in me: when I shout “Oh!” in a certain way that is expressive of the occurrence in me of a certain feeling. We must seek then for the origin of value judgements in the expressions of approval, delight, and affection, which children utter when confronted with certain experiences. (Barnes 1934, 45)

Carnap:

The rule, “Do not kill,” has grammatically the imperative form and will therefore not be regarded as an assertion. But the value statement, “Killing is evil,” although, like the rule, it is merely an expression of a certain wish, has the grammatical form of assertive proposition. Most philosophers have been deceived by this form into thinking that a value statement is really an assertive proposition, and must either be true or false. . . . But actually a value statement is nothing else than a command in a misleading grammatical form. (Carnap 1935, 24)

Russell:

If, now, a philosopher says “Beauty is good,” I may interpret him as meaning either “Would that everybody loved the beautiful” . . . or “I wish that everybody loved the beautiful” . . . The first of these makes no assertion, but expresses a wish; since it affirms nothing, it is logically impossible that there should be evidence for or against it, or for it to possess either truth or falsehood. The second sentence, instead of being merely optative, does make a statement, but it is one about the philosopher’s state of mind, and it could only be refuted by evidence that he does not have the wish that he says he has. This second sentence does not belong to ethics, but to psychology or biography. The first sentence, which does belong to ethics, expresses a desire for something, but asserts nothing. (Russell 1935, 236–7)

Ayer:

...in every case in which one would commonly be said to be making an ethical judgment, the function of the ethical word is purely “emotive”. It is used to express feeling about certain objects, but not to make any assertion about them. (Ayer 1936, 108)

Stevenson (1937), building on Ogden and Richards (1923):

Doubtless there is always *some* element of description in ethical judgments, but this is by no means all. Their major use is not to indicate facts, but to *create an influence*. (18)

When you tell a man that he oughtn't to steal, your object isn't merely to let him know that people disapprove of stealing. You are attempting, rather to get *him* to disapprove of it. Your ethical judgment has a quasi-imperative force . . . (19)

These views were set against what Austin (1962) and others later framed as the *descriptive fallacy*:

It was for too long the assumption of philosophers that the business of a 'statement' can only be to 'describe' some state of affairs, or to 'state some fact', which it must do either truly or falsely. (1)

It has come to be seen that many specially perplexing words embedded in apparently descriptive statements do not serve to indicate some specially odd additional feature in the reality reported, but to indicate (not to report) the circumstances in which the statement is made or reservations to which it is subject or the way in which it is to be taken and the like. To overlook these possibilities in the way once common is called the 'descriptive' fallacy . . . (3)

Considering now the thesis that normative sentences do not have the force of assertions, two questions confront us:

- (i) What sense of 'assertoric force' is at issue? Do we mean 'force' in some illocutionary sense? Some dynamic sense? Both?
- (ii) What is the nature of the link between the meaning of normative vocabulary and the putatively non-assertoric force of normative sentences?

Since as noted, the dynamic notion of force appeared on the scene only in the seventies, early forays into the prospects for expressivist approaches to normative language worked with conceptions of force in the illocutionary vein. As for the nature of the link between the meaning of normative vocabulary and the putatively non-assertoric force of normative sentences, the details here were often less than completely clear.<sup>12</sup>

## 15.5 Frege-Geach

However, it was clear what the opponents of these approaches, notably Geach (1965) and Searle (1969), took the view to be saying, or trying to say. They took the idea to be that the connection between the meaning of normative vocabulary and the putatively non-assertoric, non-descriptive force of normative sentences is very tight. The meaning of normative vocabulary was to be explained by appeal to the distinctive sorts of non-assertoric, non-descriptive illocutionary acts they ostensibly participate in. The meanings of 'good' and 'bad', for instance, were to be explicated

<sup>12</sup> None of the above cited authors offered anything like a detailed account, for instance. Undoubtedly Hare (1952) was the most detailed mid-century attempt at working out the details.

by appeal to the observations that ‘good’ is used to perform the speech act of commending, and the word ‘bad’ is used to perform the speech act of condemning. The semantics of normative terms was somehow to be a matter of associating them with the distinctive speech acts they enable. The usual thought was that these distinctive speech acts corresponded to the expression of distinctive non-doxastic (or not entirely doxastic) states of mind—perhaps desire-like states of mind if Carnap and Russell were on the right track; perhaps something more emotional in character, if Barnes and Ayer were on the right track; perhaps some mix, if Stevenson was on the right track.

The fundamental difficulty with this way of developing the idea, as Geach (1965) and Searle (1969) famously stressed, is that meanings are compositional, whereas illocutionary forces seem not to be. Here we can separate two related points.

First and most obviously, the mere appearance of a particular word in a construction cannot make it the case that a speech act of a particular illocutionary type is being performed with that construction.<sup>13</sup> There are ever so many sentences in which ‘good’ (etc.) appears wherein there is no plausibility that the corresponding speech act need be a commendation. On the contrary, choose virtually any speech act force you please, and we will be able to produce an example of an illocutionary act with that force which involves tokening a sentence wherein ‘good’ (etc.) figures. This point is widely appreciated, so I’ll skip examples.

The second point emphasizes the degree of disanalogy between meaning and illocutionary force in respect of compositionality. Whatever the meaning of a complex expression is, it had better be by and large fixed by the meanings of its constituent parts plus the syntax of the expression. The assumption of compositionality is part of what is required for natural language semantics to play its part in explaining the productive character of language understanding and use.<sup>14</sup> The illocutionary force associated with the utterance of a sentence, on the other hand, is not helpfully understood as something somehow fixed by the forces putatively associated with the individual primitive parts of the sentence and their combination. Generally speaking, it rarely if ever even makes sense to speak of the illocutionary force of subsentential constituents of a sentence uttered. The locus of illocutionary force is the whole utterance-in-context qua intentional act. What force an utterance has may certainly be constrained in interesting ways by the meaning of the sentence uttered, but the illocutionary force of a whole utterance is not somehow built up out of the putative forces of the sub-utterances of the constituent words.<sup>15</sup>

The classic Frege-Geach critique encourages a certain understanding of what the expressivist thesis is, or was supposed to be. It suggests that expressivism is a view that

<sup>13</sup> There may be some limited exceptions. Perhaps one performs a speech act of derogation in virtue of using a slur, no matter how deeply embedded it appears.

<sup>14</sup> I offer a more detailed discussion of this point in Yalcin (2014).

<sup>15</sup> Sometimes even anti-expressivist theorists, e.g. Geach (1965), slip into a way of talking that suggests that subsentential clauses of sentences may have illocutionary forces. But theorists who slip into this kind of talk rarely defend it or render it conceptually clear. Sometimes what theorists have in mind by this kind of thing is clearly something more easily rendered via the idea of dynamic force, about which more shortly. (Geach’s slips are mostly in connection with facts that would be described from a modern perspective as facts about presupposition projection—a famous impetus for the development of the notion of dynamic force.)

is supposed to take on the challenge of delivering a theory which derives the forces of utterances compositionally—a theory which teaches how the meaning of normative expressions can be given by explicating their connections to distinctive illocutionary forces, in such a way as to respect the compositionality of meaning. We could call this kind of approach to developing expressivism *compositional force expressivism*.

Compositional force expressivism has a close cousin. According to it, expressivism is the view that is supposed to take on the challenge of delivering a theory which associates, not forces, but attitudes—various mental state types—with sentences compositionally. We could call this version *compositional attitude expressivism*. The meaning of (e.g.) ‘good’ is not given by directly associating it with speech acts of commendation; rather, it is associated with some kind of ‘pro-attitude’, or state of preference, or something along those lines. Calling *x* good still amounts to performing a distinctive speech act of commendation; but what makes the speech act a commendation is in part that the meaning of ‘good’ is somehow, as a matter of its meaning, tied to the attitude of favoring *x*. This is roughly the sort of way that expressivism has been understood by many in more recent times. Thus for example Rosen (1998), discussing Blackburn (1993):

The centerpiece of any quasi-realist (expressivist) ‘account’ is what I shall call a psychological semantics for the region: a mapping from statements in the area to the mental states they ‘express’ when uttered sincerely. The procedure is broadly recursive. Begin with an account of the *basic states*: the attitudes expressed by the simplest statements involving the region’s characteristic vocabulary. Then assign operations on attitudes to the various constructions for generating complex statements in such a way as to determine an ‘expressive role’ for each of the infinitely many statements in the area. (387–8)

Similar characterizations of expressivism can be found in Blackburn (1998), Wedgwood (2007), Schroeder (2008c), Gibbard (2003), and Charlow (2015).

On the face of it, the prospects for working out compositional attitude expressivism are as bleak as the prospects for working out compositional force expressivism. From the point of view of the skeptic versed in modern natural language semantics and pragmatics, they both seem based on the same kind of category mistake, mislocating the locus of compositionality.

I leave it to others to sort out whether some version of compositional force expressivism or compositional attitude expressivism could be rendered viable. It seems to me that there is a more promising path for developing the kind of expressivist themes sounded by the authors cited above. It is basically the sort of path I have explored elsewhere in connection, not with normative talk, but with epistemically modal talk (Yalcin 2007, 2011, 2012a). This path does not involve any radical reconception of the notion of illocutionary force. Nor does it involve any attempt at a compositional mapping from sentences to mental states. Nor does it require a total rethinking of the foundations of semantics. But to get to this kind of view, we need to say some things about how best to conceive of the relations between compositional semantic value, content, and dynamic force.<sup>16</sup>

<sup>16</sup> I regret I lack the space in this paper to chart the ways that my take on the Frege-Geach problem differs from others in the literature. The recent literature is especially influenced by the framing of Unwin

## 15.6 The locus of compositionality

The locus of compositionality is not force, and neither is it attitude—so in a nutshell goes the Frege-Geach critique of textbook expressivist views. But Frege and Geach themselves could be critiqued for mislocating the locus of compositionality. Specifically, they could be chided for failing to observe the distinction between content and compositional semantic value, and for misconstruing the relationship between content and the demands of compositionality.

The semantic value–content distinction is stressed in various ways in Dummett (1973, 1993), Lewis (1980), and Stanley (1997a,b), and more recently in Yalcin (2007, 2012a, 2014), Ninan (2010), Rabern (2012a,b, 2013), and Yli-Vakkuri (2013).<sup>17</sup> These works differ in where they place the stress, and in the terminology used (in Dummett, the distinction appears as that between ‘ingredient sense’ and ‘assertoric content’). My preferred take appears in Yalcin (2014) (it owes significantly to Lewis (1980), Rabern (2012b), and to conversations with Ninan). Without rehearsing the full story told there, the basic thought is that the requirements on a notion of content suitable for modeling the mental states we traditionally call ‘propositional attitudes’ are importantly different from the demands appropriate to the notion of linguistic meaning (semantic value), and in such a way that we shouldn’t expect the realizers of the content role to line up in some particularly straightforward way with the realizers of the semantic value role. In particular, there is little reason to theorize under the assumption that the semantic values of sentences (in context) are identical with the objects we find useful to call, in the theory of mental content, ‘contents’ (‘propositions’, ‘propositional content’, etc.).

This isn’t to say that we don’t or can’t assert propositions in something like the traditional sense using ordinary declarative sentences. One can of course still have that view compatible with respecting the distinction between semantic value and content. It’s just that in such cases, it sows less confusion to see the matter like this: the compositional semantic value of the sentence *determines*, as a function of context, the item of content asserted (cf. Lewis 1980). That is all that is necessary to uphold the idea that we can assert propositions using declarative sentences—viz., that there be some bridge principle, understood as a feature of the pragmatics or ‘post-semantics’ of the language, connecting the target class of sentences to propositional contents via the semantic values and relevant features of context. Thus the suggestion isn’t that semantic value and content are wholly disconnected. On the contrary, most will naturally want to take the theory of mental content and the theory of linguistic meaning to have deep and important interconnections. The point is merely that we should distinguish these theoretical concepts and their associated explananda.

Semantic values are the locus of compositionality in natural language. It is hardly a contestable thesis that they are compositional, since it is their job to do their part in explaining the productive character of language understanding and use, and this job

(2001): see in particular Gibbard (2003), Dreier (2006), Schroeder (2008a,b,c). See also Charlow (2014), Ridge (2014), Woods (2017).

<sup>17</sup> Burge (1979) argues that theorists have exaggerated the extent to which Frege pressed senses into work as linguistic meanings. In a sense, Burge could be read as arguing that Frege was alive to what I am here calling the semantic value–content distinction. This isn’t the place to pursue the exegetical question, but I discuss the issue briefly in (Yalcin 2015, 242).

appears undoable without compositionality. Content, on the other hand, may not be compositional in anything like that sense. The assignment of contents to mental states may be a more global, holistic matter—so in fact go the kind of pictures of mental content advanced by theorists like Stalnaker (1984), Dretske (1988), and Lewis (1994), for instance. (Of course, even the theory of content implicit in Kaplan (1977/1989) is (despite intentions) noncompositional, as Rabern (2013) shows.) What view one prefers here depends on what one expects the notion of mental content to do—what explanatory work it is supposed to perform—and it seems rarely the case that theorists are working with just the same conception of that work. In any case, the claim is not that it is necessary to sign up for one of these particular views of content to proceed. It is enough to see that we can still intelligibly debate the question whether content is compositional in some sense, even after having agreed that linguistic meaning is compositional.

When we separate the notion of content and of semantic value, we clear conceptual space for the possibility that:

The semantic values of declarative sentences are of a uniform type in a manner conducive to compositionality, even though the communicative role of some declarative sentences is such as not to determine truth-conditional propositional content of some traditional sort.

That is, we make room for the possibility that the role of some sentences—stereotypically ‘descriptive’ sentences—in communication may be to express ordinary, world-characterizing truth-conditional propositional contents as a function of their semantic values, along traditional lines, whereas the communicative role of other declarative sentences—normative sentences, perhaps—may not be.

And that idea in turn can be sharpened via the notion of dynamic force. Even if the semantic values of declarative sentences take a uniform shape, there may yet be semantically notable subtypes of declarative sentences, subtypes that correspond to distinctive dynamic forces. A coherent possibility is that sentences we intuitively describe as having normative import, while being of the same semantic type as non-normative sentences, nevertheless form a semantically natural and distinctive class; and moreover that this class is associated with a distinctive kind of conversational update in a way that vindicates the intuitive thought that normative sentences are different in some communicatively important way from straight factual assertion.

## 15.7 Static and dynamic expressivist paths

What does it mean to say that declarative sentences could be of a ‘uniform semantic type’ despite also dividing into ‘semantically distinctive classes’ in such a way as to allow them to correspond to distinctive dynamic forces? This can be—indeed, has been—made precise in a variety of ways. Let me mention two possible paths, without presuming there aren’t others. My objective here is not to lay out very detailed semantic-pragmatic proposals about specific expressions; rather it is to clarify two shapes that detailed proposals could take.<sup>18</sup>

<sup>18</sup> The basic contours of the picture I present here appear in Yalcin (2007, 2011, 2012c). I discuss normative language in particular in Yalcin (2012a,b).

The first trail was in essence blazed by Gibbard (1986, 1990, 2003), though notably he did not employ the notion of dynamic force I am recommending. Let me describe a version of the idea. Assume a textbook intensional semantics of the usual sort—say, in the style of the appendix of Kaplan (1977/1989). But in addition to parameters for context, possible world, and variable assignment, take it that semantic values are relativized also to ‘systems of norms’ in the style of Gibbard (1990). (Or better, to ‘hyperplans’ in the style of Gibbard (2003), and about which more later.) Thus formally speaking, all expressions have as their semantic values functions from this kind of tuple of parameters to extensions. The semantic values of declarative sentences in particular are functions from such tuples to truth-values. In this sense, declarative sentences are of uniform semantic type.

Nevertheless, we can isolate the subclass of sentences whose truth-values are sensitive to the value of the hyperplan parameter. The modeling idea is that the normative sentences will correspond to this class. We can isolate this class just as we can, e.g. semantically isolate the class of open sentences in first-order logic. The situation is analogous. Semantically speaking, the open sentences are (to an adequate first approximation) the ones whose truth-values are sensitive to the variable assignment; the closed sentences are the ones whose truth-values do not vary with the choice of variable assignment.<sup>19</sup> The analogy here is worth underlining. There is no Frege-Geach problem about open sentences—no problem about how it could be that the open and closed sentences of first-order logic, despite their very real semantic differences, can nevertheless intelligibly appear in the same places in a compositional way. So it is, too, with norm-sensitive sentences on Gibbard’s approach. The fact that these sentences are sensitive to value of the norm parameter makes this class ‘semantically distinctive’ in the relevant sense.

Once we acknowledge this semantically distinctive subclass of declarative sentences, we are free to hypothesize that the sentences of this class may have a distinctive dynamic role in conversation—that their dynamic force is not, say, simply to add truth-conditional information to the conversational state along textbook Stalnakerian lines. Their role may be to change a different aspect of the conversational score, or anyway, to change *more than* just the ‘world-describing’ component of the conversational state. To use Gibbard’s way of talking, normative sentences—or anyway, those of a paradigm sort—would serve to update conversational states in respect, not (or not only) of how things are, but also in respect of what is to be done. That is a way of getting at the expressivist idea, voiced in the quotes above, that normative talk is different from ordinary ‘descriptive’ assertion. To spell out this thought in detail, we would want to postulate the relevant component of the conversational state tracking “to be doneness”—perhaps Portnerian To-Do Lists are what is needed; perhaps the plans of Lewis (1979b); perhaps the hyperplans of Gibbard (2003) (explored in Yalcin (2012a); see also Pérez-Carballo and Santorio (2016)); perhaps all of the above; perhaps something else—and articulate how sentences of the target class serve to change that feature of conversational states.

<sup>19</sup> ‘First approximation’ because, of course, an open sentence like  $(Fx \vee \neg Fx)$  may nevertheless be technically insensitive to the value of the assignment function. A more sophisticated definition of ‘sensitive to (assignments, norms, etc.)’ could be given, but the exercise isn’t necessary here.

To an account like this we could add, further, that the relevant kind of conversational state change is often exploited to perform illocutionary acts we could call something like “expressing norms”, acts that correspond to the expression of states of mind that are not “prosaically factual”. (Just as we could say that the sort of conversational state change Stalnaker described is often exploited to perform illocutionary acts we could call “expressing beliefs” or “expressing knowledge”.) Altogether, this path would seem to lead a coherent package of views about normative discourse, a package that could reasonably be called ‘expressivist’.

This obviously is not yet to establish that this kind of path is the right one to take for some particular subfragment of English; that is an empirical matter that needs to be fought out in the usual way. It is just to clarify a coherent possible form a semantics-pragmatics for some language could take, a form that is recognizably expressivist in spirit.

Call this kind of expressivist plan ‘static’, as the details of implementation involve the assumption of a static intensional semantics. A second, alternative implementation would use the resources of dynamic semantics. The meaning of a sentence on the dynamic semantic approach is given by the way it is apt to change the state of the conversation. Formally, the compositional semantic values of all sentences take the form of functions from conversational states into conversational states—context change potentials. The semantic values of declarative sentences on this approach would again be of uniform semantic type in the sense relevant to compositionality. Still, we could isolate interestingly different subclasses of sentences, grouping sentences into characteristically similar context change potentials. Again the basic thought would be that normative sentences invoke a characteristic sort of change to a component of the conversational state, one not merely adding more information to the state about what the world is like. That they induce this kind of change would be something reflected, on the dynamic approach, directly in the semantic values of sentences. And again, the thought would be that the relevant kind of conversational state change is often exploited to perform illocutionary acts we could call something like “expressing norms”.<sup>20</sup>

(The main difference between the static and dynamic approach is that on the static approach, we need some ‘bridge principle’ mapping the normative sentences to their context change potentials, whereas no such principle is needed on the dynamic approach (as it equips sentences with context change potentials directly in the compositional semantics). For example, if we accept Gibbard’s idea that normative sentences are semantically the sort that determine a nontrivial condition on (centered) world-hyperplan pairs, then we’d want a rule, akin to Stalnaker’s assertion rule, telling us how this condition is supposed to serve to change the state of the conversation—a rule that maps the output of the semantics to a context change potential. The simplest version of such a rule would perhaps postulate that the conversational state just is represented by a set of (centered) world-hyperplan pairs, and that update is just intersection.

<sup>20</sup> See Charlow (2015), Starr (2016), Willer (2016, 2017) for developments of expressivist ideas about normative language using tools from dynamic semantics. Veltman’s work on the dynamic evolution of expectation patterns in discourse, in connection with words like ‘normally’, is one important relevant precedent (Veltman 1996).



On that simple kind of picture, normative sentences will be conversationally distinctive, because they will serve to eliminate (centered) world-hyperplan pairs in part as a function of the hyperplan-component.)

Expressivism is often thought of as a special kind of semantic theory. The version of expressivism I am now suggesting is not well-described that way. I am suggesting that on the most plausible development, expressivism is not a kind of semantic theory as model-theoretic truth-conditional semantics is a kind of semantic theory, or as (say) Heim (1982)'s dynamic semantics is a kind of semantic theory. Expressivism is not an alternative to these frameworks; on the contrary, an expressivist view can be developed entirely within the context of such frameworks. The misconception that expressivism must be seen as a special kind of semantic theory stems in part from the tendency to conflate semantic value with content, together with the idea that items of content each determine a 'factualist' truth-condition, understood as fixing a 'way the world might be'. Expressivism will seem radical if one thinks orthodoxy in semantics requires making these assumptions. But to think that is to misunderstand truth-conditional semantics in the familiar model-theoretic style.

Since the basic expressivist idea can be realized in the context of quite different compositional semantic theories, it is not itself well-characterized as a thesis about the shape of a compositional semantic theory. In general, one cannot necessarily read an expressivist view directly off a compositional semantic theory. Better to think of expressivism as a view in pragmatics, or at the semantics-pragmatic interface. It is a kind of view that may be seen as imposing some high-level constraints on semantics. Expressivism about a fragment of language comes in (or doesn't) when we take a certain stand on the relation between the compositional semantic values of the fragment and their dynamic force, and in the relation between dynamic force and the sorts of states of mind they are apt to express in various contexts.

## 15.8 Normative states of mind in an expressivist setting

To make some of the preceding slightly more concrete, let me walk through one conception of the underlying normative states of mind that normative language is, according to the expressivist, in the business of expressing. The story is substantially inspired by Gibbard (2003), though certain aspects will depart from Gibbard's preferred development.

The story begins with a model of normative states of mind. Formally speaking, Gibbard's way of modeling attitudes begins with the abstract, idealized model of attitudes given by Lewis (1979c), in which a state of belief is represented as a set of centered worlds, intuitively the centered worlds compatible with the way the agent takes herself to be situated in the world. Gibbard adopts this much structure to model what it is to have a factual view, a view about how the world is (and about where one is within it). He then adds further structure to model what it is to have normative view, which he takes fundamentally to be a view concerning what to do. To have a view about what to do is to have a plan. To model normative, plan-like states of mind, Gibbard introduces a technical notion, the *hyperplan*. As a single centered world might be used to characterize a state of mind completely opinionated concerning every matter of fact, so a hyperplan can be used to characterize a state of mind

completely opinionated about what is okay to do in any situation. Thus a hyperplan is a maximal contingency plan:

[it] covers any occasion for choice one might conceivably be in, and for each alternative open on such an occasion, to adopt the plan involves either rejecting the alternative or rejecting rejecting it. In other worlds, the plan either forbids an alternative or permits it. (56)

A hyperplan can thus be construed as a mapping from a set of available options to some subset of those options—the options deemed by the hyperplan as permissible.<sup>21</sup>

Equipped with this notion of a hyperplan, Gibbard modifies Lewis's model of belief states. Instead of modeling them as sets of centered worlds, he proposes we model them as sets of centered world-hyperplan pairs. We can say these are models of *plan-laden* states of belief, states of mind that intertwine a view about how things are with a view about what is to be done. One's view about purely descriptive, worldly matters of fact is settled by the centered worlds that one's plan-laden belief state leaves open. But one's normative views—for instance, one's views about what ought to be the case—depend at least in part on the hyperplans that one's plan-laden belief state leaves open. Just as what one believes in the prosaically factual sense is reflected in what is true at every centered world compatible with what one believes, so what one believes about what to do is reflected in what is common to every hyperplan left open by one's state of belief. If I plan to pack, for instance, then every hyperplan left open by my state of belief calls for packing relative to the options I take myself to have.<sup>22</sup>

To believe I ought to pack more or less just is, for Gibbard, to plan to pack. The story is aimed *inter alia* at clarifying the tie between normative thought and motivation, at removing the mystery of why the belief that I ought to pack motivates packing.

Planning states are understood broadly: one can have plans, both about what one is to do in the situation one takes oneself to be in, and also plans about what to do in situations that one does not take oneself to be in. Even if one recognizes that one is not Obama, one can have a view about what to do if faced with Obama's options. This view is a planning attitude in the relevant sense.<sup>23</sup>

We could describe Gibbard as offering a model of the contents of belief states, a model aimed at capturing what is distinctive about normative thought. As we have

<sup>21</sup> Schroeder (2008a) writes that Gibbard “assumes that hyperplanners are always decided either to do A or to not do A, for any action A” (53). Not so: a hyperplan (and thus any hyperplanner whose state is modeled by the hyperplan) may deem both A and  $\neg$  A permissible.

<sup>22</sup> Note that the options one takes oneself to have are fixed by what one believes in the prosaically factual sense. In this way, one's view about what ought to be the case is sensitive to what one takes the facts to be.

<sup>23</sup> More carefully: the relevant sense of ‘plan’ and ‘planning’ here is quasi-technical. To plan in the target sense is to have a take on what is permissible to do in some class of situations. We are understanding “thinking what to do” in the sense of “thinking what is to be done” or “thinking what should be done”. One might have a plan in this sense for a situation (a take on what is permissible in it) and yet still be undecided about what one will in fact do in that situation. Resolving that latter form of indecision is something we could call (following Gibbard 2006) forming a *strategy*. I am inclined to agree with Scanlon (2006), that the connection between planning and normative judgment shouldn't be overstated (and needn't be for the expressivist's purposes). Specifically, there isn't a need to reduce normative thinking to planning if we are going expressivist. There is just a need to model normative thinking as something other than prosaically factual belief, and to be able to specify the characteristic functional role of this sort of state in nonnormative terms.

stressed in earlier pages, a model of content still leaves much open from the point of view of compositional semantics. There are many ways one might try to bring these abstract ideas to bear on the semantics and pragmatics of normative language. But I first want to consider some questions about Gibbard's model of content of a more foundational nature.

Some theorists may have the feeling that the appeal to hyperplans in modeling normative states of mind engenders only an illusion of explanation or understanding. 'Hyperplan' may seem to be a name for a mystery, or a merely formal widget that distracts from the real philosophy. One aspect of the concern may trace to the postulation of a new primitive element. Most theorists who use (centered) possible worlds to model belief content would have them in their ontology anyway—the concept of a possible world is very plausibly intelligible independently from their particular use in modeling content. (Indeed, many theorists, notably Stalnaker (1984) and Lewis (1986, 1994), hope to explain the intentionality of the mental in part by relying on a non-intentional conception of modality.) Hyperplans, on the other hand, seem to arrive on the scene as a new primitive without independent motivation; and it can seem we have no independent grip on what they are that is analogous to the independent grip we have on the notion of a possible world. Gibbard eases us into the notion by an appeal to the idea of a completely decided agent—a hyperplanner. But if hyperplans are explained entirely in terms of normative states of mind—if all that can be said by way of clarifying the notion of a hyperplan is that it approximates the state of mind of a completely decided agent—then we seem to be modeling normative states of mind and their interrelations with the help of...a primitive notion of normative states of mind. This circle is rather tight. It can be hard to see any path here for understanding normative states of mind and their properties as explicable in other terms, as grounded in more basic facts.

Confronted with this kind of request to say more about the idea of plan-laden belief, I am inclined to react as follows. There are two parts to the response.

The first part of the response stresses that we needn't take hyperplans as primitive elements of the model. These objects can be constructed from antecedently available resources (as Gibbard himself observes (2003, 100)): sets and possibilia. We said that a hyperplan is basically a mapping from options to a subset of those options. Let's now take this literally: hyperplans are functions on sets of options. The question then arises what options are. We can model an option as a centered worlds proposition (a set of centered worlds)—the conditions under which the option can be said to be realized. Thus a hyperplan is a function from a set of centered worlds propositions (understood as a set of available options for some possible agent) to some subset thereof. So we can construct hyperplans from independent resources, resources we have a grip on independently of their application in modeling normative states of mind.<sup>24</sup> This does something to demystify hyperplans.<sup>25</sup>

The second part of the response concerns the issue of how philosophically to gloss the formal model. It begins by framing the question:

<sup>24</sup> Even simpler, we could take a hyperplan to be a function from the union of the set of options to a union of some subset of those options. (So I suggested in Yalcin 2012a.) But this leaner description may be inadequate for cases where a plan-laden state seems somehow sensitive to the way the options are distinguished.

<sup>25</sup> Let me add another point of clarification. I am taking it (following Gibbard) that the plans that one's state leaves open will generally interact with/be sensitive to what worlds one's belief state leaves open

In virtue of what is an agent's state of belief well-modeled by a given set of centered world-hyperplan pairs?

This we could call the foundational question about plan-laden content. Any formal model of attitude states faces a foundational question of this sort. It asks what it is about a situated agent that makes it the case that their state of mind is representable by some formal object in the model rather than some other formal object. It may be that some of the concern about the extent to which Gibbard's approach to normative states of mind is adequately explanatory stems from unclarity about what the shape of the answer to this question is supposed to be. If one has no grip at all on what it could be about an agent that makes it the case that their state of mind has one body of plan-laden content rather than another, the account is apt to seem mysterious.

Theorists who agree about the shape of a formal model of the attitudes may nevertheless disagree about the answer to the corresponding foundational question. So we can expect there to be various responses that expressivists who enjoy Gibbardian modeling tools might have. But to sketch one direction in order to give a sense of things, it is natural to consider the corresponding foundational question as it arises for Lewis (1979c), especially since as noted, Gibbard's formal model is an extension of Lewis's.

Again, Lewis (1979c) models a state of belief as a set of centered worlds. We intuitively describe this as the set of centered worlds compatible with where, for all the agent believes, she is. Lewis also puts it like this: to believe that *P* is to *self-ascribe* the property corresponding to the set of worlds *P*. But of course, these intuitive descriptions are not meant to supply deep explanations: they just show how the formal talk is meant to connect with informal belief talk. If, given an agent *A* in *w* modeled with centered worlds belief content *P*, we ask:

What is it about *A* in *w* that makes it the case that *P* is the content of *A*'s belief state in *w*?

(alternatively, sensitive to how one apportioned credence across logical space), and suggesting that it is easier to understand the way they interact if hyperplans are not taken as primitive but built out of (formalized with) possibilities. But I am thinking of the plan structure of an agent's state as additional structure—structure that we can, as theorists, usefully separate out from the set of worlds (or probability space) that corresponds to the agent's doxastic state. So from the point of view of the modeling theory, we are separating the planning structure from the doxastic structure—much as, e.g. the decision theorist wants to separate out a state corresponding to belief and another state corresponding to preference.

Now this might seem confusing when it comes to approaching the semantics of natural language, because I want to allow that certain belief ascriptions (e.g. 'John thinks he ought to pay his taxes') might sometimes place conditions on the plans the subject's state leaves open (in addition to the worlds their doxastic state leaves open). So there is not a straight line from what the theorist calls a "doxastic state", which is to do only with modeling what the world is like according to the agent, and object-language belief ascription, which can mix ascription of factual and plan-like content. But there is no problem here. You might compare this to the semantics of 'wants'. As theorists, we find it fruitful to model preference as some sort of ordering over possible states of the world. But when it comes to the semantics of 'wants', it may turn out that 'wants'-ascriptions also place conditions on the subject's state of belief (as e.g. Heim (1992) suggests, drawing *inter alia* on Stalnaker (1984)). So ordinary talk of what an agent wants admixes features of what they prefer and what they believe. Still, again as theorists, there's a sense in which the agent's preference ordering alone gets at a natural psychological joint, and models the agent's desires in a pure but abstract sense.

As for whether to call the additional planning structure 'cognitive', the answer is 'no' if by 'cognitive' we just mean *doxastic*. But if 'cognitive' means something like *intentional mental state*, then certainly planning states are cognitive (as are states of preference). (Thanks to Mahrad Almotahari for discussion.)

Lewis has an answer. The basic contours of his answer are given in Lewis (1984, 1986, 1994). Not really needing a full review (and lacking space for it anyway), I stick to a crude highlight reel. Like many theorists, Lewis takes it that the contents of mental states are supposed to be (at least) causal-explanatory properties of those states. Particular hypotheses about the belief and desire contents of an agent generate *ceteris paribus* predictions about how the agent will be disposed to act in various circumstances. Belief and desire states are, Lewis thinks, constitutively rational. These states are the occupants of certain functional roles, and it is just part of the functional roles associated with belief and desire that beliefs and desires tend to cause behavior that serves the subject's desires according to her beliefs. Roughly, belief, desire, and action are constitutively related at least as follows:

If agent *A* is in a belief state with the centered content *B* and in a desire state with centered content *D*, then *A* is disposed to act in ways that would tend to bring it about that he is located within *D*, were it the case that he occupied a centered world within *B*.

Belief and desire are causally efficacious inner states whose content is constitutively rational in (at least) this way. Roughly, the full belief-desire state of an agent is the one most apt to produce the agent's dispositions to act compatible with the above, and which otherwise maximizes the extent to which the agent's belief content is *eligible*—that is, sensitive to reality's objective structure.<sup>26</sup>

That is Lewis's approach to grounding facts about contentful states of mind. This story is responsive to the foundational question raised by the centered worlds modeling framework. It does much to clarify the subject matter of that model—the phenomena being modeled by it. In a certain sense, it offers a way of interpreting the model.

I don't have the aim of arguing that Lewis's particular foundational picture is correct. Rather, what I want to suggest is that the sort of expressivism I have described can approach the foundational question framed above along the same pattern. If we can see this expressivist and the Lewisian as basically on par—as facing similar foundational questions, and as having similar styles of answers at their disposal—that will suggest that no qualitatively different foundational or 'metasemantic' challenges are faced by the expressivist *per se*; and it will help to further demystify the role of hyperplans in the expressivist's model.

So how can the expressivist who models with Gibbard's tools approach the foundational question framed above in the same basic way as the Lewisian? She can do this by articulating the constitutive functional interconnections between belief (or credence), desire (or preference), and her postulated planning states. Moving beyond the familiar belief-desire framework assumed by Lewis, she can suppose that rational action centrally involves also states of planning. We can expect the detailed functional

<sup>26</sup> The appeal to eligibility is motivated by the fact that without it, many intuitively incorrect assignments of belief-desire content would nevertheless preserve constitutive rationality and make the correct predictions about action. The basic contours of that worry go back to Putnam (1980). See Stalnaker (1984) for a different approach. See also Stalnaker (2004) for a critique of Lewis's particular way of appealing to eligibility.

interconnections between these three states to be elaborate, but it seems safe to hypothesize that the functional interconnections will include at least something like the following. Agents are disposed to act in ways which would conform with their plans, in centered worlds with respect to which their (purely factual) belief content is true. Where such plans leave several options open, agents tend to elect those options which would serve best to satisfy their preferences. And where agents find themselves in unplanned-for situations, we appeal only to belief and desire. That is, we say that such agents will be disposed to act in ways that would tend to satisfy their desires, in centered worlds where their beliefs are true—leaving intention out of it.

Of course, that's brief—we should try to say more. But then, so should the Lewisian. The point is that Gibbard's kind of model of content need not, if explained in this way, be regarded as more mysterious than the sort of (unmysterious) accounts favored by theorists in the possible worlds tradition upon which he is building, such as those of Lewis (1994) or Stalnaker (1984). Moreover, as I read it, it seems to me much of Gibbard (2003)—not to mention important earlier work by Bratman (1987)—is anyway concerned to draw out and explore the rich functional interconnections between belief, desire, and planning—so that this is hardly work left entirely undone.

Gibbard himself embraces a very different way of philosophically glossing his own model. Theorists in the orthodox tradition of possible worlds modeling, like Lewis or Stalnaker, would say that to believe that grass is green is to be in a state which rules out possible worlds wherein grass is not green, and they would hold that these possible worlds that the state rules in or out are explicable independently of intentional mental states. Gibbard, however, prefers to explain the “possibilities” that mental states rule out *as themselves mental states*. (He has this view quite apart from his proposal to model in terms of hyperplans—this is how he would want to think about an ordinary, hyperplan-free possible worlds model of belief.) Gibbard will agree that to think grass is green is to “rule out a possibility”, but fundamentally he will explain this state as the state of ruling out another mental state, the state of rejecting grass is green—the mental state of *rejecting that grass is green* is the “possibility” ruled out. Similarly, he would describe the state of believing that grass is not green as “disagreeing with believing” that grass is green—as rejecting believing grass is green. The centered worlds of the model are interpreted by Gibbard as maximally opinionated states of (factual) belief—they are not, as Lewis or Stalnaker would have it, maximally specific ways things might have been, understanding the relevant modality as fundamentally non-mental. Likewise, he glosses the hyperplans of his model as maximal states of decision.

This is apt to look like a pretty tight circle: it is a model of mental states whose basic resources for modeling mental states include...mental states. Out is the idea of characterizing propositional attitudes as relations to contents, if the latter are understood in the traditional way as some sort of mind-independent abstracta—sets of possibilities, for instance, as Lewis and Stalnaker would have it. We don't arrive on this picture at a conception of mental content giving us a handle on it in other terms. This seems to leave it mysterious, for instance, what makes it the case, when it is the case, that one content is incompatible with another. For example: the state *believing grass is green* and the state *believing grass isn't green* “disagree” with each other; they are in logical tension. In virtue of what?—Not, says Gibbard, in virtue

of their having incompatible contents, if content elucidated in orthodox fashion, via a non-intentional notion of possibility. What, then? Gibbard says he has no further explanation of such disagreement facts; he takes them as primitive.

Gibbard argues that this is not a disadvantage, however:

Proceeding this way might seem to be philosophical theft. The scheme amounts just to helping ourselves to the notion of disagreeing with a piece of content, be it a plan or a belief. A negation, we say, is what one accepts when one disagrees—and this explains negation. Now I wish, of course, that I could offer a deeper explanation of disagreement and negation. Expressivists like me, though, are not alone in such a plight. Orthodoxy starts with substantial, unexplained truth, eschewing any minimalist explanation of truth. I start with agreeing and disagreeing with pieces of content, some of which are plans. It's a thieving world, and I'm no worse than the others.

(Gibbard 2003, 74)

Here I want to depart seriously from Gibbard. It seems to me a mistake to suggest that orthodoxy starts with “substantial unexplained truth”, if that is meant to imply that it involves brute appeals to intentional relations between mental states and their contents, with no sense of a pathway for reduction or further clarification. We can model states of belief as having plan-laden content. Items of plan-laden content stand in various familiar logical relations. Inconsistency between states of mind traces to their inconsistent content. States have content in virtue of their functional interconnections to each other and in virtue of subsidiary requirements on prosaically factual belief (such as that its content be suitably eligible along the lines of Lewis, as sketched above; or that it be a state that normally carries information, as suggested by, e.g. Dretske (1981), Stalnaker (1984); or perhaps some two-dimensional admixture of these; or perhaps something else—debates continue). There is an array of choice-points about the details, familiar from much of the philosophical work on intentionality in the 1980s and early 90s, but it is hard to deny that there are well-trodden paths of analysis in this vein. Expressivists like Gibbard—the ones who embrace, not just his abstract formal model of normative states, but also his distinctive metatheoretic gloss on the model—are, I think, relevantly alone in their plight, and are at an explanatory disadvantage compared to orthodox rivals, like the account of this essay.

Gibbard seems to suggest in places that the expressivism of his account chiefly resides in his preferred metatheoretical gloss on the model—so that to reject that gloss just is to reject expressivism. He seems to be identifying expressivism with something like what I earlier called ‘compositional attitude expressivism’, giving a version of it wedded to deflationary views about truth and meaning. But I think this is the wrong way of styling expressivism. Expressivists do “explain in terms of mental states”, but we need not take this in the direction of Gibbard’s style of metatheory.

What is distinctive of expressivism, I suggest, is the way it exploits the strategy of psychological ascent. To go expressivist about  $\phi$ , you first reject the question, “What is the world like when  $\phi$  is the case?” You replace it with the question, “What is the state of mind of accepting  $\phi$  like?” You answer this question in such a way that the state of mind is understood as not tantamount to ordinary factual belief that something is the case. You then approach the target discourse from this perspective: you seek a way to elucidate the semantics and pragmatics of  $\phi$  consistent with the idea that accepting  $\phi$  is being in this not-fully-factual state of mind.

The story I have advanced so far is expressivist in this sense. Normative thought is styled as not fully factual in character. To believe that one ought to pack is not to be belief-related to some possible worlds truth-condition, to a way things might have been. It isn't merely to represent the world, or one's position in the world, as being a certain way rather than other. It is to have a view about what is to be done, in the sense formalized and functionally explicated above. What still remains is to connect this model of normative states of mind to language—to matters of meaning and communication. But on natural ways of forging those connections, it is not hard to see how to end up with a view vindicating the traditional expressivist thought that normative talk is not purely factual in character.

### 15.9 Normative language in an expressivist setting

In a big picture way, we have already said how that story can go. Let me restate. To have a model of the state of the conversation that can comport nicely with the plan-laden conception of belief, we can make conversational states themselves plan-laden. We can, for example, model conversational states as sets of centered world-hyperplan pairs.<sup>27</sup> We then hold that normative talk is in the business of eliminating such pairs—or more broadly, of changing the state—in ways that are sensitive to the plan component. This characteristic difference in the way that normative sentences would be apt to change the state of the conversation is a difference in dynamic force.

Where does talk of “expressing a normative state of mind” come in? Here again it pays to be mindful of the distinction between illocutionary and dynamic force. In changing the conversational state in respect of the plans it leaves open, one is often expressing one's normative beliefs—the latter understood expressivistically as above. When we change the conversational state in this kind of way, is it *always* to express the normative view one in fact endorses? No—no more than ordinary factual conversational update is always a matter of expressing one's true state of belief. We stressed earlier that the dynamic notion of force prescinds largely from one's distal objectives in communication. Often we do change the state of the conversation in something like the way Stalnaker taught, and often (not always) we do that with the mutually understood aim of transferring belief, knowledge, a view about how the world really is, and the like. When those subsidiary elements are in place—aspects that come into focus when we approach the performance as an illocutionary act, locating it in a particular kind of space of human interests and objectives—we might then sensibly identify the speech act move made as an “expression of belief”. We need the same kind of subtlety when it comes to talk of “expressing one's normative view” on an expressivist approach.

It could be right to say that “expressing one's normative view” is what is happening in most ordinary cases we'll want to model dynamically in terms of elimination

<sup>27</sup> Or, as I prefer, a pair of a set of centered worlds and a set of hyperplans; see Yalcin (2012a). Some orthogonal subtleties arise here in connection with using centered worlds in a model of the conversational state; see, e.g. Egan (2007), Stalnaker (2008). It may be preferable to generalize, using multi-centered worlds in place of centered worlds. For further discussion of that approach, see Stalnaker (2008, 2014), Ninan (2012).



of hyperplans. When it is one's normative state that is prompting one's utterance, when the condition on hyperplans determined by the speech act accords with the speaker's actual normative state, and when it is mutually recognized that the speaker is attempting to engender coordination in respect of normative states, then it will seem natural to call what is happening "expressing one's normative view". When all these background factors (and perhaps more) are in place, the expressivist might want to further claim that normative sentences so deployed characteristically involve a distinctive illocutionary force—"norm expression", say. I myself am not sure what value it would add to make this kind of declaration, since I am not sure there is much of an interesting general theory of illocutionary force to be given;<sup>28</sup> but we needn't get into that. The point is to separate the vaguely illocutionary idea of "norm expression" from the idea that some fragments of language interact especially with the planning aspect of the conversational state. These will often go together, according to the sort of expressivist described, but they are not the same thing.

We should, in particular, be alive to other ways we might find ourselves messing with plan-laden conversational states. For example, just as the worlds compatible with a context set might be taken, in context, to be characterizing a fictional universe rather than the actual world, so the plans a conversational state leaves open might, in context, be mutually recognized as characterizing what is so according to some particular plan that no party to the discourse in fact endorses (cf. Lewis 1979b). That is logically possible, anyway. If there are such cases, we would want to describe the states of mind expressed as descriptive or factual in character, despite their similarity in dynamic respects to cases which clearly do involve the expression of normative states. There is no problem here for the expressivist—just distinctions to avoid tripping upon in imposing the vague word "express" upon an otherwise clear theory.

It should be acknowledged that the point we have reached is far in important respects from some of the early expressivists cited above. Normative talk is not much like exclamation (pace Barnes 1934); nor does it serve to express preference (pace Carnap 1935 and Russell 1935) or feeling (pace Ayer 1936). But in accord with these authors, and in the spirit of the quote from Austin above, the view is that normative talk serves express states of mind that are not straightforwardly factual in character.

## 15.10 Empirical plausibility

If the preceding seems to leave the impression that expressivism about normative discourse faces only smooth sailing, it is time to bring the bad news. Metaethical expressivism is, *inter alia*, supposed to be a thesis about the meaning of some fragment of natural language. About this sort of thesis we can separate two issues:

- (i) LOGICAL POSSIBILITY Is the thesis even in principle compatible with a compositional semantics for some elementary possible language? Can we even make sense of a communication system that works along expressivist lines?

<sup>28</sup> There is the worry that such a theory verges on a "theory of everything" of the sort derided by Chomsky (2000). See footnote 10.

- (ii) **EMPIRICAL PLAUSIBILITY** Can the thesis be well-motivated for a fragment of some actual natural language?

I have largely been concerned with describing affirmative answers to the **LOGICAL POSSIBILITY** questions, seeing these questions as prior. I have not said anything very substantive about **EMPIRICAL PLAUSIBILITY**. But of course, our expressivist means to say that a fragment of natural language actually works in the way described—it interacts with the plan structure of the conversational state, and is apt for giving voice to plan-laden states of mind. This places some abstract constraints on the semantics of natural language—in particular, normative expressions need to be given semantic values which interact with hyperplans in the right ways, generating sentences whose context change potentials are adequate for the expression and transmission of plan-laden states of mind. (Normative predicates would seem to need plan-sensitive extensions; deontic modal operators would seem to need to involve quantification over hyperplans, etc.) It is perfectly possible for there to be such semantic values. But is there empirical motivation for the thesis that sentences of natural language have such semantic values?

It is interesting that Gibbard (2003) never attempts to put hyperplans to non-trivial compositional semantic work—for instance, by articulating the compositional semantics of deontic modals, or the attitude verbs ‘decides’ or ‘believes’, etc., by appeal to hyperplans in the semantic metalanguage. This is the sort of thing that would be required to motivate the added structure from semantics-internal considerations.<sup>29</sup> A natural place to start might be with sentences like these:

- (1) John thinks that he ought to pack.

The first move the expressivist makes is to psychologically ascend: her story begins with a model of some mental states. It is natural to start here because our expressivist offers special truth-conditions for this sentence: on her theory, its truth turns on the hyperplans that John’s state rules in and out. But to give truth-conditions is not yet to give a compositional semantics. The task for the expressivist is to show how to determine these truth-conditions via the compositional semantics of the sentence using the advertised hyperplans, and in a way that meshes with what is already known about the semantics of the constituent expressions of the sentence. (See Yalcin (2012a) for one start at this task.) A similar task awaits the expressivist for normative predicates in general. Much work remains to be done here. The expressivist approach needs to prove itself in the details in natural language semantics and pragmatics. But recent work developing expressivist ideas using tools from formal pragmatics suggests some cause for optimism—or anyway, an open mind.<sup>30</sup>

Further, theorizing in semantics and pragmatics with plans may teach us more about what plans must be like, what constraints they are subject to. Assigning plans semantic and pragmatic work to do is a way of constraining a theory of them. Seeking reflective equilibrium between a theory of normative language and a theory of normative thinking, we can make progress on both.

<sup>29</sup> This perhaps only highlights how different the present conception of expressivism is from Gibbard’s.

<sup>30</sup> Besides the many works already cited in this vein, see also Santorio (2016), MacFarlane (2016).

## References

- Austin, John Langshaw. Performative utterances. In J. O. Urmson and G. J. Warnock, editors, *Philosophical Papers*, pages 220–39. Oxford University Press, Oxford, 1961.
- Austin, John Langshaw. *How To Do Things With Words*. Harvard University Press, Cambridge, MA, 1962.
- Ayer, Alfred J. *Language, Truth and Logic*. Dover, New York, NY, 2nd edition, 1936. Second edition, 1946.
- Bach, Kent and Robert Harnish. *Linguistic Communication and Speech Acts*. MIT Press, Cambridge, MA, 1979.
- Barnes, W. H. F. A suggestion about value. *Analysis*, 1(3):45–6, 1934.
- Belnap, Nuel and Thomas Steel. *The Logic of Questions and Answers*. Yale University Press, New Haven, 1976.
- Blackburn, Simon. *Essays in Quasi-Realism*. Oxford University Press, 1993.
- Blackburn, Simon. *Ruling Passions*. Oxford University Press, 1998.
- Bratman, Michael. *Intention, Plans, and Practical Reason*. MIT Press, Cambridge, 1987.
- Burge, Tyler. Sinning against Frege. *The Philosophical Review* 88(3): 398–432, 1979.
- Carnap, Rudolf. *Philosophy and Logical Syntax*. Kegan Paul, London, 1935.
- Charlow, Nate. The problem with the Frege-Geach problem. *Philosophical Studies* 167(3): 635–65, 2014.
- Charlow, Nate. Prospects for an expressivist theory of meaning. *Philosophers' Imprint* 15(23): 1–43, 2015.
- Chomsky, Noam. *New Horizons in the Study of Language and Mind*. Cambridge University Press, 2000.
- Dreier, James. Negation for expressivists: A collection of problems with a suggestion for their solution. *Oxford Studies in Metaethics* 1: 217–33, 2006.
- Dretske, Fred. *Knowledge and the Flow of Information*. MIT Press, Cambridge, MA, 1981.
- Dretske, Fred. *Explaining Behavior: Reasons in a World of Causes*. MIT Press, Cambridge, MA, 1988.
- Dummett, Michael. *Frege: Philosophy of Language*. Harvard University Press, Cambridge, MA, 2nd edition, 1973.
- Dummett, Michael. *The Logical Basis of Metaphysics*. Harvard University Press, Cambridge, MA, 1993.
- Egan, Andy. Epistemic modals, relativism and assertion. *Philosophical Studies* 133(1): 1–22, 2007.
- Gazdar, Gerald. *Pragmatics: Implicature, Presupposition and Logical Form*. Academic Press, New York, 1979.
- Geach, Peter. Assertion. *Philosophical Review* 74(4): 449–65, 1965.
- Gibbard, Allan. An expressivistic theory of normative discourse. *Ethics* 96(3): 472–85, 1986.
- Gibbard, Allan. *Wise Choices, Apt Feelings*. Harvard University Press, Cambridge, MA, 1990.
- Gibbard, Allan. *Thinking How to Live*. Harvard University Press, Cambridge, MA, 2003.
- Gibbard, Allan. Reply to critics. *Philosophy and Phenomenological Research* 72(3): 729–44, 2006.
- Green, Mitchell. Speech acts. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2015 edition, 2015.
- Grice, H. P. Meaning. *Philosophical Review* 66: 377–88, 1957.
- Groenendijk, Jeroen and Martin Stokhof. *Studies on the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, University of Amsterdam, 1984.
- Hamblin, Charles L. Questions. *Australasian Journal of Philosophy* 36(3): 159–68, 1958.

- Hamblin, Charles L. Mathematical models of dialogue. *Theoria* 37(2): 130–55, 1971. ISSN 1755–2567.
- Hamblin, Charles L. Questions in Montague English. *Foundations of Language*, 10(1): 41–53, 1973.
- Hare, R. M. *The Language of Morals*. Oxford University Press, New York, 1952.
- Heim, Irene. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts, 1982.
- Heim, Irene. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3): 183–221, 1992.
- Hulstijn, Joris. Structured information states: raising and resolving issues. In A. Benz and G. Jäger, editors, *Proceedings of MunDial97*, University of Munich, 1997.
- Kamp, Hans. A theory of truth and semantic representation. In Jeroen A. Groenendijk, Theo Janssen, and Martin Stokhof, editors, *Formal Methods in the Study of Language*, pages 277–322. Mathematisch Centrum, University of Amsterdam, Amsterdam, 1981.
- Kaplan, David. Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes from Kaplan*, pages 481–563. Oxford University Press, Oxford, 1977/1989.
- Karttunen, Lauri. Discourse referents. In *Proceedings of the 1969 Conference on Computational Linguistics*, pages 1–38. Association for Computational Linguistics, 1969.
- Karttunen, Lauri. Presupposition and linguistic context. *Theoretical Linguistics* 1(1–3): 181–94, 1974. ISSN 0301–4428.
- Karttunen, Lauri. Syntax and semantics of questions. *Linguistics and Philosophy* 1: 607–53, 1977.
- King, Jeffrey C. Tense, modality, and semantic values. *Philosophical Perspectives* 17: 195–245, 2003.
- Kölbel, Max. Conversational score, assertion, and testimony. In Jessica Brown and Herman Cappelen, editors, *Assertion: New Philosophical Essays*, pages 49–78. Oxford University Press, Oxford, 2011.
- Lewis, David K. A problem about permission. In *Essays in Honour of Jaakko Hintikka*, pages 163–75. Reidel, Dordrecht, 1979a.
- Lewis, David K. Scorekeeping in a language game. *Journal of Philosophical Logic* 8 (1): 339–59, 1979b.
- Lewis, David K. Attitudes de dicto and de se. *Philosophical Review* 88(4): 513–43, 1979c.
- Lewis, David K. Index, context, and content. In S. Kanger and S. Ohman, editors, *Philosophy and Grammar*, pages 79–100. Reidel, Dordrecht, 1980.
- Lewis, David K. Putnam's paradox. *Australasian Journal of Philosophy* 62: 221–36, 1984.
- Lewis, David K. *On The Plurality of Worlds*. Blackwell, Malden, MA, 1986.
- Lewis, David K. Relevant implication. *Theoria* 54(3): 161–74, 1988a.
- Lewis, David K. Statements partly about observation. *Philosophical Papers* 17(1): 1–31, 1988b.
- Lewis, David K. Reduction of mind. In S. Guttenplan, editor, *A Companion to the Philosophy of Mind*, pages 412–31. Blackwell, Oxford, 1994.
- MacFarlane, John. Vagueness as indecision. *Proceedings of the Aristotelian Society* 90: 255–83, 2016.
- Ninan, Dilip. Two puzzles about deontic necessity. In Bernard Nickel, Seth Yalcin, Jon Gajewski, Valentine Hacquard, editors, *New Work on Modality*, volume 51, pages 149–78. MIT Working Papers in Linguistics, 2005.
- Ninan, Dilip. Semantics and the objects of assertion. *Linguistics and Philosophy* 33 (5): 355–80, 2010.
- Ninan, Dilip. Counterfactual attitudes and multi-centered worlds. *Semantics and Pragmatics* 5(5): 1–57, 2012.
- Ogden, Charles Kay and Ivor Armstrong Richards. *The Meaning of Meaning*. Kegan Paul London, 1923.

- Pérez-Carballo, Alejandro and Paolo Santorio. Communication for expressivists. *Ethics* 126(3): 607–35, 2016.
- Portner, Paul. The semantics of imperatives within a theory of clause types. In Robert B. Young, editor, *Proceedings of SALT XIV*, pages 235–52, Northwestern University, 2004. Linguistics Society of America.
- Portner, Paul. Imperatives and modals. *Natural Language Semantics* 15(4): 351–83, 2007.
- Portner, Paul. Commitment to priorities. In Daniel Fogal, Daniel Harris, and Matt Moss, editors, *New Work on Speech Acts*. Oxford University Press, this volume forthcoming.
- Putnam, Hilary. Models and reality. *The Journal of Symbolic Logic* 45(03): 464–82, 1980.
- Rabern, Brian. Against the identification of assertoric content with compositional value. *Synthese* 189(1): 75–96, 2012a.
- Rabern, Brian. *Monsters and Communication: The Semantics of Contextual Shifting and Sensitivity*. PhD thesis, Australia National University, Canberra, 2012b.
- Rabern, Brian. Monsters in Kaplan's logic of demonstratives. *Philosophical Studies* 164(2): 393–404, 2013.
- Ridge, Michael. *Impassioned Belief*. Oxford University Press, Oxford, 2014.
- Roberts, Craige. Information structure in discourse: Towards an integrated formal theory of pragmatics. In Jae-Hak Yoon and Andreas Kathol, editors, *OSU Working Papers in Linguistics*, volume 49, pages 91–136. The Ohio State University, Department of Linguistics, 1996.
- Roberts, Craige. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5(6): 1–69, 2012.
- Rosen, Gideon. Blackburn's *Essays in Quasi-Realism*. *Noûs* 32(3): 386–405, 1998.
- Rothschild, Daniel and Seth Yalcin. On the dynamics of conversation. *Noûs*, Online first version. DOI: 10.1111/nous.12121:1–25, 2015.
- Rothschild, Daniel and Seth Yalcin. Three notions of dynamicness in language. *Linguistics & Philosophy*. Online first version available at DOI: 10.1007/s10988-016-9188-1:1–23, 2016.
- Russell, Bertrand. *Religion and Science*. Oxford University Press, Oxford, 1935.
- Santorio, Paolo. Nonfactual know-how and the boundaries of semantics. *Philosophical Review* 125(1): 35–82, 2016.
- Scanlon, T. M. Reasons and decisions. *Philosophy and Phenomenological Research*, LXXII(3), May 2006.
- Schroeder, Mark. *Being For: Evaluating the Semantic Program of Expressivism: Evaluating the Semantic Program of Expressivism*. Oxford University Press, Oxford, 2008a.
- Schroeder, Mark. How expressivists can and should solve their problem with negation. *Noûs* 42(4): 573–99, 2008b.
- Schroeder, Mark. What is the Frege-Geach problem? *Philosophy Compass* 3(4): 703–20, 2008c.
- Searle, John R.. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, 1969.
- Searle, John R. and Daniel Vanderveken. *Foundations of Illocutionary Logic*. Cambridge University Press, Cambridge, 1985.
- Stalnaker, Robert. Pragmatic presuppositions. In Milton K. Munitz and Peter Unger, editors, *Semantics and Philosophy*, 197–213. New York University Press, 1974.
- Stalnaker, Robert. Assertion. In Peter Cole, editor, *Syntax and Semantics 9: Pragmatics*, pages 315–32. Academic Press, New York, 1978.
- Stalnaker, Robert. Indexical belief. *Synthese* 49(1): 129–51, 1981.
- Stalnaker, Robert. *Inquiry*. MIT Press, Cambridge, MA, 1984.
- Stalnaker, Robert. Possible worlds and situations. *Journal of Philosophical Logic* 15(1): 109–23, 1986.
- Stalnaker, Robert. Lewis on intentionality. *Australasian Journal of Philosophy* 81(1): 119–212, 2004.

- Stalnaker, Robert. *Our Knowledge of the Internal World*. Oxford University Press, Oxford, 2008.
- Stalnaker, Robert. *Context*. Oxford University Press, Oxford, 2014.
- Stalnaker, Robert. Dynamic pragmatics. In Daniel Fogal, Daniel Harris, and Matt Moss, editors, *New Work on Speech Acts*. Oxford University Press, this volume forthcoming.
- Stanley, Jason. Rigidity and content. In Richard G. Heck Jr, editor, *Language, Thought, and Logic: Essays in honor of Michael Dummett*, pages 131–56. Oxford University Press, Oxford, 1997a.
- Stanley, Jason. Names and rigid designation. In B. Hale and C. Wright, editors, *A Companion to the Philosophy of Language*, pages 555–85. Blackwell, Oxford, 1997b.
- Starr, William. Dynamic expressivism about deontic modality. In Nate Charlow and Matthew Chrisman, editors, *Deontic Modality*, pages 355–94. Oxford University Press, 2016.
- Stevenson, Charles Leslie. The emotive meaning of ethical terms. *Mind* 46(181): 14–31, 1937.
- Strawson, Peter F. Intention and convention in speech acts. *The Philosophical Review* 73(4): 439–60, 1964.
- Unwin, N. Norms and negation: A problem for gibbard's logic. *The Philosophical Quarterly* 51(202): 60–75, 2001.
- Veltman, Frank. Defaults in update semantics. *Journal of Philosophical Logic* 25(3): 221–61, 1996.
- Wedgwood, Ralph. *The Nature of Normativity*. Oxford University Press, 2007.
- Willer, Malte. Dynamic foundations for deontic logic. In Nate Charlow and Matthew Chrisman, editors, *Deontic Modality*, pages 324–54. Oxford University Press, 2016.
- Willer, Malte. Advice for noncognitivists. *Pacific Philosophical Quarterly* 98: 174–207, 2017.
- Williamson, Timothy. Knowing and asserting. *Philosophical Review* 105(4): 489–523, 1996.
- Williamson, Timothy. *Knowledge and its Limits*. Oxford University Press, Oxford, 2000.
- Woods, Jack. The Frege-Geach problem. In Tristram McPherson and David Plunkett, editors, *Routledge Handbook of Metaethics*, pages 226–42. Routledge, 2017.
- Yalcin, Seth. Epistemic modals. *Mind* 116(464): 983–1026, 2007.
- Yalcin, Seth. Nonfactualism about epistemic modality. In Andy Egan and Brian Weatherson, editors, *Epistemic Modality*, pages 295–332. Oxford University Press, Oxford, 2011.
- Yalcin, Seth. Bayesian expressivism. *Proceedings of the Aristotelian Society*, CXII(2): 123–60, 2012a.
- Yalcin, Seth. Comments on Mark Schroeder, *Being For*. Delivered at American Philosophical Association Central Division Meeting, 2012b.
- Yalcin, Seth. Context probabilism. In Maria Aloni, Vadim Kimmelman, Floris Roelofsen, Galit Weidman Sassoon, Katrin Schulz, and Matthijs Westera, editors, *Proceedings of the Eighteenth Amsterdam Colloquium*, volume 7218 of *Lecture Notes in Computer Science*, pages 12–21. Springer, 2012c.
- Yalcin, Seth. Semantics and metasemantics in the context of generative grammar. In Alexis Burgess and Brett Sherman, editors, *Metasemantics*, pages 17–54. Oxford University Press, Oxford, 2014.
- Yalcin, Seth. Quantifying in from a Fregean perspective. *Philosophical Review* 124(2): 207–53, 2015.
- Yalcin, Seth. Semantics as model-based science. In Derek Ball and Brian Rabern, editors, *The Science of Meaning*. Oxford University Press, forthcoming.
- Yli-Vakkuri, Juhani. Propositions and compositionality. *Philosophical Perspectives* 27(1): 526–63, 2013.

