Communication for Expressivists

Alejandro Pérez Carballo and Paolo Santorio

1 Introduction

According to *metaethical expressivists*, moral statements do not aim at describing the world. Rather, the function of moral talk is to express nonrepresentational mental states.

Expressivists face the challenge of giving an adequate semantics for our moral language that is compatible with these claims. The jury is still out on whether any such semantic theory can be made to work. In this paper we set this question aside. Our focus is on a related, but different question, one about the expressivist's pragmatics: how can the expressivist account for the assertion and the uptake of normative claims?

Suppose that Anne and Brad are discussing the permissibility of tax evasion. Anne utters 'Tax evasion is wrong'. She does this because she has an intention to produce in Brad a certain attitude about tax evasion. Upon hearing Anne's utterance, Brad recognizes her intention and acts accordingly: either he will come to have the relevant attitude or he will challenge Anne's claim. All this has the semblance of a rational and purposeful activity. Can the expressivist account for this intuition?

To see that the expressivist has at least a *prima facie* problem, consider how an analogous question is answered for the case of uncontroversially descriptive discourse. Suppose that Anne asserts that a common acquaintance, Tom, evaded taxes last year, and Brad comes to believe the corresponding proposition. How can we explain Anne's assertion and Brad's uptake of the relevant information? Answers may vary, but they will share a common core. Anne makes her assertion because she thinks this is a good way of imparting some relevant knowledge to Brad. Brad comes to believe what Anne says because it's a way of promoting his goal of learning facts about the world.

This simple explanation crucially relies on the appeal to notions like truth, knowledge, and the actual world. But it is part of the expressivist's picture that, at least in an important sense, normative statements are not truth-apt and don't convey information about the world.

Some expressivists suggest that we engage in conversation about normative matters to achieve a kind of coordination. This answer has the advantage of applying to both the normative and the non-normative cases. Speakers make and uptake claims because

they want their attitudes to align, whether those attitudes are beliefs or attitudes of other kind.

This might be a step in the right direction, but it can't be the end of the story. If coordination was all we wanted out of conversation, our communicative practices would look very different from the way they do. Flipping a coin, for example, is a perfectly acceptable way to coordinate. But we don't solve normative disagreements by flipping coins or via other similar coordination strategies. Hence speakers are not *merely* looking to coordinate when engaging in normative talk. Something more than mere coordination must be at stake. For the case of communication about non-normative matters, truth can be again invoked to the rescue. Speakers want to coordinate by having their beliefs align with what things are actually like. But, again, the expressivist cannot appeal to truth in her account of normative talk. So, once more: how can she produce a plausible account of what speakers do in communication?

This is the question we take up in this paper. Our answer starts from some basic assumptions about the rationality of speakers and about our practice of assertion. We show that, on these assumptions, it follows that the communication of normative claims mirrors very closely the communication of non-normative claims. More specifically, we argue for the following:

> In any conversation where certain minimal assumptions are satisfied, it is presupposed that there is a (unique) normative standard on which the participants' attitudes ought to converge.

The *ought* here is a practical one. It concerns the rational requirements that apply to agents, in view of their goal of taking part in communication: it is the familiar *ought* capturing the normative force of Grice's Cooperative Principle. Our claim, then, is this: to successfully engage in communication, speakers have to presuppose that they're subject to a certain basic requirement that applies to them in virtue of their being communicators. This requirement states that their normative attitudes should converge on a joint normative standard. We'll explain the precise content of this requirement, together with the 'minimal assumptions', the notion of a presupposition, and the nature of the attitudes at stake, in the course of our discussion.

Our aim here is not to provide a full account of communication on behalf of the expressivist. Indeed, we are going to take as given certain basic facts about assertion that an expressivist will need to explain further. Instead, our goal in this paper is twofold. First, we will raise a problem for expressivist theories of moral discourse that, to our knowledge, has not been explicitly formulated before. Second, we will provide the beginning of a solution to this problem on behalf of the expressivist.

2 Background: expressivism and communication

2.1 Expressivism

Throughout the paper, we take expressivism about normative discourse as our running example. Ultimately, we would want our conclusions to apply to expressivism about a large variety of discourses. But different varieties of expressivism call for tweaks and extensions of our claims; hence, for the purposes of this paper, we set them aside.[1]

We start from a minimal conception of expressivism about normative discourse, consisting of a negative and a positive claim:[2]

Minimal Expressivism

a. normative claims are not apt to describing, stating, or reporting facts;
b. normative claims express a non-cognitive (nonrepresentational) attitude of some sort.

Of course, this minimal characterization needs to be fleshed out to be turned into a proper theory. One needs to say more about, first, the nature of the non-cognitive attitudes in play and their role in a general philosophy of mind; and second, about the expressivist's semantics for normative discourse.

For present purposes, we assume the version of expressivism developed by Allan Gibbard in his (1990). Gibbard's expressivism has a number of virtues. It is a very general theory of normative discourse; it yields a semantics that is compositional and fully compatible with standard syntactic views; it yields a simple account of logical consequence for normative discourse.[3] Moreover, Gibbard's semantics for normative discourse dovetails well with the framework for modeling communication we are going to use, namely the one developed by Stalnaker (1978).

[1] In particular, so-called expressivism about epistemic and probabilistic discourse (see, among others, Yalcin 2011, Swanson 2011, Rothschild 2012, Yalcin 2012) might require some substantial changes to the argument. It's not clear to us whether, on these accounts, the update of shared information happens via set-theoretic intersection of some formal objects, as it happens for Gibbard, or whether they employ a different update operation (as it happens in the account given by Veltman 1996, which is an obvious source of inspiration for these views). Hence it's not clear to us whether these accounts can be modeled as a generalization of the Stalnakerian model of assertion we adopt.

[2] For a similar characterization, see Price (2011). Notice that we will reserve the term 'expressivism' to what Price calls '20th century expressivism'.

[3] In particular, we take Gibbard's semantics to have the best chances of overcome the notorious Frege-Geach problem. For the initial formulation of the problem, see Geach (1960, 1965). For an overview of the literature and discussion, see Schroeder (2008a,b, 2010), among many others. As we note below, there is still considerable skepticism that Gibbard's picture can provide a real solution to the problem.

2.2 Gibbard's norm-expressivism

Gibbard starts by introducing a new kind of attitude, which he calls 'accepting a norm': this is the kind of attitude that determines what an individual regards as mandated, permissible, or forbidden. Your judging that cannibalism is wrong amounts to your accepting a norm that forbids cannibalism; Tom's judging that tax evasion is okay amounts to his accepting a norm that allows for tax evasion. Gibbard doesn't define the notion of accepting a norm: rather, he assumes that this attitude will be part of a (yet to come) fully developed empirical psychology, on a par with beliefs and desires. The notion of accepting a norm is supposed to play a central role in an evolutionary explanation of individuals' coordination in a social context.

One crucial part of Gibbard's enterprise consists in specifying a formal model for the contents of normative attitudes. To do this, Gibbard employs *complete systems of norms*. Think of a complete system of norms n simply as (something that determines) a three-way partition of possible courses of actions: those that are forbidden according to n, those that are permitted (but not required) according n, and those that are required according to n. Complete systems of norms (henceforth, for short: norms) play the same role in characterizing normative attitudes that possible worlds play for non-normative attitudes in possible worlds semantics. If you believe that Tom is a cannibal, we characterize the content of your belief state with worlds where (among other things) Tom eats human flesh. Similarly, if you accept that cannibalism is wrong, Gibbard models the content of your normative attitudes with a set of norms all of which forbid (among other things) eating human flesh.

With this model of mental content in place, it is easy to formulate a semantics for normative language that works in the usual recursive way.[4] So far, we have assumed that descriptive attitudes make distinctions between possible worlds and that normative attitudes make distinctions between norms. But it is best to let all sentences in a language denote formal entities of the same kind, for a number of reasons (for example, accounting for 'mixed' sentences like *If Tom eats people, he does something wrong*). So instead of assigning semantic values of different kinds to each fragment of the language, we assign *sets of world-norm pairs* to each sentence. Hence the semantic values of

(1) Tom is a cannibal.

(2) Eating people is okay.

are given by, respectively:

(3) $\{\langle w, n\rangle$: Tom is a cannibal in $w\}$

[4] Gibbard's own way of doing so yields the wrong results for attitude reports. But those are wrinkles that can be easily ironed out. Cf. Dreier, 1999, p. 571.

(4) $\{\langle w, n\rangle$: Eating people is permitted by $n\}$

That said, we will proceed under the simplifying assumption that all purely normative sentences get assigned a set of norms as semantic value. We do this in order to reduce clutter, but nothing in what we say below hinges on it. Indeed, it is straightforward to move back and forth between sets of norms and the sets of world-norm pairs that get assigned to purely normative sentences: simply identify each set of norms A with the set of world-norm pairs $\langle w, n\rangle$ such that n is in A.

Two final clarifications are in order. First, recent literature has raised some foundational questions about Gibbard's framework. The main worry is that Gibbard cannot specify (in a non-stipulative and non-circular way) what it is for two norms to be inconsistent. Here, we set this worry aside—we will assume that the problem can be solved one way or another.[5] Second, for the purposes of this paper, we are going to assume that formal objects like (3) and (4) work both as the semantic values recursively assigned to (1) and (2) and as the contents expressed when (1) and (2) are uttered in a conversation. There are good reasons to think that this is a conflation,[6] but one that is harmless for our purposes.

2.3 Assertion, communication, common ground

Throughout the paper, we help ourselves to Stalnaker's framework for modeling assertion and communication (see e.g. Stalnaker 1973, 1978, 2002, as well as Lewis 1979b). We choose this framework because it is a rigorous and increasingly familiar way of thinking of communication. But it's important to point out that endorsing the general framework is not strictly required to make the argument go through. (So far as we can see, there is only one feature of it that really *is* needed, namely the appeal to Stalnaker's notion of acceptance, illustrated below.) We start by presenting the framework as it applies to conversations about non-normative matters. We then suggest how to generalize it in a way that is friendly to expressivists' assumptions.

Stalnaker's first observation is that conversation takes place against a background body of information, which he calls the 'common ground'. Roughly, the common ground involves what is commonly taken for granted for the purposes of the conversation. (See below for a sharpening of this claim.) The purpose of an assertion, on this picture, is to expand this shared stock of information. A successful utterance of a sentence will bring about that the content of that sentence gets added to it.[7]

[5] Or perhaps, with one of the authors, that the problem turns out to be spurious on closer analysis. [CITATION OMMITTED FOR BLIND REVIEW]

[6] For discussion of this point, see Lewis (1980), Dummett (1981), Ninan (2010), Rabern (2012).

[7] In general (though not always) something else is also recorded in the background information. For example, it will go on record that the speaker made an utterance, that some particular words were uttered, and so on.

We can make this picture more concrete if we make a few basic assumptions about content, assertion, and communication. We represent a body of information as a set of possible worlds—the worlds that are compatible with that body of information. Following Stalnaker, we call the set of worlds modeling the background information in a particular context the 'context set'. The context set is a simple way of modeling the information that is part of the common ground. We also think of the content of a declarative utterance as a set of possible worlds, namely those worlds in which the utterance is true. Finally, we model the effects of an assertion on the context set as set-theoretic intersection: The effect on the context set of an utterance whose content is modeled with a set S of possible worlds is that of eliminating from the context set those worlds not in S.

Consider a simple example. Zoe and Yael are having a conversation about Tom. They jointly assume a number of propositions about him: for example, the proposition that Tom lives in Boston. We model the context set of their conversation as a set of worlds in which (among other things) Tom lived in Boston. Now, suppose Yael utters the sentence 'Tom didn't pay taxes last year'. If her assertion is successful, the context set will now get updated with this new piece of information: hence the new context set contains only worlds where Tom lives in Boston and didn't pay taxes last year.

We can specify what it is for some information to be "commonly taken for granted" in more detail. First, we need to introduce the notion that Stalnaker (2002) dubs 'acceptance'. This notion is entirely distinct from Gibbard's notion of accepting a norm that we introduced in section 2. The terminological similarity is just an unfortunate coincidence; the two notions emerged in distinct bodies of literature. (We will worry about how to deal with this clash in a few paragraphs.)

Acceptance is, quite simply, the attitude of taking a proposition for granted for the purposes of a conversation. Notice that, in some cases at least, this notion diverges in interesting ways from the notion of a belief. Consider the following example:

> Smith and Jones find themselves in a conversation. They both believe (suppose, truly) that, many years ago, Smith tried to steal money from Jones in a moment of desperation; moreover, they both believe that they believe this, and believe that they believe that, etc. In short, that proposition is the object of common belief between them. However, they both prefer avoiding to acknowledge that proposition. So, for the purposes of the conversation, they behave as if they did not believe this proposition.

That Smith once tried to steal money from Jones is not accepted and is not validated by the context set of their conversation. Notice that acceptance has a peculiar 'transparent' nature: if an agent represents herself as accepting p, then, as a matter of fact, she does accept p for the purposes of the conversation. One cannot lie about one's acceptances. The transparent nature of acceptance will play a crucial role in some steps of our argument.

The contents of the common ground are defined on the basis of acceptance. In particular, the common ground will be identical with what is commonly believed to be accepted in a conversation. Common belief is an iterated notion of belief: it is common belief that p just in case all the members of the conversation believe that p, all believe that they believe that p, etc. So, in summary:[8]

> It is **common ground** that p in a group if all members accept (for the purpose of the conversation) that p, and all believe that all accept that p, and all believe that all believe that all accept that p, etc.

The context set is an object that models the contents of the common ground so defined.

Using similar tools, we can also define a notion of speaker presupposition. Differently from common belief, presupposition is an individual attitude. Each speaker enters a conversation with their own presuppositions. Roughly, a presupposition is a proposition that the speaker takes for granted in the course of the conversation. We can capture this pretty simply with the current apparatus by identifying the presupposition that p with the belief that it is common ground in the conversation that p:

> Speaker S **presupposes** that p in a conversation iff S believes that it is common ground among participants in the conversation that p.

For the purposes of our argument, we will often take presupposition, much more simply, as mutual acceptance. But we will revert to the more precise notion where it matters.

2.4 Issues

The model of conversation and update that we've sketched contains a number of idealizations. In particular, it requires that we think of agents as seeking to rule out any possible world that differs, no matter how trivially, from the actual world. This idealization is often harmless. But it gets in the way of our fully understanding the goals of conversational interactions, something which will be relevant to our arguments below.

Fortunately, the literature contains a number of converging proposals on how to modify the common ground framework in order to abandon this idealization. Here we will give a brief and informal overview, which is inspired by Hulstijn (1997) and Roberts (2012). But the central ideas are ubiquitous in the literature and (as we're going to point out) they are foreshadowed by Stalnaker himself.

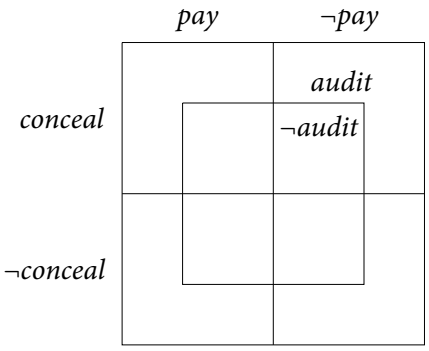
[8] As Stalnaker points out, we cannot just use iterated acceptance to define the common ground (Stalnaker, 2002, p. 716):

> The more general notion of common ground should not be just an iterated version of a broader notion of acceptance. Successful communication is compatible with presuppositions that are recognized to be false, but the information that they are being presupposed must be actually available, and not just assumed or pretended to be available. Even the liar, if he really intends to communicate, has to believe that the information needed to interpret his lies will really be common ground.

In any conversation there are number of questions that are of mutual interest to the participants in the conversation. These questions can be used to define a notion of *relevance*: a proposition is relevant in a context c iff it does not say more than what is needed to answer all of the questions. But they allow us to think of the common ground framework in a less idealized way.

If we model the possibilities compatible with the common ground with a set of worlds (the context set), we can model each question of interest—or 'question under discussion'—as a *partition* of this set: a collection of subsets of the context set that are pairwise disjoint and jointly exhaustive. Two worlds are in the same set iff they agree on the complete answer to the question.[9]

For illustration, suppose that Yael and Zoe are having a conversation with three questions under discussion: whether Tom paid taxes last year; whether he concealed anything from the IRS; and whether he is going to be audited. Here is a graphic representation of the partitions corresponding to each of these questions:



Given a set of questions under discussion, we can define a unique partition, which groups together worlds that agree on the answer to *all* the questions under discussion.[10] In the figure above, the partition is composed of those regions that are not divided by any of the lines.

On this picture, the goal of conversational participants is no longer to locate the actual world within logical space, but (more modestly) to individuate the true answers to all of the questions under discussion. This is equivalent to ruling out all the cells in the partition, aside from the one that contains the actual world. For example, suppose that Tom has indeed paid taxes, did not conceal anything from the IRS, but will nonetheless be audited. Then Yael and Zoe's goals are accomplished when all cells other than the bottom left L-shaped cell are ruled out.

This modification of the common ground framework is very much in the spirit of the original proposal. As Stalnaker puts it (1981, p. 136):

[9] Cf. Groenendijk and Stokhof, 1984; Hamblin, 1958, 1973; Lewis, 1988, *inter alia*.

[10] More precisely: this is the coarsest partition that is a refinement of all the questions under discussion.

> The alternative possibilities used to define propositions must be exclusive alternatives which are maximally specific, relative to the distinctions that might be made in the context at hand. [...] One might think of possible worlds as something like the elements of a partition of a space, rather than as the points of the space. The space might be partitioned differently in different contexts, and there might be no maximally fine partition.

It would be best to think of the possibilities that make up the common ground not as possible worlds, but as elements of the partition induced by the collection of questions under discussion.

Following the terminology of (Savage, 1972),[11] we can call the cells that are generated by the set of questions under discussion 'small worlds'. (They're called 'small' because, even though they look bigger in diagrams, they contain strictly less information than full-blown possible worlds.) Small worlds are different from possible worlds, but the common ground model works exactly in the same way, whether we use the ones or the others. In both cases, assertion works by ruling out open possibilities.

In what follows, we take small worlds to be the basic units of the context set. What small worlds are in play will be a function of context: we assume that, for each context, there are a number of questions under discussion that induce a partition of the context set into small worlds.

It is worth noting that further questions under discussion might be added in the course of the conversation, both explicitly, i.e. by uttering an interrogative sentence, or implicitly. To see how the latter may come about: suppose that, while discussing Tom's dealings with taxes and the IRS, Yael says: "Tom also didn't pay for his mandatory medical insurance last year". By doing this, she introduces a further question under discussion—whether Tom paid for his medical insurance last year. The introduction of this question splits each of the small worlds in play into more informative small worlds.[12] The implicit introduction of a question under discussion can be assimilated to accommodation, in the sense of (Lewis, 1979a).[13]

All considerations that we've just made can be reproduced, *mutatis mutandis*, for the case of Gibbardian norms. We will assume that the context set in conversations

[11] See also Joyce 1999 for recent dicussion.

[12] More accurately, we might represent the contribution of the assertion to the conversational score in a two-tier fashion. First, there is a refinement of the partition in play: each small world is split into more informative (larger) small worlds, by grouping together worlds within a given cell according to how they answer the new question. Second, the set of (new) small worlds where Tom does pay medical insurance are ruled out.

[13] Accommodation is the phenomenon whereby a speaker brings about a presupposition in the common ground by making an assertion that requires that presupposition to be in place. Borrowing an example from Yablo 2006: if I tell you "I'm picking up my guru at the airport", *ceteris paribus* the common ground will come to include the proposition that I have a guru, even if that proposition was not antecedently presupposed in our conversation.

about normative matters consists of 'small norms'—sets of complete systems of norms that grouped according to how the settle all normative questions under discussion.

We have claimed that speakers normally raise and add questions under discussion to the conversational scoreboard, even just by making simple assertions. But note that we have not claimed that every issue that becomes somehow salient is added to the questions under discussion. On the contrary—sometimes speakers can, more or less explicitly, raise an issue, and nevertheless that issue may not be added to the questions under discussion. This might happen for many reasons.[14]

As a first example, take the following scenario: you have an uncle that has very different political beliefs than you do. During your conversations about various topics, your uncle often makes in-passing claims that presuppose some of his views about (say) abortion—perhaps with toss-off remarks like "I'm not one of these liberals who deny that babies in the womb have a right to life". Wanting to avoid lengthy and pointless confrontation, you just don't pick up on these remarks and carry on with the main conversational topics. In this scenario, the question whether abortion is permissible, or the question whether embryos are babies, don't become part of the common ground.

Here is a second example. Suppose that you are discussing with an epistemicist about vagueness and they raise the question whether your common acquaintance Tom, who is exactly 5 ft 9, is tall. Given your own philosophical leanings, you're convinced that there is no fact of the matter as to whether Tom is tall. As a result, in your view there is no point whatsoever to discussing the question whether Tom is tall. There is no 'right' way to settle the question, hence you refuse to take it up.

We characterized the latter case as a case of indeterminacy—a case where there is no fact of the matter about the relevant question. These labels seem unfit for the expressivist's content (since they involve an appeal to notions like truth and facts). But it seems clear that there are cases of normative undecidedness that parallel the vagueness case in all relevant respects. For example, suppose you believe that it's mandatory not to eat meat or fish, but there's nothing to settle whether the correct ethical position is vegetarianism or veganism. On the view you endorse, there is simply nothing to adjudicate between the two; ethical norms 'run out' without deciding the issue. In this situation, you will refuse to take up the question whether one should be vegetarian or vegan.

As a summary:

> The set of questions under discussion in a conversation is the set of questions Q such that, for each Q, all speakers intend to engage in inquiry to settle whether Q.

[14] Thanks to an anonymous referee for pushing us here in number of ways. We also owe them the reactionary uncle example.

In short: the live questions are those that speakers intend to discuss.[15] As we said above, the questions under discussion in a conversation generate the small worlds in the context set. It follows from what we said that, for each small world in the conversation, all speakers involved are willing to take up the question whether that particular small world in the conversation should be ruled out.

Even though we take small worlds as the basic units of the context set, to avoid clutter we will keep using the simple terms 'world' and 'possibility'. In any case, most of the things we say will hold both for the case of small worlds and full-blown possible worlds. We will remind you of the explicit assumption that small worlds are the ones in play in places where it matters. Similarly, we will just talk about norms. As a result, when we say that conversation about normative matters involves distinguishing among complete systems of norms, we should be understood to mean 'complete only to the extent required by the purposes of a conversation'. We will remind you of the explicit assumption that small norms are the ones in play in places where it matters.

2.5 Expressivist assertions (thick and thin)

Gibbard-style expressivists can take on board the general picture of communication outlined above. Here is a sketch of how the resulting account could go. In a conversation about normative matters, speakers try to influence what norms others accept. There are now two parameters they track: the worlds compatible with what they commonly assume in the conversation, and the norms compatible with their common norm-acceptances. Hence the possibilities that are part of the common ground are different from those we find in the descriptivist model. But the structure of the apparatus remains the same. In particular, we still model the common ground as (roughly) the set of possibilities (norms) that are left open by speakers' attitudes, and we still model the effect of an assertion on the context set by intersecting it with the semantic value of the sentence asserted.

While the formal machinery is easily exportable, the application of the apparatus to the expressivist case raises substantial philosophical issues. These issues will be the starting point for our main line of argument in the paper. Before heading there, let us address two preliminary points.

First, a point about terminology. The notions of acceptance, common ground, and presupposition have been introduced in a standard truth conditional framework.

[15] This characterization has a variant worth considering:

> The set of questions under discussion in a conversation is the set of questions Q such that, for each Q, all speakers intend to engage in inquiry to settle whether Q, and all speakers believe that they intend to engage in inquiry to settle whether Q, and they believe that they believe it, etc.

The difference between these two characterizations is irrelevant for our purposes, so we stick to the simpler one.

Hence, on the standard way of understanding them, these attitudes concern descriptive contents. We would need to introduce new terminology to capture counterparts of them that also apply to normative contents. This would involve a lot of clutter, and presumably would introduce confusion. So we'll take a shortcut. We just stipulate that we're going to understand the standard vocabulary of theories of assertion and communication in an enlarged way. We take a proposition to be modeled either by a set of worlds or by a set of norms.[16] We use 'belief' to pick out both descriptive beliefs and attitudes of accepting a norm in Gibbard's sense. Hence we'll talk about a subject believing that (say) tax evasion is wrong, while still understanding this attitude in a noncognitivist way. Similarly for acceptance and presupposition.[17]

Next, let us make an important clarification. One might worry that the common ground model of assertion is incompatible with Gibbard's norm expressivism.[18] The notion of assertion, the objection goes, is closely connected to notions like truth and knowledge, which belong in a factualist picture of language.[19] Perhaps there are no formal obstacles to extending Stalnaker's formalism to Gibbardian contents. But doing so involves dropping some of the main philosophical commitments of expressivism.

The objection relies on a construal of assertion that is different from the one we adopt. Let us distinguish a 'thin' and a 'thick' construal of assertion. We agree that, on one construal of assertion—the 'thick' one—the notion of assertion has substantial, perhaps conceptually necessary connections to truth and knowledge. But here we understand assertion in a different way. On this construal, assertions are understood just via their capacity of updating speakers' attitudes in a certain way. Let us explain.

Start from the notion of common ground. Recall: a proposition is common ground in a conversation just in case all speakers accept it and it is commonly believed that they do. Hence what is common ground is automatically determined by speaker's individual attitudes. As a result, there are no problems in defining an expressivist-friendly notion of common ground. As soon as we specify what attitudes individual speakers have, we also immediately get their common ground; we don't need to make any commitments about the relevant attitudes being cognitive or representational in some way. Assertion, as we understand it throughout this paper, is the speech act whose functional role is to update the common ground, thus defined, in a certain way. In particular, assertions update the common ground by removing certain open possibilities from it. In a slogan: assertion is the speech act that produces update by intersection.

Notice that this explication of assertion makes no appeal to notions like truth and knowledge. We say that assertions remove possibilities from the common ground:

[16] In full generality: a set of world-norm pairs.

[17] For discussion of this and related terminological issues, see Schroeder 2010, p. 85ff.

[18] Thanks to an anonymous referee for pushing us to consider this and related issues.

[19] For some classical accounts vindicating these connections, see Dummett 1959 and Williamson 1996, 2000.

these possibilities can be understood either as possible worlds or as complete systems of norms (or better yet, as world-norm pairs). Of course, one may define a 'thick' notion of assertion that incorporates our 'thin' construal as a part, by explaining how this kind of update of the context set is connected (perhaps in normatively interesting ways) to truth and knowledge. This thick notion, for all we say in this paper, is perfectly consistent, and may even provide the best way of systematizing our intuitive talk of assertion. As a matter of fact, we do think that the 'thin' construal of assertion is more theoretically fruitful and cuts closer to the natural joints than many of its alternatives. But for present purposes we can remain neutral on this. All we want is to point out that we will understanding assertion in this 'thin' way.

The thin construal of assertion ties assertion to linguistic form. In general, assertions are the speech acts that are performed by declarative sentences. We say 'in general' because not all utterances of declarative sentences will do: sometimes we can utter a declarative sentence with no intention to produce a speech act (for example, just to practice our pronunciation). Assertions are the speech acts that are performed via those utterances of declarative sentences that are intended to update the common ground of the conversation.[20]

It seems uncontroversial that, on the thin construal, expressivists like Gibbard can help themselves to the claim that declarative utterances of normative claims count as assertions. Much like declarative utterances of descriptive claims, these utterances aim at altering the common ground by reducing the possibilities in play (only, they rule out normative rather than descriptive possibilities). And aiming at ruling out possibilities is exactly what makes a speech act a thin assertion.

3 The problem

We established that expressivists can borrow from descriptivists a formal framework for modeling content, assertion, and communication. But the formal framework leaves one important question unaddressed. Expressivists should be able to *make sense* of the

[20] We should note that this clashes with some classical views on assertion, for example the view emerging from Searle's famous taxonomy of speech acts (1976). For Searle, there are categories of speech acts that standardly involve declarative sentences, and that yet don't count as assertions—for example, promises. Suppose that, when Tom confesses to Yael that he evaded taxes, Yael reassures him by saying:

(i) I will not give you away to the IRS.

For Searle, (i) is not an assertion. In short, and roughly, for Searle assertions are devices for reporting facts, while (i) is a way of bringing about a new fact. (Searle puts the point in terms of direction of fit: assertions have a mind-to-world direction of fit, while promises—or, in his terminology, commissives—have world-to-mind direction of fit.) But, within a common ground picture, an utterance of (i) in the relevant circumstances fully counts as an assertion, since it rules out possibilities in the usual way. Whatever other effects an utterance of (i) brings about, these are distinct from its basic role in conversation.

practice of conversation. They need to explain why it's reasonable for a speaker to have certain expectations about the uptake of what she says, and why it's reasonable for a hearer to actually take up what the speaker says.

In the descriptivist version of our model of communication, the explanation goes as follows. Communication is a self-locating enterprise. Speakers aim to find out what portion of logical space they occupy. Exchanging information is just a way of jointly narrowing down the set of candidates for the actual world. But the expressivist can't help herself to a normative counterpart of the actual world. It is a basic assumption of her view that there is no such thing. Hence the expressivist owes us a story about the point of engaging in communication about normative matters, one that she can't simply extract from the descriptivist framework she's using.[21]

As we flagged in the introduction, expressivists have gestured towards a way of making sense of communication that generalizes to descriptive and normative discourse alike. The idea, introduced by Gibbard himself, is to think of communication primarily as an exercise in coordination. We are social creatures: we engage in conversation because coordinating on some particular family of attitudes—beliefs, say, or norm-acceptances—is likely to help us meet our goals. Here is Gibbard explicitly making this point:

> The biological function of the mechanisms underlying our normative capacities is to coordinate. Hence the psychic mechanisms that produce normative judgments are not systems of natural representations, they are coordinating systems. Their biological function is not to put something in the head in correspondence with their subject matter; it is to coordinate what is in one person's head with what is in another's. (Gibbard, 1990, p. 110)

Seth Yalcin, who endorses an expressivist account of epistemic modal talk, makes a proposal along similar lines:

> [I]n modeling the communicative impact of an epistemic possibility claim, we construe the objective as one of coordination on a certain global property of one's state of mind—the property of being compatible with a certain proposition—not one of coordination concerning the way the world is. (Yalcin, 2011, p. 310)

[21] One might worry that, at bottom, this is not really a question about normative *language*. The worry goes as follows. The basic datum seems to be that some notion of correctness seems to be in place. But correctness will ultimately attach to the *attitudes* that speakers have towards the expressivist's contents. Hence a full answer to the questions just raised will have to go through a general story about the correctness of normative attitudes and it's a useless exercise to focus exclusively on language. Although we are sympathetic to this worry, it relies on a number of assumptions that here we don't want to make. It's not clear to us that one needs an antecedent notion of correctness for attitudes to make sense of communication. Perhaps there is a way to do without it. In any case, it's an interesting question whether we can get an account of normative communication as a rational activity without relying on an antecedent notion of correctness for normative attitudes. This paper can be seen as an attempt at seeing how far one can go under this constraint.

To a first approximation, this seems like a good strategy. The point of engaging in conversation about non-normative matters is to coordinate on a body of beliefs. The point of engaging in conversation about normative matters is to coordinate on a system of norms. We work out, as a community, what to think about the world. We work out, as a community, what system of norms to accept. Communication can be understood as a way to foster such coordination.

Unfortunately, this is far from a full account. There are many ways of achieving coordination that differ from communication in crucial ways. A newly established community might coordinate on what side of the road to drive on via a random procedure like flipping a coin. This is a perfectly acceptable way of achieving coordination. But we don't rely on coin flips when working out what to do, no more so than we rely on coin flips when working out what the world is like. So saying that communication is a kind of coordination is just the beginning of an account. There is a big gap to fill by explaining exactly what kind of coordination this is and what constraints apply to it.

One natural thought is that, when picking a side of the road to drive on, all we want is that we all drive on the same side of the road. None of us cares *which* side that is. Perhaps the case of communication is different because agents approach communication with a number of initial beliefs (normative or not). Besides jointly coordinating on a sector of logical and normative space, they care about making it the case that this sector not be disjoint from the sector that is already individuated by their initial attitudes. In other terms, agents care about maximizing the extent to which their initial views are preserved after communication.

There is something right about this suggestion, but it also falls far short of an account. If communication functioned this way, it would be mostly an exercise in bargaining. In other words, communicating would be about proposing compromises that have a high enough payoff for all agents involved.

Of course, we can and do engage in bargaining of this sort. Suppose that you and a friend care about spending the evening together, but have a number of conflicting preferences. In that situation, it would not do to just flip a coin to settle each detail of your plan. There would be some give and take: you eat at the diner rather than the fancy restaurant, as per your own preferences; but you go to the opera rather than the roller derby, as per your friend's. But this is not what we do when we discuss descriptive and normative matters. Then we try to reach a set of shared attitudes by giving reasons to each other. The idea that we could compromise by just trading up some of our beliefs and acceptances for agreement is out of place.[22]

Thus we grant that coordination should play a part in a general account of com-

[22] Incidentally, this observation also shows that the expressivist's acceptances, although they are a conative attitude of some sort, cannot be construed too closely on the model of preferences. We take it that this is also one of the main lesson of Dorr 2002—though, unlike Dorr, we don't take this disanalogy to doom expressivism.

munication. But a desire for coordination on our beliefs or our normative acceptances cannot explain the complexities of our communicative practices. The expressivist needs much more to get a full account.

4 Explaining communication

We now turn to the question of how this idea of communication as coordination needs to be developed if it is to explain why communication works the way it does.

In order to spell out our answer, we first examine the assumptions agents need to make in order to purposefully and rationally engage in communication about normative matters. We argue that, if it makes sense for an agent to engage in communication, she must take for granted that there is a kind of objectivity about normative matters. As a result, in each context, it is presupposed that speakers are required to converge on a unique norm, at least insofar as they choose to remain engaged in communication. This is going to explain why our practice of communicating and debating normative claims is exactly analogous to our practice of communicating and debating non-normative claims. In normal conversations, speakers presuppose that there is an actual world: hence they presuppose that there is a point in the context set to which, at least ideally, their acceptances should converge. Similarly, we claim that they jointly presuppose that there is an 'actual' norm—a norm to which, at least ideally, their beliefs ought to converge.

4.1 The claim: convergence

More precisely, we argue for the following claim:

> CONVERGENCE. In any 'normal' conversation, it is common ground that there is a (unique) norm on which the participants' normative acceptances ought to converge.

We argue that CONVERGENCE follows—surprisingly, perhaps—from some minimal assumptions about the rationality of agents engaging in communication and our practice of assertion.[23]

[23] Notice that CONVERGENCE does not amount to a capitulation to a form of realism or objectivism. A realist will presumably endorse the following thesis:

> CONVERGENCE*. There is a unique norm (the 'true' norm) to which, in any 'normal' conversation, participants' acceptances ought to converge.

But CONVERGENCE* is crucially distinct from CONVERGENCE: besides important scope differences between the operators involved, CONVERGENCE concerns what is common ground in ordinary conversation, while CONVERGENCE* concerns what is actually the case. Also, notice that CONVERGENCE differs from

Let us say a bit more about the nature of the *ought* figuring in CONVERGENCE. This is an *ought* that is supposed to capture norms regulating communication between rational agents: i.e., norms that specify what a rational agent who has the goal of engaging in a communicative exchange ought to do. Hence it is an *ought* of practical, as opposed to epistemic, normativity. A rough gloss of '*A* ought to φ' would go 'In view of the requirements of practical rationality applying to her qua agent involved in a communicative exchange, *A* ought to φ'.

As we mentioned in the introduction, this is the *ought* that captures the normative force of Grice's Cooperative Principle. In its most general formulation, the Cooperative Principle says:

> Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged.

Grice never gave a full explication of why the Cooperative Principle is in place. But he is very explicit about what kind of rule he takes it to be:

> . . . I would like to be able to show that observance of the Cooperative Principle and maxims is reasonable (rational) along the following lines: that anyone who cares about the goals that are central to conversation/communication (such as giving and receiving information, influencing and being influenced by others) must be expected to have an interest, given suitable circumstances, in participation in talk exchanges that will be profitable only on the assumption that they are conducted in general accordance with the Cooperative Principle . . . (Grice (1975), pp. 25-6.)

Grice claims that anyone who cares about the central goal of communication is rationally required, in light of their own goals as a communicator, to make their conversational contribution adequate in certain ways. Similarly, we claim that, in a normal conversation, speakers presuppose that there is a unique norm such that they are rationally required, in light of their own goals as communicators, to have their attitudes converge on that norm. Exactly like the *ought* regulating the cooperative principle, this *ought* regulates what agents *do* in conversation, and hence is an *ought* of practical rationality.[24]

---

the following as well:

> CONVERGENCE** There is a unique norm such that, in any 'normal' conversation, it is common ground that the participants' acceptances ought to converge to that norm.

While CONVERGENCE** does entail CONVERGENCE—at least, given some reasonable assumptions—it is not entailed by it.

[24] There are traditional worries about the idea that requirements of practical normativity may apply to an agent's cognitive states. These worries are connected to skepticism about the possibility of doxastic control. But notice that the notion of acceptance we are using is different from the notion of belief (see below). Acceptances are propositions that agents choose to take for granted for the purposes of the conversation. Agents may not be able to decide what to believe, but they're able to decide what to accept in a conversation.

Before moving on to our main argument, we should emphasize that 'norm' in CONVERGENCE is short for 'small norm', i.e. a set of norms that is a cell of the partition generated by questions under discussion. Hence CONVERGENCE amounts to the requirement that, in 'normal' conversations, speakers presuppose that there is a unique answer to all questions under discussion such that that they ought to converge on that answer. This obviously falls short of a presupposition that there is a fully specified normative standard that agents ought to converge on. But it still is a very substantial step towards endorsing a kind of presupposition of objectivity in normative conversations. If CONVERGENCE is right, then, whenever speakers successfully raise a normative question, and indeed whenever they successfully make an assertion about normative matters,[25] they are presupposing that there is an answer to that question they all ought to accept.

Beyond being a step in the direction of objectivity, this arguably establishes a full parallel with conversations about non-normative matters. Also in that case, speakers presuppose that there is a unique 'small world' (the one containing the actual world) to which their acceptances ought to converge. Plausibly, in real conversations they do not presuppose that there is a unique, maximally specific, possible world to which their acceptances ought to converge. Such a presupposition would presumably be in conflict with Gricean assumptions about relevance. Questions under discussion are all that is relevant in a conversation, and speakers have no interest in irrelevant information.[26]

4.2 Overview of the argument

Let us start by introducing a bit of formalism. We use 'n' and 's' as variables ranging over systems of norms and speakers, respectively. We use 'Acc_s' to denote the set of systems of norms compatible with what s accepts, and we use 'Acc_s' as an 'acceptance' operator, saying that the embedded proposition is accepted by s. These notational conventions will allow us to make some helpful abbreviations as we go on. To begin with, we can reformulate our main claim in a shortened form:

CONVERGENCE. The following is common ground in any conversation:

$$(\exists n)(\forall s)\mathsf{Ought}(\{n\} = Acc_s).$$

Our argument for CONVERGENCE will proceed by establishing a weaker claim. In intuitive terms, this claim is that, in normative conversation, it is common ground that

[25] Recall that, as we pointed out in section 2.4, making an assertion is a simple and effective way to implicitly raise a question (in terms of accommodation, to make it the case that a question is accommodated in the common ground).

[26] Thanks to an anonymous referee for pushing us to consider the apparent conflict between CONVERGENCE and the maxim of relevance.

there is the possibility of a mistake—i.e., it is common ground that, by participating in normative conversations and ruling out normative possibilities in accordance with what is said, speakers might rule out a normative possibility they ought not rule out. This claim can be formalized as follows:[27]

CAN GO WRONG. The following is common ground in any conversation:

$$(\exists n)(\forall s)\mathsf{Ought}(n \in Acc_s).$$

This claim constitutes the center of our argument. We will spend the bulk of this section arguing for it.

Once we have done that, we will be in a position to establish our main conclusion. We will rely on another (much more straightforward) claim about what is common ground in conversation. This is just the claim that, as long as speakers engage in conversation, they presuppose that they ought to narrow down the set of relevant possibilities to one.[28] In symbols:

UNIQUENESS. The following is common ground in any conversation:

$$\mathsf{Ought}((\exists n)(\forall s)Acc_s = \{n\}).$$

Our main claim, CONVERGENCE, will follow from UNIQUENESS and CAN GO WRONG.

4.3 Not anything goes

The main premise in our argument, CAN GO WRONG, is that, in every conversation, it is common ground that there is a norm that ought not be ruled out from the acceptance states of the participants. At an intuitive level, this premise can be regarded as a minimal requirement of objectivity. It must be common ground that at least one of the normative possibilities that speakers regard as open ought not be ruled out.

Notice that this is much weaker than the claim that speakers' attitudes should converge on a unique norm. It only requires that they keep regarding one of the possibilities as open: i.e., that they not be allowed to accept normative propositions incompatible with it.

Our argument for CAN GO WRONG is somewhat involved. The basic point, however, can be stated quickly and informally. Consider a conversation between two agents, *A* and *B*, and suppose that the following is the case:

The 'anything goes' situation. *A* accepts that there is no norm such that *A*

[27] Henceforth, we drop the qualifier 'normal' for the sake of readability.

[28] Here again it is worth reiterating that the norms in question are 'small norms'—maximally specific relative to the normative questions under discussion. Cf. §4.5.

ought not rule out that norm. In other words, the following is the case, according to A:

$$\neg\exists n \mathsf{Ought}(n \in Acc_A)$$

We claim that, in the 'anything goes' situation, it would be pointless for B to engage in conversation with A.

It's helpful to consider an example. Suppose A and B have narrowed down the common ground to a set consisting of two norms (by which, once more, we mean two cells in the partition determined by their questions under discussion). On one of these norms cannibalism is wrong, while on the other it is okay. And now, suppose that A and B are indeed in an 'anything goes' situation, so that, according to A, she is allowed to accept that cannibalism is wrong or to accept that cannibalism is not wrong. B might then say (for example) "Cannibalism is wrong". A can agree that there's nothing wrong at all with B's saying or thinking that—she would not be willing to say that B said anything she ought not have said. Further, A could agree that there is no reason for A not to accept B's answer—she takes herself to be allowed to rule out the norm according to which cannibalism is okay. And yet, she may stick to her belief that cannibalism is okay, since she also takes herself to be allowed to rule out the norm according to which cannibalism is wrong.

We claim that, in this kind of situation, it would make no sense for B to engage in conversation with A. (Or at least: it would make no sense for B to engage with A in a conversation as long as the question whether cannibalism is wrong is a question under discussion. We get back to this shortly.) There would be no point to it. Hence, if B agrees to engage in communication with A, she must do so under the assumption that A doesn't take herself to be in the 'anything goes' situation.

Let us clarify and generalize the foregoing.

First, let us highlight the basic theoretical point brought out by the example. Suppose B takes a question Q to be a question under discussion. Then, we claim, this entails that B is accepting that A accepts that there is an 'objective' answer to Q—i.e., that B is accepting that A accepts that there is an answer to Q that she ought to accept. In other words, questions under discussions are (among other things) questions that we take to have an answer that we ought to endorse. We cannot take a question to be a question under discussion and at the same time be permissivists about it (i.e. think we may accept either answer to it).[29]

[29] Strictly speaking, our example shows something slightly weaker. It shows that, *by attempting to settle one of the questions under discussion*, B is ruling out the possibility that A takes it to be permissible to adopt either answer. But, one might object, maybe there are questions that we take to be under discussion, but that we assume will never be settled. Won't our claim fail in this case? We think that this case can be safely ignored. In taking Q to be a question under discussion, one acts as if one believes one will get to address it. Much like in placing an item on the agenda one is acting as if one believes one will get to it during a

Second, our example concerns a simple case where there is only one question under discussion. But the point generalizes. In every conversation, there is a set of questions under discussion $\{Q_1, Q_2, Q_3, ...\}$.[30] The norms that are in the context set correspond to conjunctions of answers to all these questions. Now, suppose that, according to A, it is not the case that there is a norm A ought to accept. Then there is at least one question under discussion Q_i such that, according to B, it would make no sense for her to attempt to answer that question.

Third, as we hinted at already, the issue whether it makes sense for a speaker B to engage in conversation with A should be relativized to what questions are under discussion. To see the point, it helps to consider the following objection.[31] Suppose we have three (mutually exclusive and jointly exhaustive) small norms: n_1, n_2, and n_3. For vividness, you can identify these norms with three ethical stances about diet: n_1 allows omnivorism, n_2 mandates vegetarianism, n_3 mandates veganism. Now, suppose that Alice and Bob are discussing ethical issues relating to diet. Alice feels strongly about n_1; it's important to her that n_1 be ruled out. But she thinks that nothing settles whether one should adopt n_2 or n_3. In her view, ethical rules 'run out' before settling which of n_2 or n_3 is right. In this case, one might worry, the context set of the conversation between Alice and Bob fits our characterization of the 'anything goes' situation, since it is not the case that there is one particular norm that Alice ought not rule out. Yet Alice does not take herself to be allowed to accept any answer whatsoever to the questions under discussion. In this case, the worry goes, it might still make sense for Bob to engage in conversation with Alice, despite the fact that we are in an 'anything goes' scenario.

We agree with the objector that, in the case described, it makes sense for Bob to engage in conversation with Alice. But we deny that the questions under discussion are the ones she describes. If Alice really thinks that norms 'run out' before settling whether n_2 or n_3 is right, then the relevant question under discussion in the conversation is not which of n_1, n_2, or n_3 to adopt. Rather, the relevant question is which of n_1, or (n_2 or n_3) is right. Recall from section 2.4: a question is under discussion just in case all speakers intend to engage in inquiry to settle it. But Alice does not intend to engage in inquiry to settle whether n_2 or n_3 is correct. She thinks ethical rules 'run out' before deciding which is correct. So the subquestion whether n_2 or n_3 is right is not on the agenda of the conversation.[32]

---

meeting, even if one believes one may not get to it, in taking Q to be a question under discussion one is acting as if one of the participants in the conversation will attempt to answer Q.

[30] For simplicity, we will assume that all questions under discussion are simple yes-no questions. Nothing in our argument hinges on this.

[31] Thanks to an anonymous referee for pushing us to consider this case.

[32] Of course, Alice can, for the sake of the argument, act as if she thinks that ethical rules do settle whether n_2 or n_3 is correct. In this case, she *will* intend to engage in inquiry to settle whether n_2 or n_3 is correct, at least for the purposes of the conversation, and it makes sense to engage with Alice in conversation about the issue.

Let us recap our basic point. *A* and *B* are in conversation. In the background, there are a number of questions they intend to discuss. *B* makes a claim that amounts to an answer to one of those questions—say 'Omnivorism is ethically forbidden'.[33] For *B*'s speech act to make sense, she must think that (a) it is under discussion whether omnivorism is wrong; and (b) *A* is not a permissivist about the relevant question, i.e. *A* doesn't think that she may accept either answer to the question whether omnivorism is forbidden. This is going to be our starting point in developing the rest of our argument.

Before moving on, let us make an important clarification. We are *not* claiming that, whenever an agent is in an 'anything goes' situation, it makes no sense to try to coordinate with her. Consider other instances of coordination, like choosing which side of the road to drive on. This case arguably does display the permissibility that we want to deny for the case of communication. Neither agents has a reason to pick one side rather than another and they take this to be so. Yet it still makes sense for them to make proposals about how to coordinate. So there is a point to coordinating activities even when agents take themselves not to be subject to any normative constraints.

What we *are* claiming, by contrast, is the following: in an 'anything goes' situation, coordination may not proceed *via assertions of declarative sentences*. An 'anything goes' situation allows for a gap between recognizing that a proposal to coordinate is legitimate and accepting that proposal. Our practice of making and accepting assertions doesn't recognize this gap. To see the point, contrast the following two conversations:

A: I suggest that we drive on the right-hand side of the road.
B: Nah, no way.

C: The Brits drive on the left-hand side of the road.
D: Nah, no way.

Both *B* and *D* are rejecting what their interlocutor said. But notice a crucial difference. By rejecting *A*'s utterance, *B* is not taking issue with the appropriateness of *A*'s speech act. *B* can (and normally will) deem *A*'s speech act perfectly successful; *B* needs find no fault with *A*'s advancing the suggestion that they drive on the right. This is fully compatible with not accepting that suggestion. (Another way to see this: notice that *B* could express his rejection as: "Okay, you've made a perfectly fine suggestion. But I refuse to accept it.") But the situation is different for the case of *C* and *D*. *D*'s rejection amounts to a rejection of *D*'s speech act. *D* cannot think that there is nothing wrong with *C*'s speech act, and at the same time think it's okay to not update their belief as suggested by *C*. (Notice the awkwardness of saying: "Okay, you've made a perfectly fine assertion. But I refuse to accept it.") For the case of assertion, there is no gap between finding an assertion appropriate and taking up the content of that assertion.

[33] Or perhaps something weaker, i.e. a disjunction of answers to one or more questions.

Notice that we are *assuming* (on the basis of simple empirical observation) that assertions of moral claims work in this way.[34] This is one of the starting points of our argument. We are not trying to explain, on behalf of the expressivist, why this is so. Thus we're not trying to give a full account of communication on behalf of the expressivist. What we're doing is part of this enterprise, but we must leave the completion of the task to a different occasion.[35]

In summary: while coordination is possible (and indeed it does often happen) in an 'anything goes' situation, in such a situation it would make little sense to attempt coordination by way of utterances of declarative sentences.

4.4 Establishing CAN GO WRONG

We have argued that, for it to make sense for B to engage in conversation with A, B must accept that A is not in an 'anything goes' situation:

$$\neg \mathsf{Acc}_B(\mathsf{Acc}_A \forall n(\mathsf{May}(n \notin Acc_A)).$$

We now claim that this is enough to establish the following claim:

(AO) $\quad \mathsf{Acc}_B\mathsf{Acc}_A(\exists n\, \mathsf{Ought}(n \in Acc_A))$

In plain English, this says that B accepts that A accepts that there is a norm A ought not rule out.[36] Notice that (AO) says that B presupposes a *descriptive* proposition about A's attitudes: namely, the proposition that A accepts that there is one norm that she is not allowed to rule out.

Of course, it is not in general true that if B must believe that it is not the case that p, B must therefore believe that that $\neg p$. We argued that if B is convinced that, according to A, anything goes for her, then it makes no sense for her to engage in communication. But, wouldn't it be enough for B to *leave open* the possibility that A takes herself to be

[34] Remember that we are using 'assertion' in the thin sense: our assumption is thus that this is the way utterances of declarative sentences work, at least in 'normal' cases.

[35] Let us just briefly flag a natural direction the expressivist might go to answer this challenge. She might argue that the speech act of assertion is connected to epistemic practices requiring us to provide shared reasons for backing assertions or for rejecting them. This idea connects with a pretty different line of thinking about assertion than the one we're using, namely the one developed in Brandom (1994). We think that, in principle at least, these lines of thinking may be reconciled; they seem to investigate different and complementary aspects of assertion. Roughly: Stalnaker's account focuses on the informational aspect of assertion, and on how assertions affects the speakers' mental states in conversation; a Brandomian theory of assertion spells out the epistemic norms that are associated to assertion.

[36] There are modes-of-presentation-kind issues with this formulation. What we want to say is that B accepts that A accepts that there is a norm that *she herself*, individuated in the self-locating way (see Lewis 1979a; Perry 1979, among many), ought not rule out. These issues are orthogonal to our main concern here and can be put on the side.

subject to some constraints in what she accepts? Why think that (A0) is required for it to make sense for B to engage in conversation with A?

It is here that the peculiar aspect of the notion of acceptance becomes crucial to our argument. We grant that B may not *believe* that it's not true that A accepts that anything goes for A. Nevertheless, for the purposes of conversation, she must take that for granted. To make an assertion is not to *try* to make a demand that the hearer changes her attitude. Rather, it just is to make a demand to that effect. By making an assertion, the speaker represents herself as accepting that it's not true that anything goes. Hence she takes that for granted for the purposes of the conversation.

Three intermediate premises. Once we have (A0), it's easy to get a further claim that will be useful. If A is rational, she will see the reasoning explained above and realize that (A0) is required for communication to function.[37] Hence she will accept (A0), which gives us the following:

(A1) $\mathsf{Acc}_A\mathsf{Acc}_B\mathsf{Acc}_A(\exists n\,\mathsf{Ought}(n \in Acc_A))$.

The next step consists in arguing for the claim that each speaker accepts that there is a norm that she ought not rule out. In symbols, this involves eliminating two iterations of the 'Acc' operator from (A1):

(A2) $\mathsf{Acc}_A(\exists n\,\mathsf{Ought}(n \in Acc_A))$.

The argument is pretty linear. Suppose that, for any proposition (descriptive or normative) φ, we have both that B accepts that A accepts φ, and that A accepts that B accepts that A accepts that φ:

(i) $\mathsf{Acc}_B\mathsf{Acc}_A\varphi$

(ii) $\mathsf{Acc}_A\mathsf{Acc}_B\mathsf{Acc}_A\varphi$

Now, assume that, despite this, A doesn't accept φ, i.e. $\neg\mathsf{Acc}_A\varphi$. Then A has a choice. She can either point out that actually she does not accept φ; or she can act as if she accepts it for the purposes of the conversation. If she goes for the first option, B will stop accepting that A accepts φ. If she goes for the second option, she will actually come to accept φ for the purposes of the conversation. (Notice that here, once more, the particular nature of acceptance and its divergence from belief is crucial to make the point.)

Now, our claims (A0) and (A1) are just instances of schemas (i) and (ii). We established above that (A0) and (A1) are going to hold with respect to any 'normal' conver-

[37] Of course, this is not to say that A will have explicit knowledge of (A0), or that she would be able to report that she accepts (A0) for the purposes of the conversation. Knowledge and acceptance of the principles discussed in this section may be implicit.

sation. Hence, in any normal conversation, speakers are faced, at any point, with the alternative of denying that they accept that there is a norm they ought not to rule out, or going along with it. If they make the former choice, then they will essentially opt out of the conversation. In that case, in fact, (A0) will stop holding and communication will lose its point. Hence, if the conversation goes on, then they effectively accept that there is a norm they ought not rule out.

In symbols, by running the reasoning for both speakers, we get the following two claims:

(A2) $\mathsf{Acc}_A(\exists n\, \mathsf{Ought}(n \in Acc_A))$.

(A3) $\mathsf{Acc}_B(\exists n\, \mathsf{Ought}(n \in Acc_B))$.

These say that each participant in a conversation accepts that there is a norm that she ought not rule out. To get CAN GO WRONG from here, we still have to show that they all accept that there is a *unique* norm that they *both* ought not rule out.

The final step. The argument for this conclusion is not dissimilar from the argument we gave for (A0). Suppose that it's not true that speakers accept that there is a norm they both ought to accept. For example, suppose that B doesn't accept that. By the previous argument, we know that B accepts that there is a norm that she ought not rule out: call that norm 'n^*'. According to B, it might be that A is allowed to rule out n^*. This allows for a situation of the following kind. Suppose that A utters a normative claim: for example, 'Tax evasion is wrong'. Now, B may accept that A's assertion is correct in a number of respects: it is sincere, it doesn't violate any constraints on A's attitudes, etc. Nevertheless, since B doesn't accept that she and her interlocutor are coordinating on the same norm, she has no reason to assent to A's assertion. It might be that the proposition expressed by 'Tax evasion is wrong' is okay to accept for A, but not for B—i.e., it might rule out n^*, the norm that B ought not rule out.

If speakers don't accept that there is a unique norm that they ought not rule out, they cannot trust each other. It might be that one utters a sentence that is perfectly okay by her standards, but is unacceptable for the hearer. But trust is essential for successful communication. (Famously, Lewis 1975 even argued that a convention of trust, intended as taking-as-true-in-a-language, is part of the basic structure of our communicative practices.[38]) Hence, if speakers want to engage in communication with each other, they will accept that there is a unique norm they ought not rule out.

Taking stock, this means that all speakers involved in conversation will accept the following:

[38] Lewis (1969) did without a convention of trust, but the general account he proposes is in the very same spirit.

(A4) $(\exists n)(\forall s)\mathsf{Ought}(n \in Acc_s)$

From here, it's a short step to claim that (A4) has common ground status. After all, we did not need any specific assumptions about what speakers happened to accept in order to establish (A4). Thus, if common belief in minimal rationality is in place, it will be common belief, at least tacitly, that (A4) is accepted by all participants in the conversation. This gives us exactly CAN GO WRONG, which we repeat below:

CAN GO WRONG. The following is common ground in any conversation:

$$(\exists n)(\forall s)\mathsf{Ought}(n \in Acc_s)$$

Summary. It's time to take stock. We have bootstrapped our way into our conclusion from a number of intermediate premises. In essence, we have claimed, first, that while engaging in conversation each speaker must assume that 'not anything goes' for her interlocutor. From here, we have argued, exploiting the peculiar public nature of acceptance, that each speaker must assume that 'not anything goes' for her own attitudes as well. Finally, we have argued that, if speakers have to trust each other, they must assume that 'not anything goes' for them in the same way—essentially, that the normative constraints applying to all of their attitudes are the same.

The main assumptions we have made are that acceptance is a public attitude and that speakers need to trust each other to engage in communication. Both seem to us very minimal requirements. If this is right, then a minimal objectivity requirement—i.e., CAN GO WRONG—is in place with regard to all normal conversations.

4.5 Establishing UNIQUENESS

To get our final result, we need one more claim. This is the claim that speakers presuppose that it ought to be the case that their acceptance states are shrunk to a unique point:

UNIQUENESS. The following is common ground in any conversation:

$$\mathsf{Ought}((\exists n)(\forall s)Acc_s = \{n\}).$$

UNIQUENESS may seem more controversial than it is. At first sight, you might worry that it enforces a requirement of objectivity that is even stronger than that posed by CAN GO WRONG. But notice the difference in the scope of the operators. UNIQUENESS doesn't say that there is a norm on which acceptances ought to converge; it just says that speakers ought to converge on some norm or other. Hence UNIQUENESS amounts to no more than the claim that, when engaged in conversation, we take for granted that we ought to agree on one among a number of live alternatives (and that this is common

belief between us). But now, it seems a truism that the point of conversation is to achieve consensus as to which of the live alternatives is the one we ought to coordinate on. Hence we take UNIQUENESS to be uncontroversial.

Let us address two worries. First, one might complain that speakers normally don't hold that all uncertainty ought to be eliminated. They can opt out of conversation for a number of reasons—they're uninterested, or it's too demanding, or it's clear that agreement won't be reached. But this is compatible with UNIQUENESS. We are claiming that UNIQUENESS is something that speakers hold *as long as they are engaged in conversation*. In other words, as long as they think that there is a point to engaging in conversation, they must think that they ought to converge on some live possibility. Converging on some live possibility is just what the point of conversation is. Speakers may give up on UNIQUENESS, but by doing so they give up on the very point of staying in a conversation.

The second worry is that UNIQUENESS is excessively fine-grained. Do speakers really assume that they ought to narrow down the context set to a unique element? Once more, the distinction between full-blown worlds and norms on the one hand, and small worlds and norms on the other, is relevant here. We agree that it's unrealistic to assume that ordinary speakers take up the task of narrowing open possibilities to a point in logical space. But we are not making this assumption. All we need is that, in every context, speakers take up the task of narrowing down the context set to one of the elements of the partition determined by the questions they take as relevant. This much seems perfectly acceptable. By raising a question, whether implicitly or explicitly, a speaker seems to demand that participants in the conversation make a choice between the available answers. On the small worlds construal, UNIQUENESS just amounts to the requirement that all questions under discussion receive an answer.

4.6 Convergence

We have argued for the following two claims:

CAN GO WRONG. The following is common ground in any conversation:

$$(\exists n)(\forall s)\mathsf{Ought}(n \in Acc_s)$$

UNIQUENESS. The following is common ground in any conversation:

$$\mathsf{Ought}((\exists n)(\forall s)Acc_s = \{n\}).$$

These two claims entail our wanted conclusion:

CONVERGENCE. In any normal conversation, the following is common ground:

$$(\exists n)(\forall s)\mathsf{Ought}(\{n\} = Acc_s).$$

The step from CAN GO WRONG and UNIQUENESS to CONVERGENCE is straightforward. Intuitively, the argument is just this. CAN GO WRONG establishes that there is one norm that ought not be ruled out from the context set. UNIQUENESS establishes that the context set ought to be reduced to some norm or other. The only way to satisfy the two desiderata is that the context set be shrunk exactly to the one norm that ought not be ruled out.[39]

4.7 Convergence and Error Theories

We have argued that if UNIQUENESS and CAN GO WRONG are in any normal context set, so is CONVERGENCE. This is not to say that CONVERGENCE is true. Indeed, it's not clear how the question of truth and falsity of normative claims like CONVERGENCE should be handled by an expressivist theorist. But we need not make any commitments on this. Our point is simply that speakers' endorsement of CONVERGENCE is a precondition on their engaging in communication about normative claims.

Before closing the section, we want to emphasize a consequence of our conclusion. The model of communication that results from our argument doesn't require any kind of pretense or error theory. As we pointed out at the beginning of the section, the crucial presupposition is normative. Hence we're not claiming that speakers need to make an assumption that is false by the expressivist's lights (something in the ballpark of there being normative facts) to proceed in conversation. The assumption they must make is functionally analogous. But crucially it doesn't commit them to some false claim about what the world is like.

This seems a significant advantage of the proposal we end up with. To see this, let's compare it briefly with a somewhat similar picture of communication for nonfactual discourse recently proposed by Andy Egan (2007; 2010), in the context of defending a form of relativism.

Egan endorses a form of truth relativism on which the asserted contents of certain claims, for example claims about personal taste, are centered world propositions, i.e. sets of triples of a world, a time, and an individual. On this picture, the content of

(5) Eggplants are tasty.

is the set of triples $\langle w, t, i\rangle$ such that i has (at t in w) the disposition to like eggplants (Egan, 2010, p. 259ff).

[39] Here is a more detailed proof. Assume, with CAN GO WRONG, $(\exists n)(\forall s)\mathsf{Ought}(n \in Acc_s)$. Let n^* be such an n; we get that $(\forall s)\mathsf{Ought}(n^* \in Acc_s)$. Then assume, with UNIQUENESS, that $\mathsf{Ought}((\exists n)(\forall s)Acc_s = \{n\})$. Finally, for *reductio* assume that CONVERGENCE fails, i.e. assume that $(\forall n)(\exists s)\mathsf{May}\neg(\{n\} = Acc_s)$. Instantiating the last claim with n^*, we get that $(\exists s)\mathsf{May}\neg(\{n^*\} = Acc_s)$. Let s^* be the relevant speaker: we get $\mathsf{May}\neg(\{n^*\} = Acc_{s^*})$ Given CAN GO WRONG, this means that s^*'s acceptance state can permissibly be a (proper) superset of $\{n^*\}$. But UNIQUENESS requires that $\mathsf{Ought}((\exists n)Acc_{s^*} = \{n\})$, i.e. that s^*'s acceptance state be a singleton. Contradiction.

Since contents are more fine-grained than sets of worlds, Egan has a problem analogous to that of the expressivist. He must explain how assertions of claims about taste work, given that he can't appeal to the standard idea of locating an actual point. He does this by imposing a constraint on the felicity of assertions: asserting a centered-world proposition requires presupposing that all speakers in the conversation are similar in relevant respects. For example, a felicitous assertion of (5) requires that all speakers presuppose that they're disposed to have a similar response of enjoyment or distaste towards eggplants.

This proposal seems to run into trouble rather quickly. After all, (5) seems assertable even in cases where it's entirely manifest that one's audience doesn't share one's dispositions ("What? You don't like eggplants? You're crazy! Eggplants are just *so* tasty!").[40]

An account that requires a factual presupposition of similarity to be in place is hard-pressed to explain cases like this, since this presupposition flatly contradicts information that is built into the context set.[41] To be sure, this does not show that there is no property P such that (i) the content of an utterance of (5) is the set of centered worlds in which i has property P at w (at t) and (ii) the presupposition that participants in a conversation are alike with respect to P is compatible with the recognition of differences with respect to suitably related properties. For example, it could be that the content of (5) is the set of centered worlds such that i at t is such that, suitably idealized, i would have the disposition to enjoy eggplants—one can clearly maintain that all parties to a conversation are alike with respect to *that* property while recognizing that they are not alike with respect to related properties, like the property of being disposed to enjoy

[40] Note that this worry is different from that of whether disputes over (5) are ever worthwhile. Egan rightly points out that one can be mistaken about whether one is disposed to enjoy eggplants. Suppose Alice and Barbara are in conversation. Alice utters (5) and Barbara utters its negation. On Egan's view, Alice can take Barbara to have been sincere while still thinking she is wrong. For Alice can believe (and be reasonable in so doing) that Barbara is mistaken about her own dispositions to enjoy eggplants. Our point is that Alice can explicitly recognize that Barbara lacks that disposition and continue to insist that eggplants really *are* tasty. Thanks to an anonymous referee for pushing us to clarify this point.

[41] As Lewis (1979a) points out, it is customary to bring about a presupposition in the context set by making an assertion that requires it. This is the so-called phenomenon of *presupposition accommodation* (see footnote 13). Note however that, in this case, the presupposition could not just be added to the context set, for it contradicts other information that is already in the context set. Admittedly, this worry should not be overstated. One can sometimes add something to the context set that contradicts information that was already there, so we will need some story about a repair mechanism to supplement the Stalnakerian framework. Why can't accommodation also make use of a similar repair mechanism? Still, while it may be possible to trigger accommodation with information that conflicts with the context set, it seems odd to do utter in one breath two sentences one of which presupposes something that is incompatible with the other. On Egan's picture, an utterance of 'I see you do not like eggplants, but they really are tasty' would be like an utterance of 'There is no King of France, but the King of France is bald'. Thanks here to an anonymous referee.

eggplants, etc.[42]

Whether such an account can yield the right predictions is yet to be determined. Even if it does, we are still being forced to make controversial assumptions about the level of conceptual sophistication one would need in order to believe that eggplants are tasty. In contrast, our model has no problems with cases like this. On our model, just engaging in conversation is enough to bring about the presupposition that there ought to be a point of convergence between speakers. While it might be evident that convergence is hard to achieve in some cases, this doesn't contradict other bits of common ground information.

5 Conclusion

Expressivism about normative discourse runs into a problem about communication. The expressivist wants to see communication as a form of coordination in attitudes among speakers. But it is not clear how the expressivist can explain why the communication of normative claims takes the form that it does. The descriptivist model of communication makes an obvious appeal to truth, and it's unclear how the expressivist can do without it. We argued that, just by assuming certain basic facts about communication and rationality of the speakers, we can answer this challenge. When engaging in normative conversation, speakers need to assume that there is a unique system of norm on which their attitudes ought to converge. From a functional point of view, this assumption works as the normative counterpart of the presupposition that there is an actual world that speakers are trying to locate in conversation.

Let us close with some more speculative remarks. The problem of explaining how communication works is part of a more general problem for the expressivist, i.e., the problem of accounting for the apparent objectivity of moral thought and talk. It's a vexed question exactly what 'objectivity' means in this context (see e.g. Rosen 1994). Following Gibbard, let's just say that there is a sense in which, by making a normative statement, a speaker seems to be claiming an authority of some sort. She's not merely making a suggestion as to how conversational participants could achieve coordination. Rather, she's claiming that her opinion is right in a stronger, intersubjective sense. As Gibbard puts it: "[she] is not merely exposing [her] opinion to public scrutiny; [she] is claiming to be right. [She] claims that [her] opinion is interpersonally valid."[43] One major issue for the expressivist, one that Gibbard takes very seriously, is doing justice to this idea.

Our conclusions in this paper are certainly not enough to address this issue. But we think we have taken some steps in the right direction. If our argument is sound, we've

[42] Cf. Egan, 2010, p. 270f.

[43] Gibbard 1990, p. 155.

shown that engaging with anyone's normative claims requires assuming the existence of a kind of intersubjective normative standard. Part of what it is to engage in normative conversation with an agent is to take her claims to apply some pressure on one's own normative views. This pressure comes not from a factual assumption of a common normative outlook, but rather from the normative structure of conversation itself.

What we have not shown, of course, is how this kind of intersubjectivity can stretch beyond the boundaries of conversations. For example, our account is silent on why Anne and Zoe, who have never met face to face and never been in a conversation together, should take themselves to be subjected to the same normative standard. For all we know, nothing that we've said helps answer this question. But our general strategy—i.e., deriving a kind of objectivity from the structure of our social practices—might be worth exploring as a way of producing a more general vindication of objectivity on behalf of the expressivist.[44]

References

Brandom, R. (1994). *Making it Explicit*. Cambridge, MA: Harvard University Press.

Dorr, C. (2002). "Non-cognitivism and Wishful Thinking". *Noûs* 36.1, pp. 97–103.

Dreier, J. (1999). "Transforming Expressivism". *Noûs* 33.4, pp. 558–572.

Dummett, M. (1959). "Truth". *Proceedings of the Aristotelian Society* 59, pp. 141–162.

—— (1981). *Frege: Philosophy of Language*. Second. Cambridge, Mass.: Harvard University Press.

Egan, A. (2007). "Epistemic Modals, Relativism and Assertion". *Philosophical Studies* 133.1, pp. 1–22.

—— (2010). "Disputing about Taste". In: *Disagreement*. Ed. by T. A. Warfield and R. Feldman. Oxford: Oxford University Press, pp. 247–286.

Geach, P. (1960). "Ascriptivism". *Philosophical Review* 69.2, pp. 221–225.

—— (1965). "Assertion". *Philosophical Review* 74.4, pp. 449–465.

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

Grice, H. P. (1975). "Logic and Conversation". In: *Syntax and Semantics*. Ed. by P. Cole and J. Morgan. Vol. 3. New York: Academic Press, pp. 41–58. Reprinted as Ch. 2 of Grice, 1989, pp. 22–40. Page numbers referred to the reprinted version.

—— (1989). *Studies in the Way of Words*. Cambridge, Mass.: Harvard University Press.

Groenendijk, J. and M. Stokhof (1984). "Studies in the Semantics of Questions and the Pragmatics of Answers". PhD thesis. Amsterdam: University of Amsterdam.

Hamblin, C. L. (1958). "Questions". *Australasian Journal of Philosophy* 36.3, pp. 159–168.

[44] *[Acknowledgments suppressed for blind review.]*

Hamblin, C. L. (1973). "Questions in Montague English". *Foundations of Language* 10.1, pp. 41–53.
Hulstijn, J. (1997). "Structured information states: Raising and resolving issues". In: *Proceedings of MunDial*. Vol. 97, pp. 99–117.
Joyce, J. M. (1999). *The Foundations of Causal Decision Theory*. New York: Cambridge Univ Press.
Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press.
—— (1975). "Languages and Language". *Minnesota Studies in the Philosophy of Science* 7, pp. 3–35. Reprinted in Lewis, 1983, pp. 163–189.
—— (1979a). "Attitudes *De Dicto* and *De Se*". *Philosophical Review* 88.4, pp. 513–543.
—— (1979b). "Scorekeeping in a language game". *Journal of Philosophical Logic* 8.1, pp. 339–359.
—— (1980). "Index, Context, and Content". In: *Philosophy and Grammar*. Ed. by S. Kanger and S. Öhman. Dordrecht: Reidel, pp. 79–100. Reprinted in Lewis 1998, pp. 21–44.
—— (1983). *Philosophical Papers*. Vol. 1. New York: Oxford University Press.
—— (1988). "Statements partly about observation". *Philosophical Papers* 17. Reprinted in Lewis, 1998., pp. 1–31.
—— (1998). *Papers in Philosophical Logic*. Cambridge: Cambridge University Press.
Ninan, D. (2010). "Semantics and the objects of assertion". *Linguistics and Philosophy* 33.5, pp. 355–380.
Perry, J. (1979). "The Problem of the Essential Indexical". *Noûs* 13.1, pp. 3–21.
Price, H. (2011). "Expressivism for Two Voices". In: *Pragmatism, Science, and Naturalism*. Ed. by J. Knowles and H. Rydenfelt. New York: Peter Lang, pp. 87–114.
Rabern, B. (2012). "Against the identification of assertoric content with compositional value". *Synthese* 189.1, pp. 75–96.
Roberts, C. (2012). "Information structure in discourse: Towards an integrated formal theory of pragmatics". *Semantics and Pragmatics* 5.6, pp. 1–69.
Rosen, G. (1994). "Objectivity and Modern Idealism: What is the Question?" In: *Philosophy in Mind*. Ed. by M. Michael and J. O'Leary-Hawthorne. Dodrecht: Kluwer.
Rothschild, D. (2012). "Expressing Credences". *Proceedings of the Aristotelian Society* 112.1pt1, pp. 99–114.
Savage, L. J. (1972). *The Foundations of Statistics*. Second. New York: Dover.
Schroeder, M. (2008a). *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford University Press.
—— (2008b). "How expressivists can and should solve their problem with negation". *Noûs* 42.4, pp. 573–599.
—— (2010). *Noncognitivism in Ethics*. Oxford: Routledge.
Searle, J. R. (1976). "A Classification of Illocutionary Acts". *Language in society* 5.01, pp. 1–23.

Stalnaker, R. C. (1973). "Presuppositions". *Journal of Philosophical Logic* 2.4, pp. 447–457. Reprinted in Stalnaker, 1999.
—— (1978). "Assertion". In: *Syntax and Semantics*. Ed. by P. Cole. Vol. 9. New York: Academic Press, pp. 315–322. Reprinted in Stalnaker, 1999.
—— (1981). "Indexical Belief". *Synthèse* 49.1. Reprinted in: Stalnaker, 1999, pp. 130–149. Page numbers refer to the reprinted version., pp. 129–151.
—— (1999). *Context and Content*. Oxford: Oxford University Press.
—— (2002). "Common Ground". *Linguistics and Philosophy* 25.5, pp. 701–721.
Swanson, E. (2011). "How Not to Theorize about the Language of Subjective Uncertainty". In: *Epistemic Modality*. Ed. by A. Egan and B. Weatherson. Oxford: Oxford University Press, pp. 249–269.
Veltman, F. (1996). "Defaults in update semantics". *Journal of Philosophical Logic* 25.3, pp. 221–261.
Williamson, T. (1996). "Knowing and Asserting". *Philosophical Review*, pp. 489–523.
—— (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
Yablo, S. (2006). "Non-Catastrophic Presupposition Failure". In: *Content and Modality: Themes from the Philosophy of Robert Stalnaker*. Ed. by A. Byrne and J. J. Thomson. Oxford: Oxford University Press, pp. 164–190.
Yalcin, S. (2011). "Nonfactualism about Epistemic Modality". In: *Epistemic Modality*. Ed. by A. Egan and B. Weatherson. Oxford University Press, pp. 295–334.
—— (2012). "Bayesian Expressivism". *Proceedings of the Aristotelian Society* 112.2, pp. 123–160.