Expressivism and the Problem of Communication

Alejandro Pérez Carballo and Paolo Santorio

February 21, 2013

1 Introduction

Expressivists in metaethics think that our moral talk doesn't describe or represent the world. The primary function of moral statements is instead to express nonrepresentational mental states. As is widely acknowledged, 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. Here we will assume that it can. Our question arises once this semantic project has been completed: how can the expressivist account for our practice of asserting normative claims?

Take an example. Anne and Barbara are discussing the permissibility of tax evasion. Presumably Anne and Barbara's conversation is a kind of rational and purposeful activity. Suppose that Anne utters, say, 'Tax evasion is wrong'. She does this because she has an intention to produce in Barbara a certain attitude about tax evasion and believes that this is a good way to fulfill her intention. Similarly, upon hearing Anne's utterance, Barbara normally recognizes her intention and acts accordingly: either she will come to have the relevant attitude or she will challenge Anne's claim. Can the expressivist account for these simple facts about normative conversations?

Consider an analogous question about factual discourse. Zoe and Yael are engaged in a conversation about a soccer game. Zoe missed the second half, Yael missed the first; now each is describing to the other what she saw. How can we explain their conversation as a rational and purposeful activity? Here the answer is easy. Zoe and Yael both want to have an accurate and more complete picture of what the match was like. Yael has information Zoe lacks, and Zoe has information that Yael lacks. They engage in conversation because it is an effective way of exchanging information.

The expressivist cannot rely on this kind of answer. On her view, there is no such thing as moral information. Accordingly, it makes no sense to say that Anne and Barbara want to have an accurate and more complete picture of the morality of tax evasion. On the expressivist's view, there is just no such picture to be had. So how can we account for the exchange betwee Anne and Barbara?

Some expressivists have suggested that we engage in conversation about normative matters in order to *coordinate* our attitudes. This answer is general enough to provide a unified account of both factual and normative talk. Zoe and Yael engage in conversation with one another because they want their beliefs about the game to align. Similarly, Anne and Barbara engage in conversation with one another because they want their moral attitudes to be aligned.

But the answer in terms of coordination cannot 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.

Imagine Anne saying: 'Tax evasion is wrong', and Barbara replying: 'Tax evasion is not wrong'. At this point, they could flip a fair coin: if heads, Barbara will accept whatever Anne does; if tails, *vice versa*. This is a perfectly acceptable way to coordinate. But this is not the way we talk: we don't solve disagreements with coinflips. This is especially not the way we talk about normative matters, which often give rise to the most intense disagreements. Hence speakers are not merely looking to coordinate when engaging in normative talk. If they did, tossing a coin would be enough to settle the matter. Something more than mere coordination must be at stake.

In the case of Zoe and Yael, it's not hard to see what's missing. They want to coordinate by having both of their beliefs align with what the game was actually like. Truth is what puts constraints on coordination in conversations about non-normative matters. 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 the functioning of communication?

This is the question we take up in this paper. In the next pages, we give an answer to it. We take as our starting point 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. The kind of constraints that are in place in normative and non-normative conversations are structurally analogous. Hence the expressivist can explain communication as an appropriately constrained form of coordination.

More precisely, 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.

What the 'minimal assumptions' are, the notion of a presupposition, and the nature of the attitudes at stake will be made clearer in the course of our discussion.

Stating the argument in detail will occupy us for several pages below, but it's useful to give a quick sketch of our strategy.

We start by considering the case of an agent who thinks that 'anything goes' as far as her normative attitudes are concerned—an agent that takes herself to be allowed to accept or reject any normative claim whatsoever. Engaging in communication with such an agent, we claim, would have no point. This agent could grant that there is nothing wrong with, say, your uttering "Tax evasion is wrong", and at the same time still refuse to update her attitudes accordingly. Hence engaging in communication with someone involves assuming that she's not in an 'anything goes' situation. You must assume that the attitudes of your interlocutor are constrained in *some* way or other. We show that, once this minimal claim is established, the more substantial claim we stated above follows: in normal conversations, speakers presuppose that there is a shared normative standard that their attitudes ought to converge on.

The assumption that there is a shared normative standard is something like the normative counterpart of the assumption that there is an actual world. Indeed, the expressivist's problem with communication is solved just by the fact that this shared normative standard plays a role that is functionally analogous to that of the actual world in an account of communication. One might thus worry that this assumption is too strong and sneaks in a kind of factualism in the expressivist's account.

We're going to show that this is not so. Our conclusion still leaves the expressivist's main claims intact. At the same time, the introduction of a normative counterpart of the actual world suggests that expressivism might be closer than expected to truth-conditional theories. In standard accounts of communication, talk of the actual world and talk of truth are easily interchangeable. Hence we should expect to find a way to translate back the expressivist's story about content and communication into a truth-conditional one. We investigate this issue in the last section of the paper. Although we don't have any definitive conclusions, we suggest that expressivism and the assessment-dependent relativism defended by John MacFarlane are, at the very least, extremely close theories and that it's unclear whether they differ substantially in their predictions. We don't regard this as a blow to expressivism. On the contrary, we see the potential convergence of different views into a unique framework as a mark of progress. Hence, we hope that expressivists, as well as relativists, will welcome our conclusion.

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 norm-expressivism, 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 norm-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).

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

[1] In particular, so-called expressivism about epistemic and probabilistic discourse (see, among others, Yalcin 2011, Swanson 2011, Rothschild 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, in Price's terminology, we stick to '20th century expressivism', refraining from taking the leap to global ('21st century') expressivism.

[3] In particular, Gibbard's semantics seems to us the best bet to 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.

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 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. All that is needed is one extra step. So far, we have assumed that descriptive attitudes make distinctions between possible worlds and 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) $\{\langle w, n\rangle$: Eating people is n-permitted in $w\}$

Two 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. We set aside this worry, assuming that Gibbard can find one way out for the problem (or, with one of the authors, that the problem turns out to be spurious on closer analysis[4]). 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

[4] [CITATION OMMITTED FOR BLIND REVIEW]

the contents expressed when (1) and (2) are uttered in a conversation. There are good reasons to think that this is a conflation,[5] but one that is harmless for our purposes.

2.3 Modeling communication

Conversation takes place against a background body of information. Roughly, this body of information involves what is commonly taken for granted for the purposes of the conversation. The purpose of an assertion, on this picture, is to expand this shared stock of information.[6] A successful utterance of a sentence will bring about that the content of that sentence gets added to it.[7]

We can make this picture more concrete by making 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 (1978), we call the set of worlds modeling the background information in a particular context the 'context set'. We also think of the content expressed by a declarative utterance as a set of possible worlds, namely those worlds in which the utterance is true. 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.

A word of warning about the worlds in use in the context set: for the moment, we simply assume that the notion of possible world used here will coincide with the metaphysical notion of a possible world. But notice that nothing hinges on this. In particular, what we count as a world for the purposes of our model could be strictly more coarse-grained than metaphysically possible worlds. We could think of the members of the context sets as *cells of logical space*, rather than points. Following the terminology of Savage (1972),[8] we can call these cells 'small worlds', since each of them will contain strictly less information than a full-blown possible world. If we revert to using small worlds, what speakers try to locate is the 'actual cell' (i.e. the cell containing the actual world) of a salient partition of logical space. This might give us a more realistic picture of what is going on in actual conversations. For the moments we'll stick with possible worlds, but we will come back to this alternative way of construing the Stalnaker picture later on in our argument.[9]

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

[6] This picture is due, in its essentials, to Stalnaker. See e.g. Stalnaker 1973, 1978. See also Lewis 1979b.

[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.

[8] See Joyce 1999 for recent dicussion.

[9] Indeed, we think that this is the best way to implement Stalnaker's model. As he puts it: "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.

Consider a simple example. Zoe and Yael are having a conversation about Tom. They jointly presuppose 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 'Pablo 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.

Gibbard-style expressivists can take on board this general picture of communication.The idea is quite simple. In a conversation about normative matters, speakers try to influence what norms others accept. There are now two parameters they track: roughly, the worlds compatible with what they commonly assume in the conversation, and the norms compatible with their common norm-acceptances.

Go back to Zoe and Yael and assume that their conversation moves to issues of greater significance. Hearing about Tom's behavior, Zoe utters 'Tax-evasion is wrong'. If Yael does not object, this bit of normative content will become part of the conversation's context set. Hence, after Zoe's utterance, the context set will contain only norms that forbid evading taxes. Again, this is captured by enriching the possibilities in the model. Rather than sets of possible worlds, we use sets of *pairs*—for example, for the case of our conversation, world-norm pairs $\langle w, n \rangle$ such that Tom lives in Boston in w and didn't paying taxes *and* n forbids not paying taxes. We can thus predict, as we should, that Zoe and Yael now take for granted, for the purposes of the conversation, that something Tom did was wrong.

2.4 The problem

Hence 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 practice of conversation. They should be able to explain why it is rational for agents to engage in communication in the way they actually do.

Descriptivists are well-placed to give an explanation of this kind. In the model of communication we have assumed, the explanation goes as follow. 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. Hence it's rational for speakers to engage in this kind of practice.

The problem for the expressivist is that she 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

The space might be partitioned differently in different contexts, and there might be no maximally fine partition." (Stalnaker, 1981, p. 136, emphasis added). Cf. also Stalnaker 1988, p. 152.

thing. Hence the expressivist owes us a story about communication. What purpose do speakers have when engaging in normative conversations? And why is it rational for them to do so?

No account has been given in full detail. Expressivists have gestured toward a way of making sense of communication, one 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 accepting a norm—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)

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.

We agree with this basic suggestion, but more needs to be said to get a substantial account. A vast amount of human practices are the result of coordination, yet they 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.

This gap cannot be filled easily. 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 care *which* side that is. Perhaps the case of communication is different because agents approach communication with a number of initial beliefs and acceptances. 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, besides achieving coordination, agents care about maximizing the amount of their own initial beliefs and acceptances that is preserved in the final stage of conversation.

While this is obviously on the right track, 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 do engage in bargaining of this sort at times. Suppose that you and a friend care about spending the evening together, but have a number of conflicting preferences about where to meet, what food to eat, and what to do with your time. In that situation, it would not do to just flip a coin to settle each of the details of your plan. There would 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.[10]

Thus we grant that coordination should play a part in a general account of communication, one that makes sense of both normative and non-normative discourse. 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. In the next section, we look at the ways in which this idea of communication as coordination needs to be developed if it is to explain why communication works the way it does.

3 Explaining communication

In this section we set out to solve the problem posed in the previous section: we spell out what conditions must be in place for speakers to be able to purposefully and rationally engage in communication.

[10] 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 the main lesson of Dorr, 2002—though, unlike Dorr, we don't take this disanalogy to doom expressivism.

One way to pursue the task would be to start from a notion of coordination, following Gibbard and Yalcin, and add extra constraints until we get an account that captures the actual functioning of our practices. We choose a different route. We examine the assumptions agents need to make in order to purposefully and rationally engage in communication about normative matters.

Essentially, 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 set, it is accepted that speakers are required to converge on a unique system of norms, 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 jointly presuppose that there is an actual world: hence they jointly 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' system of norms—a system of norms to which, at least ideally, their acceptances ought to converge.

More precisely, this is the claim we're going to argue for, which we label 'CONVERGENCE' (or 'C' for short):

(C) In any 'normal' conversation, it is presupposed that there is a (unique) system of norms on which the participants' acceptances ought to converge.

In particular, we will argue that (C) follows—surprisingly, perhaps—from some minimal assumptions about the rationality of agents engaging in communication and our practice of assertion.

A few remarks are in order. First, the notion of acceptance in (C) 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, due to the fact that this second notion emerged, quite independently from Gibbard's work, in a separate literature. Roughly, acceptance is the attitude of taking a proposition for granted for the purposes of a conversation (more in a few paragraphs). It has been developed in a setting quite different from metaethics (specifically, in accounts of communication, in particular by Stalnaker 2002). The notion of presupposition we use is also importantly related to acceptance: presuppositions are propositions that agents believe are taken for granted in the conversation.[11]

Second, (C) does not amount to a capitulation to a form of realism or objectivism. A realist will presumably endorse the following thesis:

[11] More precisely, and following Stalnaker (2002): an agent presupposes p just in case she believes that p is commonly accepted by all participants in the conversation. Common acceptance involves infinite iteration of mutual acceptance: all participants accept p, they all accept that they all accept p, they all accept that they all accept that they all accept p, and so on.

(C*) There is a unique system of norms (the 'true' system of norms) to which, in any 'normal' conversation, participants' acceptances ought to converge.

But (C*) is crucially distinct from (C): besides important scope differences between the operators involved, (C) concerns what is presupposed in ordinary conversation, while (C*) concerns what is actually the case. Also, notice that (C) differs from the following as well:

(C**) There is a unique system of norms such that, in any 'normal' conversation, it is presupposed that the participants' acceptances ought to converge to that system of norms.

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

Third, note that in endorsing (C) the expressivist need not impute any type of false belief to ordinary speakers. In presupposing that there is a unique system of norms to which our acceptances ought to converge we are not, according to the expressivist, taking a stance on what the world is like. So the presupposition in (C) is compatible with a perfectly accurate view on what the world is like (and in particular on whether there are normative facts).

Finally, a word on the qualifier 'normal': by 'normal' conversation (at t) we mean one in which it is presupposed (at t) that there is a point to communication. In other words, one in which it is presupposed that participants in the conversation intend to achieve coordination.

Our argument for (C) is somewhat complex. Before spelling it out in some detail, we take a moment to introduce some assumptions and outline the main moves we make along the way.

3.1 Assumptions and overview of the argument

Here are some assumptions we make for the purposes of the argument.

1. **The formal framework.** Throughout our argument, we help ourselves to a Stalnakerian framework for modeling update and communication. We adopt this framework for convenience and ease of exposition; it's not strictly required to make the argument go through. (In particular, we don't require that speakers have common beliefs or other infinitely iterated attitudes.[12]) What *is* required by our argument is that the attitude we use to characterize the context set is acceptance rather than belief. Let us say more about the latter.

[12] The notion of common belief here is a technical notion: a set of agents commonly believe p just in case they all believe p, they all believe that they all believe p, they all believe that they all believe that they all believe p, and so on.

2. **Acceptance.** Agents enter a conversation with a stock of individual and shared beliefs. Belief, however, is not always the right attitude to explain what goes on in a conversation. Consider the following example:

> Smith and Jones find themselves in a conversation. They both believe (suppose, truly) that, in the far-away past, Smith had an affair with Jones' wife; 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 had an affair with Jones' wife is not the object of common acceptance 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.

3. **Rationality.** We make the assumptions that agents engaged in conversation are rational, at least in a minimal sense. We don't have in mind a formal notion of rationality. We just want to use an intuitive notion which is strong enough to guarantee that agents will leave a conversation if they come to believe that remaining in the conversation is not conducive to achieving their purposes.

Finally, 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 presupposed 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 presupposed that there is the possibility of a mistake—i.e., it is presupposed 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:[13]

CAN GO WRONG. The following is presupposed 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 presupposed 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. In symbols:

UNIQUENESS. The following is presupposed in any conversation:

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

Once UNIQUENESS and CAN GO WRONG are in place, our main claim, CONVERGENCE, quickly follows.

3.2 Establishing CAN GO WRONG

This premise states that, in every conversation, it is presupposed 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. Speakers must presuppose that at least one of the possibilities that they 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 is rather involved, but the basic point 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 A ought not rule out. In other words, the following is the case, according to A:
>
> $$\neg\exists n\mathsf{Ought}(n \in Acc_A)$$
>
> Hence A thinks it's okay for her to accept any normative proposition whatsoever.

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

Notice that, in the 'anything goes' situation, *A* takes herself to be allowed to rule out anything that *B* accepts, even if *A* thinks that there's nothing wrong with *B*'s asserting or accepting it. As a result, the following conversation could take place:

B: Cannibalism is wrong.

A: Nothing wrong with that, yet I still accept that cannibalism is okay.

Notice that *A* is not rejecting what *B* says. She need not find anything wrong with *B*'s assertions of normative claims. But she is not constrained in any way by *B*'s utterances, since she's allowed to accept any normative proposition. Hence she can just stick to her acceptances; *B*'s claims will have no effect on her.

We claim that, in this kind of situation, it would make no sense for *B* to engage in communication. 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.

This is our basic point. In the next paragraphs we show how, once this claim is established, we can bootstrap our way into establishing CAN GO WRONG.

Before that, though, 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. Both agents have no reasons 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 utterances*. 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: I reject that.
A: Do you mean that there is something wrong with what I said?
B: No, there was nothing wrong with your *suggesting* that we do that. Your speech act was not inappropriate in any way. But I refuse to accept the content of your suggestion.

C: The Brits drive on the left-hand side of the road.
D: I reject that.
C: Do you mean that there is something wrong with what I said?
D: No, there was nothing wrong with your *claiming* that the Brits drive on the left-hand side of the road. Your speech act was not inappropriate in any way. But I refuse to accept the content of what you claimed.

B's reply is pedantic but sensible. This shows that it's okay to register *A*'s speech act as successful and at the same time reject her proposal. But the situation is different in the dialog between *C* and *D*. *D*'s reply has a contradictory feeling to it. The intuition is that she must either point out that there's something wrong with *C*'s assertion (it is inaccurate, it is unjustified, etc.), or she must accept it. This shows that the rules attaching to the speech act of assertion do not allow for a gap between finding an assertion appropriate and taking up the content of that assertion.[14]

Hence, while coordination is possible (and indeed it does often happen) in an 'anything goes' situation, communication via declarative utterances is bound to fail in a situation of that sort.

3.3 The argument, in detail

The informal claim that *B* must accept that *A* is not in an 'anything goes' situation can be formalized as follows:

(A0) $\mathsf{Acc}_B\mathsf{Acc}_A(\exists n\mathsf{Ought}(n \in Acc_A))$

In plain English, this says that *B* accepts that *A* accepts that there is a norm *A* ought not rule out.[15] (A0), notice, 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.

We've already stated our argument for (A0), so we won't return to it. But before moving on, let us address an objection. We claim we've established that *B* must accept that it's not the case that *A* accepts that anything goes for *A*:

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

[14]Notice that we are *assuming* that the assertion of normative claims works in this way. 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 general account of assertion 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.

[15]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.

But, the objection goes, our argument only shows that B must not accept that A accepts that anything goes for A:

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

The reasoning is this: if B is convinced that, according to A, anything goes for her, then it's true that it makes no sense for her to engage in communication. But, as long as B *leaves open* the possibility that A takes herself to be subject to some constraints in what she accepts, it makes sense for her to at least *try* to engage in conversation with her.

The reply is that the objection confuses acceptance with belief. 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.[16] Hence she will accept (A0), which gives us the following:

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

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

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) $\quad \mathsf{Acc}_B\mathsf{Acc}_A\varphi$

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

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

[16] 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.

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' conversation. 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) $\mathsf{Acc}_B(\exists n \mathsf{Ought}(n \in Acc_B))$.

Hence 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. Indeed, it has been argued that a convention of trust (intended as taking-as-true-in-a-language) is part of the basic structure of our communicative prac-

tices (see Lewis 1975[17]). 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. This gives us exactly CAN GO WRONG, which we repeat below:

CAN GO WRONG. The following is presupposed 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 their attitudes are the same.

The main assumptions we have used are the notion of acceptance, and in particular its public nature, and the claim 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.

3.4 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 presupposed 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. But now, it seems a truism

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

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 in 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. There are uncountably many worlds and—presumably—uncountably many norms. It's unrealistic to think that speakers assume that we ought to narrow down the context set to one of them.

Obviously, the claim that we ought to narrow our normative options down to one norm is an idealization. But we know how to step back from this idealization. As we mentioned in section 2, all talk of worlds in the Stalnaker picture of assertion can be reformulated in terms of sets of worlds which form a partition of logical space. Similarly, we can reformulate our principles in terms of a partition of the norms that are in the context set. UNIQUENESS then becomes the claim that, in conversation, speakers ought to narrow down the members of the partition that they take to be relevant for conversational purposes to one. This seems perfectly acceptable.

Having shown how we can get rid of the idealization without harm, we will proceed with it. Whether norms or cells are the right objects to use in context set is not our focus, and it will be easier to state everything in terms of worlds.

3.5 Convergence

Let's take stock once more. We have argued for the following two claims:

> CAN GO WRONG. The following is presupposed in any conversation:
>
> $$(\exists n)(\forall s)\mathsf{Ought}(n \in Acc_s)$$
>
> UNIQUENESS. The following is presupposed 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 presupposed:

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

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. [18]

3.6 CONVERGENCE and error theories

Hence, 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 different in that 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).

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. Since contents on the Egan picture are more fine-grained than sets of worlds, Egan has a problem analogous to that of the

[18] Here is a more formal 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 is a 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.

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) Eggplants are tasty.

requires that all speakers presuppose that they're disposed to have a similar response of enjoyment or distaste towards eggplants.

This proposal runs into obvious difficulties, since (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!").[19] 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.[20] By 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 presupposed information.

4 A bridge to truth relativism?

4.1 Introducing assessment-sensitivity

We have argued that, if communication is a rational activity, speakers must assume the existence of a system of norms that (ideally, at least) their normative states ought to converge on. In this section, we investigate some consequences. We want to close by suggesting that our conclusion allows us to see a close connection between expressivism and assessment-dependent relativism of the kind defended by MacFarlane. Although we don't have a watertight argument that the two positions are equivalent, our discussion shows that, once our point about communication goes through, it's very easy to move back and forth between the theories. The differences, if any remain, are rather minor.

In standard semantic frameworks, the notion of the actual world and the notion

[19] One of the authors has uttered this line almost *verbatim* more than one time.

[20] Notice that the problem is not that the presupposition is not in place. Lewis (1979a) points out that it's 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*.) But the problem is that the presupposition would conflict with other information that is already in the context set; hence it's hard to see how we can have a consistent context set.

of truth are interchangeable for many purposes. One way of explaining why speakers make assertions is saying, as we've done, that they want to narrow down the possibilities they regard as candidates for the actual world. Another way is saying that the point of assertion is increasing the stock of truths commonly held by speakers.[21] Here we don't have the space to show how a model of communication cast in terms of context sets and an actual world can be translated into one that makes explicit use of the notion of truth. But we trust that the point is intuitive enough that we can take it for granted.[22] Our question for this section is: what happens to this intuitive bridge between the 'actual' point in the context set and truth, once we switch to the expressivist framework?

As we know, the formal structure of the framework remains unaltered. Conversation takes place against the backdrop of a set of open possibilities. Moreover, our argument establishes that the point of making an assertion can be characterized in an analogous way. Both in the case of descriptive and normative assertions, speakers intend to narrow down the open possibilities to an 'actual' point. The difference is that, from the theorist's point of view, talk of an actual point for the normative case is unwarranted. There is an actual world, but there is nothing corresponding to an 'actual norm'. What consequences does this have for the mapping to truth?

At this point, it's useful to say more about the notion of truth that is in place in standard semantic frameworks. This notion goes under the technical label 'truth at a context' and was characterized in its present form by David Kaplan (1989). On Kaplan's semantic framework, the context of utterance does double duty. First, it contributes to determining what content is expressed by a sentence. Second, it fixes the so-called circumstances of evaluations, i.e. a set of parameters that, in combination with the content expressed, produce a truth-value. So the general form of a clause defining truth at a context is the following:

$$S \text{ is true at } c \text{ iff } [\![S]\!]_c \text{ is true at the circumstances (world, time, etc.) of } c$$

(We use the notation '$[\![S]\!]_c$' to denote the content expressed by S at c.)

If we construe contents as sets of worlds, the circumstances of evaluation fixed by the context correspond just to a world, i.e. the actual world. For example, Anne's utterance of the sentence "Barbara evaded taxes in 2012" expressed the content

$$\{w\text{: Barbara evaded taxes in 2012 in } w\}$$

[21] This seems at least one way of unpacking Dummett's (1958) well-know thesis that truth is the point of assertion.

[22] Let us be clear about one point. Strictly speaking, we can give both a semantic theory and a theory of communication without using a notion of truth at all. This is what a number of semantic-pragmatic packages on the market do: for example, the ones in the so-called dynamic strand. (See e.g. Groenendijk and Stokhof 1991; Veltman 1996.) But, whether our theories of meaning make explicit mention of truth or not, we think that they are all committed to a notion of truth at a deep level. A notion of truth is required to explain what the point of communication is, as the previous section suggested. Explaining communication requires assuming that agents engage in it to discover the truth about the world where they are located.

and it counts as true just in case that content is true as evaluated at the world of the context, i.e. the actual world.

Fitting an expressivist model of communication with a definition of truth at a context in this style is problematic. Assume that our semantics produces as contents sets of world-norm pairs.[23] If we treat the norm parameter on the same footing as the world parameter, the context of utterance will uniquely fix a norm that is relevant for the evaluation of the sentence. Hence we're able to assign, once and for all, a truth value to each utterance of a sentence in a context, including normative claims. This is problematic, for two reasons.

First, it conflicts with one of the main features of expressivism as we've characterized it here. The expressivist does want to capture our ordinary talk of truth as applied to normative discourse. But she also want to deny that, from her own point of view *qua* theorist, normative claims are true, at least in the sense of truth that applies to descriptive statements. If we generalize Kaplanian truth at a context, we have no way to capture this difference.

Moreover, using a Kaplan-style notion of truth is also not a good way to capture the point we've been making in this paper. We have argued that speakers must take for granted that there is a norm they ought to coordinate on. But this falls far short of there being a *real* 'actual' norm in their context set (in relation to this, recall our discussion at the beginning of section 3). By contrast, of course, the context set *does* include an actual world. Using Kaplan-style truth at a context seems to blur this crucial difference.

There is a way to sidestep these difficulties. We can adopt a relativistic definition of truth at a context along the guidelines suggested by by John MacFarlane (see, among many, 2003, 2003, 2005, 2012). Roughly, Kaplan's framework treats sentence truth as relative to some parameters, but utterance truth as absolute—each utterance gets assigned a truth-value once and for all. MacFarlane's apparatus treats *utterance truth* as relative to some further parameters. More accurately, on MacFarlane's view a sentence and a context of utterance are not enough to determine a truth-value. A second context, namely the context where the sentence is assessed for truth and falsity, is also needed. Accordingly, the definition of truth at a context now involves relativization to two contexts rather than one. Assuming that we have a world and a norm parameter, this definition will look as follows:

$$S \text{ is true at } c_U \text{ and } c_A \text{ iff } [\![S]\!]_{c_U} \text{ is true at the world of } c_U \text{ and the norm of } c_A$$

Following MacFarlane, call this definition of truth at a context *assessment-sensitive*. It's useful to illustrate how assessment-sensitive truth works with an example. Suppose that Anne, trying to chastise Barbara, utters "Tax evasion is wrong!". The content she

[23] Again, in this discussion we'll be cavalier about the distinction between semantic values and contents (see footnote 5), since it seems irrelevant in the present context.

expresses is the usual set of world-norm pairs:

$$\{\langle w, n\rangle\text{: Tax evasion is wrong according to } n\}$$

This content produces a truth-value when combined with a world and a norm. The world will be, as usual, the world of utterance. By contrast, the norm is not fixed once and for all. A different norm is provided by different assessors. Presumably, the norm provided by Anne's context declares tax evasion wrong, and hence Anne's utterance is true as evaluated by her. But the same utterance might be false as evaluated by a different assessor—for example, Barbara.

The assessment-sensitive definition of truth at a context is immune to the difficulties described above. The switch to relative truth fits the *desideratum* that the notion of truth relevant for normative matters be different, and less substantial, than the one relevant for descriptive discourse. For similar reasons, the use of relative truth allows us to capture more accurately the conclusion of our argument.

One we have contexts of assessment, it's not hard to see how the content of the 'actual norm' presupposition can be recast in terms of truth. Our conclusion can be restated as follows:

(C_{TRUTH}) In any 'normal' conversation, it is presupposed that all speakers ought to occupy a unique context of assessment at which their normative utterances are evaluated for truth and falsity.

Notice that, again, the content of the presupposition is normative, and hence it will display sensitivity to an assessment-dependent parameter.

4.2 Equivalent theories?

We started from an intuitive mapping between talk of the actual world and talk of truth in theories of meaning. We argued that speakers make a presupposition that their context sets include a normative counterpart of the actual world. Hence, given the mapping between talk of an actual point and truth talk, we should expect that this presupposition can be characterized in terms of truth. The suggestion is that MacFarlane's assessment-dependent notion of truth does the job.

If this is right, it establishes that the expressivist can easily recast her semantics as an assessment-dependent semantics. Her characterization of the contents expressed by normative utterances will remain analogous. To this, she will add a MacFarlane-style definition of truth that allows her to assign truth-values to utterances from the perspective of an assessor. The assignment of truth-values to utterances will come in useful when it comes to explaining the point of asserting normative claims. Speakers aim at asserting what is true at their context of assessment. Given the presupposition we defended in section 3, this will be enough to make communication work.

Hence expressivist and assessment-dependent semantics seem less different than is sometimes thought. In fact, it seems that we can easily move from one framework to the other, to the point that there may be no substantial difference between the two. A few qualifications are in order, though.

First, while assessment-dependence is a semantic view, expressivism is an overall package which pairs a semantics for language with a theory of attitudes. Our claims concern the semantic component of expressivism. Indeed, relativists have been mostly silent about the view of attitudes that should accompany their linguistic claims about assessment-dependence.[24] So we're not attempting a comparison in that respect.

Second, it's not clear that the two theories turn out to be fully equivalent. In particular, we have argued that speakers must *presuppose* that there is a normative counterpart of the actual world. Thus a presupposed actual normative point, or a presupposed unique point of assessment with respect to which contents are assessed for truth, play a role in explaining speakers' behavior. But the theorist herself need not assign truth-values to normative utterances. In the expressivist's overall theory, there is just no use for these truth-value assignments.

Let's illustrate the point with an example. Suppose Anne tells Barbara "Tax evasion is wrong". An expressivist helping herself to relativistic machinery will explain Anne's communicative intention as follows. Anne presupposes that she ought to occupy the same context of assessment as Barbara; she intends to say something that is true as evaluated at this context. But the theorist herself needs *not* assign a truth-value to Anne's assertion from her or his own context. Indeed, the expressivist, as we've characterized her, will want to *refrain* from calling normative utterances true or false. No such issues arise on the relativistic picture. Run-of-the-mill relativism will just predict that Anne's utterances receives a truth-value, which is determined by the theorist's own context of assessment.

Hence, on a first pass, it seems that expressivism and relativism *might* disagree at least on this. The relativist, but not the expressivist, may assign truth-values to normative utterances at her own context of assessment. We leave it to future work to assess whether this difference really obtains and how substantial it is.[25] For present purposes, we're content with having flagged that, once we endow expressivism with a full account of communication, expressivism and relativism turn out to be closer than we would

[24] As is probably clear by now, we are tempted by the idea of a merger of the theories and hence think that the relativist should adopt the expressivist story. But we will let the relativist make up her own mind about this.

[25] Let us just flag an obvious possibility for fully reconciling the predictions of the two theories. The relativist might allow for *gappy* contexts of assessments: i.e. contexts of assessment that leave the value of some parameters undefined. Then she could say that the context of assessment of the theorist formulating a semantics for Anne and Barbara's utterances is gappy in this sense with respect to the norm parameter. Whether this story can and should be defended is a question for the metasemantics associated to relativism, and it goes beyond the purposes of this paper.

have expected at first.

It shouldn't be surprising that this is so. In classical theories of meaning, the notion of truth (and in particular, the formal notion of truth at a context of utterance) works as the joint between a compositional semantics for language and a theory of assertion and communication. So far expressivists, who aim at dispensing with truth for normative utterances, have been quite reticent about how communication works on their picture. It is unsurprising that, once we start digging into the matter, we unearth a notion of truth or, at least, a close surrogate of it.

5 Conclusion

We started by considering the problem of communication that arises for expressivism about normative discourse. Suppose that, as the expressivist wants, we see communication as a form of coordination in attitudes among speakers. It's 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.

In the last section, we have suggested that, once we add this element to the expressivist package, we can very easily move back and forth between an expressivist semantics and a relativistic semantics in the style of John MacFarlane. We see this as a fortunate convergence of two different theoretical frameworks. This convergence suggests that the models at our disposal are moving in the right direction—towards a stable, credible linguistic framework that can back antirealism about a certain kind of discourse.

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."[26] 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 might have moved some steps in the right direction. If our argument is sound, we've 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. 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.[27]

[26] Gibbard 1990, p. 155.

[27] Thanks to John MacFarlane, Agustín Rayo, and Bob Stalnaker for feedback in the early stages of the project, and to audiences at Harvard, MIT, the Australian National University, and the University of Sydney.

References

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

Dummett, M. (1958). "Truth". In: *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". In: *Philosophical Studies* 133.1, pp. 1–22.

Egan, F. (2010). "Computational models: a modest role for content". In: *Studies in History and Philosophy of Science* 41.3, pp. 253–259.

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

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

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

Groenendijk, J. and M. Stokhof (1991). "Dynamic predicate logic". In: *Linguistics and Philosophy* 14.1, pp. 39–100.

Joyce, J. M. (1999). *The Foundations of Causal Decision Theory*. New York: Cambridge Univ Press.

Kaplan, D. (1989). "Afterthoughts". In: *Themes From Kaplan*. Ed. by J. Almog, J. Perry, and H. Wettstein. Oxford University Press, pp. 565–614.

Lewis, D. (1969). *Convention*. Cambridge, MA: Harvard University Press.

—— (1975). "Languages and Language". In: *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*". In: *The Philosophical Review* 88.4, pp. 513–543.

—— (1979b). "Scorekeeping in a language game". In: *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.

—— (1998). *Papers in Philosophical Logic*. Cambridge: Cambridge University Press.

MacFarlane, J. (2003a). "Future Contingents and Relative Truth". In: *The Philosophical Quarterly* 53.212, pp. 321–336.

—— (2003b). "Three Grades of Truth Relativity". Unpublished ms., University of California, Berkeley.

—— (2005). "Making Sense of Relative Truth". In: *Proceedings of the Aristotelian Society (Hardback)* 105.1, pp. 305–323.

—— (2012). *Assessment Sensitivity: Relative Truth and Its Applications*. Unpublished book ms., University of California, Berkeley.

Ninan, D. (2010). "Semantics and the objects of assertion". In: *Linguistics and Philosophy* 33.5, pp. 355–380.
Perry, J. (1979). "The Problem of the Essential Indexical". In: *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". In: *Synthese* 189.1, pp. 75–96.
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". In: *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". In: *Noûs* 42.4, pp. 573–599.
—— (2010). *Noncognitivism in Ethics*. Oxford: Routledge.
Stalnaker, R. C. (1973). "Presuppositions". In: *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". In: *Synthèse* 49.1. Reprinted in: Stalnaker, 1999, pp. 130–149. Page numbers refer to the reprinted version., pp. 129–151.
—— (1988). "Critical Notice of Lewis, OPW". In: *Mind* 97.385, pp. 117–128.
—— (1999). *Context and Content*. Oxford: Oxford University Press.
—— (2002). "Common Ground". In: *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". In: *Journal of Philosophical Logic* 25.3, pp. 221–261.
Yalcin, S. (2011). "Nonfactualism about Epistemic Modality". In: *Epistemic Modality*. Ed. by A. Egan and B. Weatherson. Oxford University Press, pp. 295–334.