How to be an ethical expressivist

Silk, Alex

Link to publication on Research at Birmingham portal

How to Be an Ethical Expressivist[*]

Alex Silk

a.silk@bham.ac.uk

Penultimate draft. Please consult the official version
in *Philosophy and Phenomenological Research*.

Abstract

Expressivism promises an illuminating account of the nature of norma-
tive judgment. But worries about the details of expressivist semantics have
led many to doubt whether expressivism's putative advantages can be secured.
Drawing on insights from linguistic semantics and decision theory, I develop a
novel framework for implementing an expressivist semantics that I call *order-
ing expressivism*. I argue that by systematically interpreting the orderings that
figure in analyses of normative terms in terms of the basic practical attitude of
conditional weak preference, the expressivist can explain the semantic prop-
erties of normative sentences in terms of the logical properties of that attitude.
Expressivism's problems with capturing the logical relations among normative
sentences can be reduced to the familiar, more tractable problem of explain-
ing certain coherence constraints on preferences. Particular attention is given
to the interpretation of wide-scope negation. The proposed solution is also
extended to other types of embedded contexts — most notably, disjunctions.

1 The expressivist program

Expressivism is something of a four-letter word in some circles. ("Is that not triply a lie?") And yet there is something persistently alluring about its research program. Though expressivism was originally developed as a metaethical position about normative language and judgment, its appeal has extended across philosophical disciplines. Expressivist semantics have been given for a diverse class of expressions — not just for normative terms like 'ought', 'wrong', and 'rational', but also for epistemic terms like 'might' and 'probably', for attitude verbs like 'knows' and 'believes', and even for 'means' itself.[1] Expressivism promises an illuminating account of the meanings of such expressions, the nature of judgments involving them, and the connection between such judgments and motivation or action. But foundational worries about the details of expressivist semantics have led many to doubt whether these advantages can be obtained.

On a standard truth-conditional semantics the semantic properties of sentences (e.g., inconsistencies, entailments) are, in the first instance, explained in terms of properties of the contents of those sentences: To a very rough first approximation, the sentences 'Grass is green' and 'Grass is not green' are inconsistent because they have incompatible truth-conditions. The primary explanatory weight is placed on an assignment of contents to sentences and on relations among those items of content. Expressivists take a different tack. Though it is somewhat contentious how best to understand expressivism, I take as my starting point the following familiar characterization. At the explanatory outset, the expressivist attempts to account for the semantic properties of sentences in terms of properties of the attitudes or states of mind that utterances of those sentences conventionally express.[2] Contents or truth-conditions, even if or when they are assigned to sentences, do no work in fundamental explanations of the semantic properties of those sentences. Expressivists can be understood as accepting the following requirement on fundamental

---

[1]See, e.g., HARE 1952, 1981; PRICE 1983; BLACKBURN 1984, 1998; GIBBARD 1990, 2001, 2003, 2012; FIELD 2000; HORGAN & TIMMONS 2006; YALCIN 2007, 2012; SCHROEDER 2008a; SWANSON 2012; MOSS 2012. I will treat expressivism as a thesis primarily about natural language expressions rather than concepts.

[2]See, e.g., ROSEN 1998: 391–392; UNWIN 2001: 62, 72; SCHROEDER 2008b: 576, 580, 586; DREIER 2009: 97. There may be good reasons to think there are positions deserving to be called 'expressivist' that don't proceed in this way (see CHARLOW 2013, SILK 2013b). But since this is a standard way of understanding expressivism, I accept it here for the sake of argument.

semantic explanation:

THE EXPRESSIVIST PROGRAM
>The semantic properties of sentences are to be explained, fundamentally, in terms of properties of the attitudes conventionally expressed by utterances of those sentences.

In the case of ordinary factual sentences, this requirement might not seem overly difficult to satisfy. For example, the meaning of 'Grass is green' would be explained in terms of the belief that grass is green, or what it is to bear the belief relation toward a certain representational content. And the inconsistency between 'Grass is green' and 'Grass is not green' would be explained in terms of the incoherence in both believing that grass is green and believing that grass is not green. There is of course something to be explained here, but this, as they say, is everyone's problem.

Fortunately for us who like a good metaethical puzzle, expressivists standardly deny that all declarative sentences express ordinary factual beliefs. (Probably why expressivist treatments of 'Grass is green' tend not to get a lot of press.) It is with these sentences that expressivism gets its teeth. An *expressivist* about some linguistic expression $\mathcal{E}$ claims (perhaps *inter alia*) that $\mathcal{E}$-sentences don't conventionally represent how the world, narrowly construed, might be, or even where or when one might be located in the world. $\mathcal{E}$-sentences don't conventionally determine ordinary possible worlds propositions; they don't have truth-conditions in the canonical sense. Insofar as $\mathcal{E}$-expressions can figure in valid reasoning and be embedded in complex linguistic environments, the expressivist must give an alternative account of what their meanings are, and how these meanings compositionally interact with the meanings of other expressions to determine the meanings of expressions of arbitrary complexity. This is a technical challenge, but, given THE EXPRESSIVIST PROGRAM, it isn't merely technical. Whatever formal solution is given, the expressivist must provide a systematic way of interpreting the formalism, and of characterizing what states of mind correspond to what formal objects in such a way that explains the semantic properties of the sentences that express those states of mind.

This is the root of the so-called Frege-Geach problem.[3] Many have regarded it as fatal to expressivism. But I am more optimistic. In this paper I develop a novel framework for implementing an expressivist semantics that I call *ordering expressivism*. Drawing on developments in linguistic semantics and decision theory, I argue that ordering expressivism provides a formally and interpretively adequate

---

[3]"So-called" because it was prominently raised by GEACH (1960, 1965), who attributed it to Frege's distinction between content and assertoric force. See also SEARLE 1962.

solution to the Frege-Geach problem. For concreteness I focus on expressivist treatments of normative terms, but the points can be applied to expressivist treatments of other expressions. Following work by Jamie Dreier, I treat the semantic properties of sentences as explained, fundamentally, in terms of coherence constraints on attitudes. I then show how to implement this idea in an existing, general compositional semantic framework. The resulting view addresses complications concerning indecision, indifference, and incomparability, and generalizes to cover normative, non-normative, and mixed sentences. This avoids limitations in previous accounts. Ordering expressivism constitutes a more secure semantic basis for a broader expressivist theory of language and judgment.

The structure of the paper is as follows. I focus primarily on one prominent instance of the Frege-Geach problem: the interpretation of wide-scope negation ("the negation problem"). §2 characterizes the problem. §3 considers three prominent accounts — those by Allan Gibbard, Mark Schroeder, and Jamie Dreier — in terms of which I will situate my proposed solution. §4 motivates and develops the ordering expressivist framework and shows how it can solve the negation problem. §5 extends the proposed treatment of negation to the case of other types of embedded contexts, with particular attention to Mark Schroeder's recent problem of "mixed disjunctions." Ordering expressivism combines the advantages of Gibbard's and Dreier's accounts while avoiding their limitations.

2   The negation problem

The negation problem is the problem of saying in general what attitude corresponds to wide-scope negation — what attitude is expressed by sentences of the form '$\neg\phi$' — and of doing so in such a way that captures how a normative sentence and its negation are logically contradictory, as per THE EXPRESSIVIST PROGRAM.

Let's fill this in. Consider the following normative sentence.

MUST   Alice has to help the poor.

For the expressivist, the meaning of MUST is, in the first instance, to be given, not in terms of its content or truth-conditions, but in terms of the attitude or state of mind it conventionally expresses. For concreteness, say MUST expresses the attitude of requiring Alice to help the poor. Now consider MUST NOT.

MUST NOT   Alice has to not help the poor.

Since the deontic modal 'have to' is the primary operator in MUST NOT as it is in MUST, MUST NOT, the expressivist tells us, expresses the same sort of attitude as MUST. No fancy footwork necessary: MUST NOT expresses the attitude of requiring Alice not to help the poor. The internal negation — the negation within the scope of 'have to' — poses no new challenges.

We have accounts of what attitudes are expressed by MUST and MUST NOT. But what about NOT MUST?

NOT MUST   Alice doesn't have to help the poor. (/It's not the case that Alice has to help the poor.)

In NOT MUST we have an external negation, a negation that takes scope over 'have to'. What attitude should we say that NOT MUST expresses? Because of THE EXPRESSIVIST PROGRAM, not just any answer will do. Since NOT MUST is the contradictory of MUST, the attitude expressed by NOT MUST must be related in the right sort of way to the attitude of requiring Alice to help the poor to reflect the logical inconsistency between MUST and NOT MUST.

First, does NOT MUST express the requiring attitude toward Alice not helping the poor? No. Both requiring ϕ and requiring $\neg\phi$ might be incompatible in such a way as to capture the inconsistency between a sentence and its negation. But as we just saw, the attitude of requiring Alice not to help the poor is the attitude expressed by the internally negated sentence MUST NOT. And MUST NOT and NOT MUST don't have the same meaning or express the same attitude.

Second, does NOT MUST express the lack of the requiring attitude, the attitude of failing to require Alice to help the poor? No. Both requiring ϕ and failing to require ϕ might be incompatible in such a way as to capture the inconsistency between a sentence and its negation. But the attitude of failing to require Alice to help the poor is the attitude ascribed in NOT BELIEVE — with, say, Bert as subject — but the attitude we need to explain is the attitude ascribed in BELIEVE NOT.

NOT BELIEVE   Bert doesn't believe that Alice has to help the poor.

BELIEVE NOT   Bert believes that Alice doesn't have to help the poor.

NOT BELIEVE, but not BELIEVE NOT, is true if Bert has no views on whether Alice has to help the poor.

Third, does NOT MUST express the attitude of permitting Alice not to help the poor. Yes! But because of THE EXPRESSIVIST PROGRAM we aren't off the hook just yet. What we need to explain is how MUST and NOT MUST are *logically* inconsistent. But if we take the attitudes of permitting and requiring as basic, we must stipulate how

they are logically related, i.e., that requiring ϕ is inconsistent with permitting $\neg\phi$. But this is precisely what we need to explain. Sure enough, NOT MUST expresses an attitude of permitting. But the problem is to say how this attitude is logically related to the attitude of requiring, and to do so in such a way that doesn't presuppose an interpretation of external negation.

In sum, we started by saying that the unembedded sentence MUST expresses an attitude of requiring. We then asked what attitude its negation NOT MUST expresses that captures the logical inconsistency between the two sentences. Neither way of inserting a negation into the requiring attitude — failing to require ϕ and requiring $\neg\phi$ — captures the attitude expressed by NOT MUST. So we might posit a distinct attitude, an attitude of permitting, and say that it corresponds to the external negation of 'have to'. Though NOT MUST does express an attitude of permitting, we cannot stop here lest we leave opaque the logical relations between the distinct attitudes of requiring and permitting and thus, given THE EXPRESSIVIST PROGRAM, between the sentences that express them. So, either we capture the inconsistency between MUST and NOT MUST but get the attitudes expressed wrong, or we get the attitudes expressed right but fail to capture the inconsistency. This is the negation problem.[4]

3 A stylized history

A brief recap — or at least a stylized history — of the current state of play will be instructive. For purposes of situating my positive account, I will focus on developments stemming from work by Allan Gibbard and Jamie Dreier, understood in part in the context of work by Mark Schroeder. In this section I will highlight what I take to be the crucial advances in their accounts, and will raise several problems and potential limitations. In §§4–5 I will develop a framework for implementing an expressivist semantics that combines the virtues of Gibbard's and Dreier's accounts while avoiding their costs.[5]

3.1 Enter Gibbard

In his response to the Frege-Geach problem, Allan Gibbard (1990, 2003) introduces an extension of ordinary possible worlds semantics. Rather than treating the con-

---

[4]See especially UNWIN 1999, 2001, GIBBARD 2003, DREIER 2006, 2009, SCHROEDER 2008a,b,c, HORGAN & TIMMONS 2009.

[5]Thanks to Jamie Dreier and an anonymous referee for helpful discussion about how best to situate my positive account.

tents of sentences as (determining) sets of possible worlds, Gibbard (2003) treats the contents of sentences as (determining) sets of pairs of possible worlds and *hyperplans*.[6] A hyperplan is a maximal contingency plan, a plan that, for any occasion for choice one might conceivably be in and for any action open on that occasion, either forbids or permits — either rejects or rejects rejecting — that action on that occasion (2003: 56); it represents the plan of someone who is maximally decided. The content of an attitude or judgment is given in terms of the world-hyperplan pairs — the possible ways things might be factually and normatively — that it rules out.

On first glance it might seem that the model theory itself suffices for a response to the negation problem. One might say the following. The content of **MUST** is a certain set of hyperplans, those that require Alice to help the poor. Since negation means set complementation, the content of **NOT MUST** is the complement set of hyperplans. Since these two sets of hyperplans are disjoint, we have incompatibility of content. So we predict the inconsistency of the sentences. We're home free.

Or not. As Gibbard himself recognizes — along with many interpreters after him — the formal semantics isn't sufficient for an expressivist response to the negation problem. As per THE EXPRESSIVIST PROGRAM, the expressivist cannot help herself to content or truth-conditions, even of a more fine-grained sort, to explain the semantic properties of sentences. Rather, the fundamental explanation of the inconsistency of **MUST** and **NOT MUST** must be in terms of the nature of the attitudes expressed. The expressivist must systematically interpret the formal objects in question by mapping them onto psychological attitudes whose individual natures and interrelations do the explaining of the semantic properties in question.[7]

Then what *is* Gibbard up to? The crucial contribution of Gibbard's account, on my view, is that it *starts* with an existing framework for doing formal semantics — possible worlds semantics — *and then systematically provides that framework with an expressivist interpretation*. The compositional semantics proceeds roughly as usual; the abstract objects that Gibbard uses as semantic values — sets of world-hyperplan pairs — function like sets of possible worlds. What makes Gibbard's use of the framework distinctively "expressivist" is the interpretation he gives for it. Gibbard's expressivist extension of possible worlds semantics thus provides a crucial piece of apparatus for what Simon Blackburn (1988) once called a "fast track" solution to the general Frege-Geach problem. By systematically interpreting elements of a familiar formal apparatus in terms of states of mind, Gibbard offers a general

---

[6] The theory in GIBBARD 1990 is couched in terms of "systems of norms."

[7] *Cf.* UNWIN 2001: 72; DREIER 2006: 221, 2009: 95, 97; SCHROEDER 2008b: 586.

strategy for implementing an overall expressivist semantic theory.

The prospects for Gibbard's account turn on the details of the expressivist interpretations he gives for the various elements of the formalism. For the particular case of negation, one way of interpreting Gibbard is as attempting to solve the negation problem through the completeness of hyperplans (see esp. 2003: 53–59, 71–75). A hyperplan represents the plans of an agent who is fully decided about what to do. For any relevant circumstance and available action, a hyperplan either rejects the action or permits it — that is, either *disagrees* with performing it or *disagrees with disagreeing* with performing it. A hyperplan that fails to require $\neg\phi$ (fails to disagree with ϕ) ipso facto permits ϕ. We can thus interdefine the attitudes of requiring and permitting without external negation: permitting ϕ is defined as *disagreeing* with requiring $\neg\phi$. Negation corresponds to a basic attitude of disagreeing. Roughly, MUST expresses an attitude of requiring Alice to help the poor, and NOT MUST expresses an attitude of disagreeing with requiring Alice to help the poor. These attitudes are plausibly related in the right sort of way as to explain the logical inconsistency between the sentences that express them. Permitting is interdefined with requiring so as to illuminate the logical relations among sentences that express those attitudes.[8]

In §2 we noted that we must distinguish thinking that one doesn't have to do something — thinking its contradictory is permitted — from not thinking that one has to do something — not thinking it is required. I have two worries with Gibbard's account of negation, which stem from two ways in which these attitudes can come apart: indecision and decided agnosticism.

First, human agents aren't hyperplanners. We can be undecided. So, the expressivist must be able to represent a state of indecision, and distinguish it from states of permitting or indifference. The challenge is to do so in a manner amenable to expressivism, i.e., not in terms of agents' beliefs. One must be able to say more than that, fundamentally, indifference between alternatives is believing them to be equally good (in the relevant sense), and indecision about alternatives is having no

---

[8]DREIER (2006: 222–224), SCHROEDER (2008b: 585–586), and CHARLOW (2011b: 241, 265–266) interpret Gibbard's completeness constraint on hyperplans as requiring hyperplans to either require or forbid every available action. I view this interpretation as incorrect. Gibbard is clear that a hyperplan can merely permit an action (permit but not require it); contrary actions can be tied for best according to a hyperplan (see GIBBARD 1990: 88n.3; 2003: 56). Schroeder (2008b: 598n.7) claims that Gibbard's argument that normative terms pick out natural properties relies on the claim that hyperplanners cannot be indifferent. I disagree. The argument relies on there being a natural property P that describes what is permitted by the hyperplan in a situation; but I don't see why this property couldn't be a disjunctive property. In any case, it isn't my aim to be doing Gibbard exegesis. What will be important is that the view described in the main text (which I am attributing to Gibbard) is insufficient.

belief about which is better (in the relevant sense). The worry isn't that Gibbard's formal framework doesn't have enough structure to represent the distinction between undecided and indifferent states of mind. It does. But we need some story about what it is about an agent's state of mind such that the one type of abstract object rather than the other represents that state of mind. Absent such a story, the worry is that any adequate one will be incompatible with expressivism (*cf.* DREIER 2006: 227–228).

Second, even restricting our attention to decided states of mind, Gibbard's way of interdefining permitting and requiring mischaracterizes these attitudes and hence the attitudes intuitively expressed by MUST and NOT MUST. Even if one is fully decided, one can disagree with requiring ϕ because one rejects taking any attitude toward ϕ. One might take neutrality to be the stance to take; one might be decidedly agnostic. Disagreeing with requiring ϕ isn't equivalent to permitting $\neg\phi$, even in the context of a hyperplan. One might reply that the relevant attitude of disagreement is a decided state of mind that is incompatible with agnosticism. But this raises the worry that we may no longer be trading in a familiar general notion of disagreement in attitude. The worry is that this assumes precisely what needs to be explained, namely, that there is a such an attitude and that it needn't be understood in terms of belief.[9]

3.2 Enter Dreier

In response to the problem of distinguishing indifference from indecision, Jamie Dreier (2006; *cf.* 2009: 105–107) proposes to take the attitude of *strict preference* as basic and then define the attitude of indifference in terms of preference. On Dreier's definition, someone is *indifferent* ("defined indifferent") between two options ϕ and ψ iff (*a*) she doesn't prefer ϕ to ψ and doesn't prefer ψ to ϕ, and (*b*) for any option χ, she prefers ϕ to χ iff she prefers ψ to χ and prefers χ to ψ iff she prefers χ to ϕ (2006: 228; *cf.* BROOME 2004: 21). One is indifferent between two alternatives iff one prefers each of them to exactly the same things, and one prefers exactly the same things to each of them. One is undecided between two alternatives iff one isn't indifferent between them and one fails to prefer either alternative to the other. According to these definitions, whereas indifference is a transitive relation, indecision is not. (I might be undecided between eating ten jellybeans and eating a donut, and undecided between eating a donut and eating eleven jellybeans, though I prefer eating eleven jellybeans to eating ten.) So, says Dreier, we can cash out the attitude of

---

[9]*Cf.* UNWIN 2001: 65–67; HAWTHORNE 2002: 176; GIBBARD 2003: 71–75.

indifference, and distinguish it from indecision, in terms of the more basic practical attitude of strict preference.

The crucial advance in Dreier's account, in my view, is its move to interdefine various relevant attitudes in terms of a more basic practical attitude — in Dreier's case, strict preference — the logic of which is used to explain the semantic properties of normative sentences. To see the importance of this move, it can be helpful to situate Dreier's account in the context of Mark Schroeder's (2008a, 2008b) way of framing the negation problem.[10] Schroeder distinguishes two types of ways that states of mind can be inconsistent. States of mind are A-type inconsistent iff they are the same attitude toward inconsistent contents. Believing that ϕ and believing that $\neg\phi$ is an example of A-type inconsistency. States of mind are B-type inconsistent iff they are "*distinct* and apparently *logically unrelated* attitudes toward the *same* content" (2008b: 581; emphasis in original). Requiring ϕ and permitting ϕ is (allegedly) an example of B-type inconsistency. According to Schroeder, expressivists can help themselves to A-type inconsistency but not to B-type inconsistency; helping themselves to B-type inconsistency would be helping themselves to everything they need to explain, namely, how certain attitudes can be inconsistent in a way that doesn't ultimately reduce to their being inconsistent beliefs.

What Dreier's account highlights is that A-type inconsistency isn't the only promising model of inconsistency in attitude. Not all logical relations among attitudes need be explained in terms of logical relations among the objects of those attitudes. If we can interdefine requiring and permitting in terms of preference, then the inconsistency between sentences like **MUST** and **NOT MUST** needn't be explained either in terms of distinct fundamental attitudes or a single attitude toward inconsistent contents. The inconsistency could be explained in terms of familiar coherence constraints on preferences (more on which in §4). Like belief, preference has a logic; combinations of preferences can be incoherent. Suppose you prefer to go back to sleep, but also prefer not to lose your job. If you realize that the only way to keep your job is to get up, then something in your attitudes "has to change," as Dreier (2009: 106) puts it, like how something would "have to change" if you believed that all ravens are black and that Tweety is a raven but see that Tweety isn't black. If normative sentences express preferences, then **MUST** and **NOT MUST** would be inconsistent, the thought goes, because they express an incoherent body of preferences. Dreier himself only sketches how to apply this model, and only for simple cases. I won't ultimately be taking up his apparent suggestions about how to do so. What is im-

---

[10]Thanks to an anonymous referee for encouraging me to frame the importance of Dreier's account in this way.

portant here is that Dreier's proposal brings into a relief a strategy for responding to the negation problem that may avoid the putative problems detailed in SCHROEDER 2008a for expressivist accounts framed in terms of *A*-type inconsistency.

Despite these contributions, I would like to raise three worries for Dreier's account. The first two aren't intended to be decisive. I raise them simply to highlight potential limitations in Dreier's account. It would be preferable if we could develop an expressivist theory that avoided them. The third limitation is more serious.

Call indifference according to Dreier's definition *D-indifference*. Suppose Chip has the following preference structure. He prefers eating vanilla ice cream to being stabbed, he prefers chocolate ice cream to being stabbed, he prefers winning the lottery to eating vanilla ice cream, and he prefers winning the lottery to eating chocolate ice cream. However, he has never even considered the question of how to compare vanilla and chocolate ice cream. So, he fails to prefer chocolate to vanilla and fails to prefer vanilla to chocolate. Chip is D-indifferent between eating chocolate and eating vanilla. Both are ranked below the same alternatives — winning the lottery — and both are ranked above the same alternatives — getting stabbed. But Chip isn't indifferent between vanilla and chocolate; he is undecided between them. So being D-indifferent isn't sufficient for being indifferent. One can be D-indifferent while still being intuitively undecided.

Dreier recognizes that cases like Chip's pose a problem for his definition (2006: 230). He notes that his definition "relies on there being a sufficiently rich field of preference" to exclude treating cases like Chip's with incomplete preferences from representing indifference. But cases like Chip's cannot be regarded as "don't cares" by the theory. They are precisely the sorts of preference structures that distinguish indifference from indecision. (Or, as we will see shortly, they are at least *one* such sort of preference structure.) Two alternatives can be D-indifferent either by being ranked the same, or by failing to be ranked relative to one another but bearing the same relations to every other alternative. Only the former represents indifference. Unless we assume that the preference relation is complete — and thereby rule out indecision from the start — satisfying D-indifference isn't sufficient for indifference.

There is certainly more to be said in reply. I won't press the worry any further here.[11] Suffice it to say that Dreier's account requires that agents have a suitably rich

---

[11]Dreier invokes preferences among lotteries — "objects engineered to enrich the field of preference in just the way we need" (2006: 231) — to motivate that cases like Chip's may be rare enough that conflating indifference and indecision in these cases won't be such a serious bullet to bite. The problem is that lotteries can do precisely what Dreier says: they can *enrich* agents' preferences. Delineating an enhanced prospect, and forcing choices that involve it, can artificially refine an agent's (incomplete) preference structure. It can function precisely to remove the basis for one's indecision.

field of preferences in order to distinguish indifference and indecision. At minimum, it is worth investigating whether we can develop an expressivist theory that avoids this requirement.

A second limitation in Dreier's account stems from a second way in which one's preferences can be incomplete. Lacking preference and not being (D-)indifferent isn't sufficient for being undecided. Suppose you are forced to make a choice about which of your children, *A* or *B*, to save and which to let die. You fail to prefer either alternative to the other. But you aren't indifferent. You treat the lives of each your children as uniquely precious. Being told, for example, that you could get a free box of tissues if, but only if, you saved *A*, wouldn't settle for you what to choose. You prefer saving *A* and getting the tissues to just saving *A* without getting them, but it's not the case that you prefer saving *A* and getting the issues to just saving *B*. Your attitude is "insensitive to mild sweetening," as they say (HARE 2010; DE SOUSA 1974). And yet, intuitively, you aren't undecided about what to prefer. You have considered the question, and, we can suppose, no further reflection would lead you to change your attitude. You decidedly treat alternatives where you save *A* as incomparable to alternatives where you save *B*. If this is right, the expressivist now has three attitudes to distinguish: indifference, indecision, and what we might call "decidedly treating as incomparable."[12]

Even if there are in fact no genuine incomparabilities in values, it isn't impossible to think otherwise. Whether it is rational is more contentious.[13] This may affect whether an attitude of decidedly treating as incomparable is the sort of attitude that needs to be characterized in giving an expressivist semantics. Nevertheless, it would be preferable if the viability of expressivism wasn't held hostage to these debates.

Third, and most importantly, Dreier's account is only directly applied to sentences that express pure practical states of mind. No attempt is made to implement the account of normative language and preference in a general compositional semantic theory. Doing so is non-negotiable. Some have thought that complex sentences combining normative and non-normative clauses raise unique problems. An adequate expressivist treatment of the purely normative fragment of the language must generalize to cover non-normative sentences and mixed sentences as well.

Hypothetical choice can explain actual preference only insofar as the agent is already sensitive to and decided among the relevant options.

[12]Ruth Chang (2002) claims that, in addition to incomparability, understood negatively as the lack of any comparative relation, there is also a fourth positive value relation of "parity" (the other three being "better than," "worse than," and "equally good"). I will put aside any purported differences between incomparability and parity in what follows.

[13]For survey discussion, see HSIEH 2008.

The lesson from this section is this: We need an account that combines the virtues of Gibbard and Dreier. We need a way of integrating a Dreier-style account of inconsistency in attitude into a general, independently motivated formal semantic framework. We need a Gibbard-style fast-track apparatus with a Dreier-style expressivist interpretation. In the remainder of the paper I will develop a positive expressivist account that provides precisely this. This account also avoids the potential limitations in Gibbard's and Dreier's accounts described in this section.

4 The negation problem: A solution

4.1 Ordering expressivism

Given THE EXPRESSIVIST PROGRAM (§1), merely offering a formalism that assigns contents to sentences won't suffice for an expressivist explanation of the semantic properties of normative sentences. Nevertheless I want to take a step back from THE EXPRESSIVIST PROGRAM for a moment and simply examine how we might formally distinguish the various attitudes in question. Perhaps the resulting formalism will suggest new possibilities for expressivists to explore.

Forget the framework of Gibbardian hyperplans for a moment. Normative terms are semantically modal. They concern not what happens in the actual world — or not merely what happens in the actual world — but what happens in certain alternative possibilities. On the consensus best theory in linguistic semantics, modal expressions are analyzed in terms of an ordering semantics.[14] Modals are interpreted as quantifiers over possible worlds. The domain of quantification is set by two parameters: a set of relevant (accessible) worlds and a preorder $\lesssim$ (a reflexive and transitive relation), where this preorder ranks worlds along a relevant dimension. Different choices of accessibility relations and preorders correspond to different readings of modals — e.g., epistemic, deontic, goal-based. A modal quantifies over the accessible worlds that rank (among the) highest in the preorder, i.e., that aren't $\lesssim$-bettered by any other world. Call these worlds the "$\lesssim$-best" worlds. To a first approximation, necessity modals like 'have to' and 'must' universally quantify

---

[14]Equivalently, a premise semantics (LEWIS 1981a). See especially LEWIS 1973, VAN FRAASSEN 1973, VELTMAN 1976, KRATZER 1977, 1981, 1991. For simplicity I will make the limit assumption (LEWIS 1973: 19–20) to ensure that there is a set of most highly ranked worlds. I will sometimes follow common usage among philosophers and use 'order' and 'ordering' to refer to preorders (LEWIS 1973: 48), though mathematicians typically reserve 'order' for specifically antisymmetric preorders, i.e. preorders that don't permit ties. Familiar examples of preorders are the relations of being at least as tall as, at least as clever as, etc.

over the set of $\lesssim$-best worlds, and possibility modals like 'may' existentially quantify over the set of $\lesssim$-best worlds. For example, 'According to the law, you have to pay your taxes' is true iff you pay your taxes in all the relevant worlds that best approximate the legal ideal. Since our concern here is with normative language, hereafter I will assume that $\lesssim$ is a practical normative ranking of possibilities; $\lesssim$ can thus be understood as reflecting the content of a practical normative view.[15]

Before getting all in a huff about how the expressivist can't appeal to such truth-conditions to do fundamental explanatory work, notice that this framework appears to make many of the distinctions we need. We could say that $\lesssim$ requires ϕ iff all the $\lesssim$-best worlds are ϕ-worlds (worlds in which 'ϕ' is true); that $\lesssim$ forbids ϕ iff none of the $\lesssim$-best worlds are ϕ-worlds; that $\lesssim$ merely permits ϕ iff some but not all of the $\lesssim$-best worlds are ϕ-worlds; that $\lesssim$ is indifferent between alternatives u and v iff $u \lesssim v \wedge v \lesssim u$; and that $\lesssim$ treats alternatives u and v as incomparable iff $u \not\lesssim v \wedge v \not\lesssim u$. (We will return to indecision below.)

This formal apparatus suggests the following strategy for the expressivist, given a normative term 'N' and preorder $\lesssim$ that figures in the interpretation of 'N'-sentences:

How to solve the Frege-Geach problem:
1. Interpret $\lesssim$ as some suitable practical attitude.
2. Use of the logic of this basic attitude to capture the semantic properties of 'N'-sentences.
3. Define any other attitudes intuitively expressed by 'N'-sentences in terms of the basic attitude corresponding to $\lesssim$.

This, schematically, is my proposal for how expressivists can solve the Frege-Geach problem. Call an expressivist theory that makes use of this strategy *ordering expressivism*. Ordering expressivism is a *framework* for developing an expressivist theory. It can be implemented in various ways depending on a range of broader linguistic and metaethical issues. What is most important for present purposes is the structure of the solution disclosed by ordering expressivism. For concreteness it will help to fill in some of the details. But I make no claim that the particular theory to follow is the

---

[15]I use expressions like '$\lesssim$-best' to emphasize that 'best' is being used neutrally to refer to the maximal elements of the preorder, regardless of the dimension along which those elements are being ranked. The preorders in question needn't rank worlds along a specifically evaluative dimension. Making use of orderings thus doesn't prejudge the question of whether to go in for a teleological normative theory, or even a maximizing theory.

only adequate way of implementing ordering expressivism — or even that the general ordering expressivist framework constitutes the only way of solving the Frege-Geach problem. Indeed, I welcome the development of alternatives with which the present account may be compared.

What attitude should we say is represented by $\lesssim$ in the formal semantics? Not just any choice of attitude will do. Here are three constraints. First, the attitude must be a *practical* attitude. If we are to capture the practical character of normative language and judgment, the attitude we choose to explain the meanings of normative terms must itself be practical in nature. It must be action-guiding and motivating; it must regulate choice and behavior. Relatedly, second, the attitude must be *expressivist-friendly*. It mustn't require being understood in terms of belief. Third, the attitude must have a *logic*. Given that normative sentences have logical properties, can stand in logical relations, and can figure in valid reasoning, the attitude expressed by normative sentences must impose certain logical constraints.

For concreteness I will follow Dreier in starting with an attitude of preference. But since $\lesssim$ is a preorder, I suggest, naturally enough, that we interpret it as the attitude of *weak* preference, or "preferring at least as much." This choice of attitude has several advantages. First, weak preference is clearly a practical attitude. Its connection with action is well studied in decision theory. Second, weak preference is plausibly expressivist-friendly. The majority view is that it needn't be construed as a belief that one alternative is at least as good as another.[16] Third, there are rich literatures in preference logic and decision theory describing and justifying the logical properties of and coherence constraints on preferences (§3.2).[17] Fourth, there is a tradition in preference logic of treating weak preference as the primitive relation. This independent research provides a promising basis for an expressivist account of the meaning of normative language.

(Some theorists who mark a fundamental distinction between evaluative vocabulary ('good', 'bad', 'beautiful', 'desirable', 'base') and deontic vocabulary ('must', 'may', 'right', 'wrong', 'permissible', 'obligatory', 'forbidden') may wish to say that the former, but not the latter, express weak preferences. Such theorists may recast the ensuing discussion in terms of capturing the logical inconsistency between, say, sentences about what is evaluatively best and their negations. Since this isn't the place to take up issues concerning the relation between deontic and evaluative language and attitudes, for expository purposes I will simply assume that our example deontic

---

[16] For relevant discussion, see, e.g., Smith 1987, Lewis 1988, Broome 1991, 1993, Byrne & Hájek 1997.

[17] Classic references include Ramsey 1928, Davidson et al. 1955, Davidson 1976. For survey discussion, see, e.g., Hansson 2001, Hansson & Grüne-Yanoff 2012.

modal sentences express weak preferences. Theorists with different commitments about specifically deontic language may feel free to recast our discussion in terms of their preferred choice of attitude, assuming it meets the constraints described above.)

An agent's weak preferences inherit the logic of the preorder. Just as we say that ordinary belief states are inconsistent if they aren't representable in terms of a non-empty set of worlds, so the expressivist may say that states of weak preference are inconsistent if they aren't representable in terms of a non-empty preorder. Of course at the end of the day she will need some explanation for why having preferences that aren't representable in this way is incoherent — just like how everyone, expressivists and non-expressivists alike, needs an analogous story in the case of belief. And she must show how the relevant sort of incoherence is of the right kind to underwrite explanations of the logical inconsistency of normative sentences — just like how an expressivist, like Schroeder, who appeals to A-type inconsistency (§3.2) must explain why a given type of attitude (belief, being for) is "inconsistency transmitting," i.e. why it is such that bearing it toward inconsistent contents is incoherent in a way that explains the logical inconsistency of sentences (or, more generally, why logical relations among the objects of the attitude place constraints on the logic of the attitude itself). Accounts of preferential incoherence are well-known, and the literatures on them vast (nn. 16–17). I won't attempt to offer such an account here. Suffice it to say that it is independently plausible *that there are* coherence constraints on preferences, and that these constraints are of the right sort to underwrite fundamental semantic explanations. At minimum we have made progress if we can reduce expressivism's problem with negation to the familiar, more tractable problem in decision theory of explaining why preferences that don't satisfy certain constraints are incoherent.

The expressivist can then use the logical properties of preferences to capture the logical properties of certain normative sentences. For instance, the familiar truth-conditions for **MUST** and **NOT MUST**, along with their ordering expressivist interpretation, are given in (1)–(2), respectively.

(1)   a.   **MUST** is true at w, according to $\lesssim$, iff Alice gives to the poor in all the $\lesssim$-best worlds

           b.   **MUST** expresses an attitude of having all of one's most weakly preferred alternatives be ones where Alice gives to the poor

(2)   a.   **NOT MUST** is true at w, according to $\lesssim$, iff Alice doesn't give to the poor in some of the $\lesssim$-best worlds

           b.   **NOT MUST** expresses an attitude of having some of one's most weakly

preferred alternatives be ones where Alice doesn't give to the poor

So, in response to the negation problem, we can say that **MUST** and **NOT MUST** are inconsistent because their conjunction expresses an incoherent set of weak preferences. It is incoherent to have all of one's most weakly preferred alternatives be ones where Alice helps the poor and some of one's most weakly preferred alternatives be ones where Alice doesn't help the poor. There is no (non-empty) preorder that represents such a body of preferences. Preferential incoherence appears to be the right kind of inconsistency in attitude to explain the inconsistency between normative sentences 'ϕ' and '$\neg\phi$'.

Taking the weak preference attitude as basic, we can define in terms of it various other attitudes intuitively expressed by normative sentences — e.g., requiring, permitting, merely permitting, and forbidding. To a first approximation: A requires ϕ iff all A's most weakly preferred alternatives are ϕ-worlds; A permits ϕ iff some of A's most weakly preferred alternatives are ϕ-worlds; A merely permits ϕ iff some but not all of A's most weakly preferred alternatives are ϕ-worlds; and A forbids ϕ iff all of A's most weakly preferred alternatives are $\neg\phi$-worlds. By defining requiring and permitting in terms of the single basic attitude of weak preference, we can avoid the worries discussed in §2 with taking, say, the attitude of requiring as basic and then defining permitting in terms of failing to require or disagreeing with requiring. The inconsistency between requiring ϕ and permitting $\neg\phi$ — the attitudes intuitively expressed by **MUST** and **NOT MUST**, respectively — is explained in terms of the incoherence of the preferences in terms of which they are defined.

In this way, the expressivist can import developments in truth-conditional semantics for modals, though with some interpretive tweaks. I have shown how she can do so in terms of a standard ordering semantics because of its familiarity. But adoption of this kind of ordering semantics, or any ordering semantics for that matter, isn't essential to the underlying expressivist maneuver. Even if an alternative way of analyzing modals proves to be superior, the expressivist will be able to implement our general strategy as long as there is some particular element in the analysis corresponding to the normative reading of the expression (e.g., the preorder in an ordering semantics), where the structure of this element captures any logical relations among sentences containing normative terms. Given requirements of compositionality and empirically adequacy, it is hard to see how a successful semantics could fail to have this property. Despite its name, ordering expressivism needn't be held hostage to an ordering semantics for modals.

This point may become relevant depending on one's broader views about the semantics of modals and the nature of preference (or whatever practical attitude

is chosen) — e.g., concerning quantificational vs. scalar semantics for modals; relations among comparative and quantitative notions of possibility and probability; information-sensitivity in normative modals; the objects of preference; the relation between exclusionary and combinative preferences; and the justification of coherence constraints on preference. A plurality of implementations of ordering expressivism will be possible depending on one's commitments on these sorts of issues.[18] How the semantics, metasemantics, philosophy of mind, psychology, and decision theory interact and mutually constrain theorizing likely won't be straightforward. Moreover, for all I have said, the prospects of ordering expressivism may vary for different normative expressions and practical attitudes. It isn't trivial that the best philosophical interpretation of the best semantics for a given expression will be consistent with expressivism.

Such complications needn't detain us here. For present purposes, in what follows I will continue to assume a standard ordering semantics for modals coupled with an interpretation of the preorder in terms of weak preference. The preorder on worlds in the semantics can be understood as reflecting preferences among maximally specific possibilities. This coheres nicely with the standard use of maximal states of affairs (point propositions, or maximally consistent sets of sentences, in the limiting case) as the objects of preference in decision theory and preference logic. However, there may be large equivalence classes in the preorder depending on which features of the world are relevant to one's preferences in the given context. Preference relations on non-maximal possibilities may be relevant in the interpretation of other kinds of normative expressions.[19]

---

[18]To take just one example: Many authors in the literature on information-sensitivity have argued that a deontic modal's domain of quantification can reflect what alternatives are best given a relevant body of information (e.g., KOLODNY & MACFARLANE 2010, CARIANI ET AL. 2011, CHARLOW 2011a, DOWELL 2013, SILK 2013a). If this is right, then certain deontic modal claims may be understood as expressing one's derived preferences *given* one's information (beliefs, credences). (More on this in §5.) How one implements this would depend on one's views about the details of the semantics and how much decision-theoretic apparatus, if any, is explicitly encoded (e.g., whether there is an explicit representation of information-independent desires/utilities/values). Thanks to an anonymous referee for raising this issue.

[19]The literature on comparative modal notions may be relevant here (e.g., KRATZER 1981, 2012, PORTNER 2009, LASSITER 2011, KATZ ET AL. 2012; for related discussion on generating orderings on propositions from orderings on worlds, see LEWIS 1973, HALPERN 1997 and references therein). Note that the appeal to preferences on maximally specific possibilities is neutral on whether preferences on maximal vs. non-maximal possibilities are more fundamental. What is important for present purposes is simply *that there are* preferences on maximal alternatives. Though decision theorists standardly treat preference as a relation on (mutually exclusive, possibly singleton) sets of worlds, this needn't hinder the expressivist from utilizing a classic ordering semantics for modals, given the

4.2 Indecision

In §2 we raised the challenge of capturing in an expressivist-friendly manner the difference between indifference and indecision. In §3.2 we considered Dreier's response and noted two limitations. First, it fails to distinguish indifference and indecision in certain agents with impoverished preferences. Second, it fails to distinguish two ways in which one's preferences can be incomplete. Indecision is distinct from having incomparabilities in one's preferences. Given the coherence of thinking that one might find oneself in a practical dilemma, there is a third attitude to be explained: the attitude of decidedly treating as incomparable. In this section I will argue that the expressivist can distinguish these attitudes by showing how they encode different practical dispositions — specifically, different dispositions in choice situations. (To be clear, the aim isn't to *reduce* these attitudes to patterns of choice behavior. Choice behavior is treated as *revealing* attitudes, not as constituting them.)

Start with distinguishing indifference from decidedly treating as incomparable. Much work in revealed preference theory has concerned how to derive a preference order among alternatives by treating choice as basic.[20] Agents are modeled by choice functions which select elements from (finite) sets of options. An agent's choices are treated as revealing her psychological state of mind, her preferences. Traditionally, choosing both alternatives u and v from an option set has been interpreted as revealing indifference between u and v. But an alternative is to interpret such a choice as revealing that no other element in the option set is strictly preferred to u or v.[21] This might be because one is indifferent between u and v, but it also might be because one treats them as incomparable. Nevertheless we can distinguish these two types of attitudes on the basis of further choice behavior. Results in Eliaz & Ok 2006 attest to the fruitfulness of the following approach. The technical details would take us too far afield, but the rough intuitive idea is this: We check if, for all choice situations, the agent chooses u whenever she chooses v and vice versa; we check if u and v are treated identically in all occasions for choice that involve them. If they are, this reveals that the agent is indifferent between u and v. If they are not, she is better regarded as treating u and v as incomparable.[22]

---

systematic availability of reductions of preorders to partial orders (i.e., antisymmetric preorders, or preorders without ties): for any preorder $\lesssim$ on worlds, there is a corresponding partial order on the set of equivalence classes of worlds with respect to $\lesssim$. Thanks to an anonymous referee for pressing me on these issues.

[20] For classic discussions, see Samuelson 1938, Arrow 1959, Richter 1966, Hansson 1968, Sen 1970.

[21] See, e.g., Nehring 1997, Sen 1997, Eliaz & Ok 2006.

[22] More precisely (*cf.* Eliaz & Ok 2006: 66–67): An agent treats u and v as incomparable if both

For example, suppose Alice must choose what beer to buy out of *x*, *y*, and *z*. The relevant factors guiding her choice are taste and smell. As for taste, she strictly prefers *x* to *y* and *y* to *z*. As for smell, she strictly prefers *y* to *z* and *z* to *x*. Since she strictly prefers *y* to *z* with regard to both taste and smell, she will not choose *z* when *y* is an option. But suppose that while Alice is still contemplating her choice Bert swoops in and takes the last of beer *y*. Left with a choice between *x* and *z*, Alice seems to have no way of choosing merely in light of her preferences; after all, *x* is preferred to *z* with regard to taste but *z* is preferred to *x* with regard to smell, and (let's suppose) she has no further preferences. Her choice set for $\{x, z\}$ might thus be $\{x, z\}$ (in the sense that she might subjectively randomize choice between *x* and *z*). However, we can determine that Alice treats *x* and *z* as incomparable and isn't indifferent between them in light of her further choice behavior. For there is some other choice situation that doesn't contain (say) *z* — namely, where the option set is $\{x, y\}$ — with respect to which *x* and *z* aren't treated symmetrically: whereas *x* is choosable out of $\{x, y\}$, *z* isn't choosable out of $\{z, y\}$, the set that results from swapping *x* with *z*, as we have seen. (In terms of the definition in note 22, it is the (a)-condition that is satisfied.)

Now turn to indecision. In theory of choice, there is a requirement that an agent must, when given a set of options, select a subset of that set. Dropping this requirement, by allowing as a possible response the action of positively refraining from choosing, seems to provide one natural way of characterizing indecision. Indecision between *u* and *v* would be revealed by (perhaps *inter alia*) deferring choice between *u* and *v*. It isn't counterintuitive that if I am genuinely undecided between *u* and *v* — e.g., because I have never considered the question of how to compare *u* and *v* — I may, if presented with a choice between them, refuse to make a choice, at least for the time being. After all, I haven't worked out my preferences, and so

---

u and *v* are choosable from $\{u, v\}$ — $c(\{u, v\}) = \{u, v\}$, for choice function *c* — and there exists a (finite, non-empty) set *S* of alternatives that contains *u* but not *v*, such that at least one of the following holds, where $S^* := (S \cup \{v\}) \setminus \{u\}$, i.e., the option set that results from swapping *u* with *v*:

    a. $u \in c(S) \wedge v \notin c(S^*)$
    b. $u \notin c(S) \wedge v \in c(S^*)$
    c. $c(S) \setminus \{u\} \neq c(S^*) \setminus \{v\}$

ELIAZ & OK 2006 prove that choice behavior that satisfies a weakened version of the Weak Axiom of Revealed Preference can be used to derive a unique, possibly incomplete preference relation (whose strict part is suitably rich). The pairs which the agent treats as incomparable in the above sense can be proven to be precisely those pairs which are incomparable relative to this preference relation. ELIAZ & OK 2006, as well as others in the tradition in theory of choice stemming from AUMANN 1962, treat themselves as characterizing indecision (via incomparabilities), but, for reasons discussed in §3.2, I prefer to interpret their results as characterizing the attitude of decidedly treating as incomparable.

any choice, even choosing both u and v, would give a false impression of my state of mind. Indecision between alternatives, as revealed in this way, amounts to a kind of unwillingness even to pick, or a preference to defer choosing among them.

I see two general ways in which this kind of indecision-revealing choice behavior could be used to generate representations of agents' states of mind: one introduces more structure into the representation of normative beliefs; the other introduces more structure into the representation of bodies of preferences themselves. For space purposes I describe only the former option in the main text; I leave discussion of the latter option for an extended note (n. 25). Before proceeding I want to reiterate that the question of how to represent indecision is subordinate for the expressivist to the question of how to characterize indecision. What is primary is our characterization of indecision as a preference to defer choice, as revealed by a refusal to choose. This is what undergirds our expressivist explanation for why a certain agent should count as being in a state of mind represented by the relevant type of formal object. Nevertheless, although questions of how to represent indecision may be of secondary importance, they are important just the same.

Our indecision-revealing choice behavior can be used to characterize a set of preference preorders, representing ways in which the agent could coherently resolve her indecision (complete her preferences) given the rest of her choice behavior. For example, simplifying by only considering two alternatives u and v, the state of mind of an agent who defers choosing between them could be modeled in terms of the set $\{\lesssim_1, \lesssim_2, \lesssim_3\}$, where $\lesssim_1$, $\lesssim_2$, and $\lesssim_3$ are characterized as follows, representing strictly preferring u to v, strictly preferring v to u, and being indifferent between u and v, respectively:[23]

$$\lesssim_1: \quad u <_1 v$$
$$\lesssim_2: \quad v <_2 u$$
$$\lesssim_3: \quad u \lesssim_3 v, v \lesssim_3 u$$

This choice of representation coheres nicely with standard representations of ordinary factual uncertainty. In ordinary possible worlds semantics, belief states are represented as sets of worlds, the agent's "doxastic alternatives," those worlds com-

---

[23] $<$ ("is strictly preferred to") is a strict partial order such that $\forall u, v: u < v \Leftrightarrow u \lesssim v \wedge v \not\lesssim u$. One way of representing the action of failing to choose might be in terms of a partial choice function that is undefined for certain sets of options. I am unaware of developments of this strategy in revealed preference theory. As noted above, indecision, when discussed at all, is represented in terms of incomparabilities, incorrectly in my opinion. This is certainly not the place to solve the logical problem of how to generate (sets of) preference preorders from partial choice functions.

patible with her beliefs.[24] Uncertainty about how the world is is represented by there being a plurality of worlds in one's doxastic alternatives. Becoming more opinionated is modeled predictably, and intuitively, in terms of set intersection: Intersect one's previous doxastic alternatives with the set of worlds characterizing what one comes to believe to yield the representation of one's new belief state. For example, if I am uncertain about whether grass is green, my doxastic alternatives include both worlds in which grass is green and worlds in which grass is not green. If I come to believe that grass is green, my new set of doxastic alternatives will be my old set minus the worlds in which grass is not green. Similarly, an agent's practical state of mind might be represented in terms of a set of weak preference preorders, the agent's "practical alternatives," those preorders consistent with her preferences. Becoming more decided about what to do can be modeled in the same way as becoming more opinionated about the world. Generalizing, an agent's total belief state might be represented as a set of pairs $\langle w, \lesssim \rangle$ of worlds and weak preference preorders.

Accordingly, even if individual preference preorders cannot fail to require ϕ without permitting $\neg\phi$, we can still represent how an agent can fail to think that ϕ is required without thereby thinking that $\neg\phi$ is permitted.

(3)   *A believes that ϕ is required iff for **all** weak preference preorders in A's practical alternatives, **all** the most weakly preferred alternatives are ϕ-worlds (i.e., iff all the weak preference preorders in A's practical alternatives require ϕ).*

(4)   *A doesn't believe that ϕ is required iff for **some** weak preference preorder in A's practical alternatives, **some** of the most weakly preferred alternatives are $\neg\phi$-worlds (i.e., iff some of the weak preference preorders in A's practical alternatives permit $\neg\phi$).*

(5)   *A believes that $\neg\phi$ is permitted iff for **all** weak preference preorders in A's practical alternatives, **some** of the most weakly preferred alternatives are $\neg\phi$-worlds (i.e., iff all of the weak preference preorders in A's practical alternatives permit $\neg\phi$).*

This reflects how failing to believe that ϕ is required is weaker than believing that $\neg\phi$ is permitted. The former holds only if some weak preference preorders consistent with one's preferences are a certain way, whereas the latter holds only if all the weak preference preorders consistent with one's preference are that way. Similarly, one decidedly treats two alternatives as incomparable only if they are incomparable in

---

[24]See Hintikka 1969, Lewis 1981b, Stalnaker 1984, 1999 for classic discussions.

all the weak preference preorders in one's practical alternatives.[25]

In sum, I have suggested that the expressivist can distinguish the attitudes of indecision, indifference, and decidedly treating as incomparable in terms of differences in choice behavior. Indifference among alternatives is revealed by treating them identically in all relevant occasions for choice. Treating them as incomparable is revealed by not treating them identically in this way but still choosing all of them when presented with all and only them. Indecision is revealed by deferring choice. It remains to be seen how best to develop these ideas more comprehensively. Nevertheless I take them to constitute a promising expressivist strategy. This strategy avoids problems and limitations of previous approaches, like Dreier's, as discussed in §3.2. It is also independently interesting and attractive.

---

[25] An alternative way of representing indecision is by enriching our representation of bodies of preferences. Suppose we represent a body of weak preferences with a set of preorders rather than a single preorder. Intuitively, we might think of each preorder in the set as representing a partial strategy that the agent accepts, and the set of preorders as representing the agent's overall plan for action. A strategy, in the decision-theoretic sense, assigns a single act to every possible occasion for choice. It represents what the agent takes to be a permissible course of action given her preferences. One could then say that ϕ is required according to a overall plan — the set of preorders — iff for all preorders in the set, all the most weakly preferred alternatives are ϕ-worlds; and that ϕ is permitted according to the overall plan iff for some preorder in the set, all the most weakly preferred alternatives are ϕ-worlds. This choice of representation entails that an overall plan can fail to require ϕ without thereby permitting $\neg\phi$; 'required' and 'permitted' won't be duals: It's not the case that ϕ is required iff for some preorder in the set, *some* of the $\lesssim$-best worlds are $\neg\phi$-worlds; but $\neg\phi$ is permitted iff for some preorder in the set, *all* the $\lesssim$-best worlds are $\neg\phi$-worlds. And an overall plan can be undecided about a single proposition without merely permitting it: ϕ is merely permitted iff for some but not all of the preorders in the set, all the $\lesssim$-best worlds are ϕ-worlds; but the overall plan is undecided about ϕ if, e.g., for none of the preorders in the set, all the $\lesssim$-best worlds are ϕ-worlds. On this representation, a subject's practical state of mind has the same structure as a body of preferences. Indecision is encoded in the representation of preference itself.

A brief example may be helpful. Suppose my overall plan Π consists of the following partial strategies, represented as sets of propositions: $\Pi = \{\{$*I eat a cookie, I give to the poor*$\}, \{$*I eat a brownie, I give to the poor*$\}\}$. (These sets of propositions can determine preorders in the usual way: given a set of propositions P, for any worlds u and v, $u \lesssim_P v$ iff all propositions in P that are true in v are also true in u.) On this overall plan, my eating a cookie and my eating a brownie are both merely permitted, since some but not all of my partial strategies entail that I take the action in question; my giving to the poor is required since all of my partial strategies entail that I give to the poor; and I am undecided about all other actions (not entailed by these), since they aren't entailed by any of my partial strategies. So, unlike hyperplans, overall plans in this sense needn't be complete (or consistent: incomparabilities may be represented by all of one's partial strategies entailing contrary propositions).

5 Beyond negation: The disjunction problem

Let's recap. Ordering expressivism combines the virtues of Gibbard and Dreier in the manner called for at the end of §3. Following Dreier, semantic properties of sentences, like inconsistency, are explained, fundamentally, in terms of (non-A-type) coherence constraints on attitudes. Given the counterintuitive consequences of expressivist accounts framed in terms of A-type inconsistency canvassed in Schroeder 2008a, it would be promising if an alternative way of solving the Frege-Geach problem was available. I take ordering expressivism to be one such alternative. Ordering expressivism takes a single practical attitude to be explanatorily basic — e.g., weak preference — and defines the attitudes intuitively expressed by various kinds of normative sentences in terms of it. The logical relations between the "*apparently* logically unrelated attitudes" of, e.g., permitting and requiring are illuminated by defining both attitudes in terms of weak preference (Schroeder 2008b: 581; emphases altered). Even if positing brute inconsistency relations between distinct attitudes is problematic, explaining inconsistency relationships between attitudes in terms of the logic of a more fundamental attitude is not. A-type inconsistency isn't the only legitimate kind of inconsistency in town.

Like Gibbard in his extension of possible worlds semantics, ordering expressivism co-opts independently motivated apparatus from formal semantics on modals, but gives it a systematic expressivist interpretation. A concern some have with expressivism is that it seems to require us to rebuild compositional semantics from the ground up. It would be unfortunate, to put it mildly, if we had to reexplain everything that contemporary truth-conditional semantics has taught us. An advantage of the general ordering expressivist strategy in §4.1 is that it allows expressivists to import developments in truth-conditional semantics for modals in giving her semantics for normative terms. This provides a recipe for giving an expressivist account of complex normative sentences, and recovering the state of mind expressed by a sentence from its compositional semantics: interpret all occurrences of $\lesssim$ in the semantics in terms of the practical attitude associated with $\lesssim$. To a first approximation, some examples with conjunction, quantifiers, and conditionals are as follows. The (a)-examples give the ordinary truth-conditions; the (b)-examples give the attitude expressed. (In (7) I give the reading where 'most' takes wide scope.)[26]

---

[26]How one develops an expressivist treatment of deontic conditionals will depend on one's broader views on the syntax and semantics of different types of implicitly and explicitly modalized conditionals. These issues, notoriously complex as they are, are complicated further by the fact that conditionals are themselves constructions that have been subject to expressivist treatments. I offer the interpretation in (8) simply for illustrative purposes.

(6)   a.   'Alice has to give to the poor, and Bert has to give to the poor' is true at w, according to $\lesssim$, iff Alice gives to the poor in all the $\lesssim$-best worlds and Bert gives to the poor in all the $\lesssim$-best worlds.

        b.   'Alice has to give to the poor, and Bert has to give to the poor' expresses the attitude of having all of one's most weakly preferred alternatives be ones where Alice gives to the poor, and having all of one's most weakly preferred alternatives be ones where Bert gives to the poor.

(7)   a.   'Most people have to give to the poor' is true at w, according to $\lesssim$, iff most people are such that they give to the poor in all the $\lesssim$-best worlds.

        b.   'Most people have to give to the poor' expresses the attitude of having all of one's most weakly preferred alternatives be ones where, for most (actual) people x, x gives to the poor.

(8)   a.   'If Alice has a job, she has to give to the poor' is true at w, according to $\lesssim$, iff Alice gives to the poor in all the $\lesssim$-best worlds where Alice has a job.

        b.   'If Alice has a job, she has to give to the poor' expresses the attitude of having all of one's most weakly preferred alternatives where Alice has a job be alternatives where Alice gives to the poor.

Perhaps there are other types of normative sentences whose semantic properties the expressivist cannot capture in this way. The devil will be in the details. But these examples should give the expressivist a license for optimism.

Perhaps not for long. Mark Schroeder has objected that "expressivists cannot directly apply the tools of truth-conditional semantics" (2011: 9). Schroeder's Exhibit A: disjunction. In the remainder of this paper I will focus on Schroeder's objection concerning "mixed disjunctions," disjunctive sentences where one disjunct expresses a pure belief state and the other disjunct expresses a pure practical state.

(Since Schroeder's objection is partly technical in nature, some parts of this section are, of necessity, a bit more formal than the others. To keep the discussion generally accessible, I only mention those technical points that are directly relevant to the philosophical issues. The important philosophical points should be clear enough, I hope, even for readers less interested in the technical details.)

As noted in §3.1, expressivists can implement their compositional semantics with an extension of possible worlds semantics, treating the contents of sentences as (determining) sets of pairs of worlds w and weak preference preorders $\lesssim$. But because of THE EXPRESSIVIST PROGRAM, she must provide a systematic way of interpreting the formalism and characterizing what states of mind correspond to what

formal objects in such a way that explains the semantic properties of the sentences that express those states of mind. In the special cases of "preference-independent" and "world-independent" contents — sets of pairs such that if one pair is in the set, so is every pair with the same w or $\lesssim$ coordinate, respectively — the expressivist has a story: The former characterize ordinary factual beliefs, like the belief that grass is green, and the latter characterize preferences, like the belief that murder is wrong. But what should we say about contents that are neither preference-independent nor world-independent? Schroeder objects that the expressivist cannot give an adequate general answer to this question, and that, as a result, the expressivist interpretation of mixed disjunctions is empirically incorrect.

Schroeder argues that, on certain natural assumptions, expressivism predicts that if 'ϕ' is a non-normative sentence and 'ψ' is a normative sentence, then it is impossible to believe $\phi \vee \psi$ without either believing ϕ or believing ψ:

> [I]f 'ϕ' is normative and 'ψ' is non-normative, there is no state of pure descriptive belief that someone who believes that $\phi \vee \psi$ is guaranteed to be in, and there is no [pure preference state] that she is guaranteed to be in. So *ipso facto* there is no combination of a descriptive belief state and [pure preference state] that she is guaranteed to be in"; "But since we're assuming that the only states of mind there are are ordinary descriptive beliefs, ordinary [preference states], and their combinations… on this view there is no room for such a state. (2011: 14, 12, variables adapted; *cf*. p. 31, 2008a: 124–127)

A bit more slowly: Giving 'or' its usual interpretation as set union, the content of the state of mind expressed by the mixed disjunction 'Alice doesn't have a job or Alice has to help the poor' is as follows.

(9) $\{\langle w, \lesssim \rangle$: Alice doesn't have a job in w, or Alice helps the poor in all the $\lesssim$-best worlds$\}$

Call this set 'S'. Bert believes that Alice doesn't have a job or she has to help the poor iff Bert's total belief state — his doxastic-practical alternatives (§4.2; *cf*. n. 25) — is a subset of S. But assuming that the only states of mind are ordinary factual beliefs and pure preference states and their combinations, then since S is the union of a world-independent content (a content that only constrains $\lesssim$) and a preference-independent content (a content that only constrains w), Bert can have this disjunctive belief only if either Bert's doxastic alternatives is a subset of S or Bert's practical alternatives is a subset of S. So Bert believes that Alice doesn't have a job or she

has to help the poor only if either Bert believes that Alice doesn't have a job or Bert believes that Alice has to help the poor. But this is incorrect. Bert can believe that Alice doesn't have a job or she should help the poor but be uncertain about which. One can believe a disjunction without believing either disjunct. So the expressivist conflates believing $\phi \vee \psi$ with either believing ϕ or believing ψ.

Like the case of the negation problem, this objection is partly technical and partly interpretive. It is technical because we need a way of blocking the entailment in question. It is interpretive because whatever story we tell about how to block this entailment, we need to be able to interpret the resulting apparatus so as to explain what states of mind are expressed by disjunctive sentences and how these states of mind are logically related to other states of mind.

As Schroeder notes, an intuitive way of thinking of the attitude expressed by a disjunctive sentence is in terms of Blackburn's (1988) notion of "being tied to a tree". A disjunction '$\phi \vee \psi$' can be understood as expressing a conditional commitment to take on the attitude expressed by 'ϕ' upon rejecting the attitude expressed by 'ψ', and vice versa. The now familiar problem for the expressivist, though, is that we have introduced another attitude — the attitude of being tied to a tree — and we must explain how it is logically related to the other attitudes in the expressivist's repertoire so as to capture the logical relations among disjunctive and non-disjunctive sentences. For example, we need to capture how '$\phi \vee \psi$', '$\neg \phi$', and '$\neg \psi$' are logically inconsistent. But the situation isn't hopeless. The attitude of being tied to a tree needn't be construed as an attitude that is "logically unrelated" to the attitudes of requiring, permitting, etc. Taking up our strategy in response to the negation problem, if we can derive these attitudes from a common core we can capture their logical relations sans stipulation.

Before turning to questions of interpretation, observe that Schroeder's objection fails if we drop the assumption that total belief states must be characterized into a purely descriptive component and a purely practical component, or that "the only states of mind there are are ordinary descriptive beliefs, ordinary [preference states], and their combinations."[27] The reason for Schroeder's assumption, of course,

---

[27]For a simple example, suppose our model contains only the following world-preference pairs: $\langle w_1, \lesssim_1 \rangle, \langle w_2, \lesssim_1 \rangle, \langle w_1, \lesssim_2 \rangle, \langle w_2, \lesssim_2 \rangle$. Suppose we have a non-normative sentence 'ϕ' and a normative sentence 'ψ' such that $[\![\phi]\!] = \{\langle w_1, \lesssim_1 \rangle, \langle w_1, \lesssim_2 \rangle\}$ and $[\![\psi]\!] = \{\langle w_1, \lesssim_1 \rangle, \langle w_2, \lesssim_1 \rangle\}$, so that $[\![\phi]\!] \cup [\![\psi]\!] = \{\langle w_1, \lesssim_1 \rangle, \langle w_1, \lesssim_2 \rangle, \langle w_2, \lesssim_1 \rangle\}$. ($[\![\alpha]\!]$ is the set of world-preference pairs that verify α.) Suppose that Alice believes $\phi \vee \psi$. Then $Bel_A \subseteq [\![\phi]\!] \cup [\![\psi]\!]$, where Bel_A is the set of $\langle w, \lesssim \rangle$ pairs that characterize Alice's total belief state. If, as Schroeder suggests, Bel_A must be a subset of the intersection of only world-independent and preference-independent sets of $\langle w, \lesssim \rangle$ pairs — that is, if Alice's belief state must be characterizable only in terms of purely descriptive beliefs and purely practical beliefs — then

is a worry that the expressivist lacks a systematic way of interpreting sets of world-preference pairs that are neither world-independent nor preference-independent. But our solution to the negation problem suggests a natural response. I suggest that we characterize these states in terms of ordinary belief and *conditional* (weak) preference. One might worry about introducing a new attitude, but the attitude of conditional preference, its logic, and its connection with belief and action are already well studied in decision theory and preference logic. Preferences — like norms, values, goals, etc. — typically don't come in the form of blunt categorical injunctions. They aren't usually of the form *No matter what, φ!*. Rather they often come with conditions under which they apply. For instance, if I want to go for a run, my preference needn't be that I go for a run, come what may. More plausibly it is that I go for a run given that it's sunny, that I'm not injured, that I didn't just eat a burrito, and so on. Our preferences are often conditional, preferences *for* certain circumstances. (Categorical preferences can be treated as preferences that hold conditional on any possibility.)

Thus far I have been treating modals as interpreted with respect to preorders. But since modals can themselves occur in intensional contexts, it is standard in ordering semantics to index preorders to a world of evaluation (written '$\lesssim_w$'). Which preorder is relevant for the interpretation of a given modal sentence can depend on how things happen to be in the actual world, or on how things could be but aren't or could have been but weren't. This suggests a natural way of representing conditional preferences. We can represent a body of conditional preference with a function from worlds to preorders. Call such a function a *preorder function*. Intuitively, a preorder function $\lambda w. \lesssim_w$ encodes one's preferences for the situation of w being actual; an indexed preorder $\lesssim_w$ represents one's preferences given the relevant circumstances in w. (Hereafter I use the unindexed '$\lesssim$' as short for '$\lambda w. \lesssim_w$' to denote preorder functions.) Suppose I prefer to go for a run given that it's sunny, and I prefer not to run given that it isn't sunny. Letting S and $\bar{S}$ be representatives of relevant equivalence classes of worlds where it's sunny and not sunny, respectively, we can represent these conditional preferences with a preorder function $\lesssim$ that is such that worlds where I run are, other things equal, $\lesssim_S$-better than worlds where I don't run, and is such that

---

it indeed follows that $Bel_A \subseteq \llbracket \phi \rrbracket \vee Bel_A \subseteq \llbracket \psi \rrbracket$, that is, that Alice believes ϕ or Alice believes ψ. (Given that $Bel_A \subseteq \llbracket \phi \rrbracket \cup \llbracket \psi \rrbracket$, the only total belief states that are the intersection of world- and preference-independent sets of world-preference pairs are $\llbracket \phi \rrbracket$, $\llbracket \psi \rrbracket$, $\{\langle w_1, \lesssim_1 \rangle\}$, $\{\langle w_1, \lesssim_2 \rangle\}$, and $\{\langle w_2, \lesssim_1 \rangle\}$, and each of these is either a (possibly improper) subset of $\llbracket \phi \rrbracket$ or a (possibly improper) subset of $\llbracket \psi \rrbracket$. So if $Bel_A \subseteq \llbracket \phi \rrbracket \vee \llbracket \psi \rrbracket$, then $Bel_A \subseteq \llbracket \phi \rrbracket$ or $Bel_A \subseteq \llbracket \psi \rrbracket$.) But dropping Schroeder's assumption blocks the entailment, for the following total belief states that are (possibly improper) subsets of $\llbracket \phi \rrbracket \cup \llbracket \psi \rrbracket$ neither entail ϕ nor entail ψ: $\llbracket \phi \rrbracket \cup \llbracket \psi \rrbracket$ itself, and $\{\langle w_1, \lesssim_2 \rangle, \langle w_2, \lesssim_1 \rangle\}$.

worlds where I don't run are, other things equal, $\lesssim_{\overline{S}}$-better than worlds where I do (n. 19).

In this way, preorder functions can be interpreted as preferences conditional on maximally specific possibilities. As in §4.1, this isn't to deny that agents have preferences conditional on non-maximal possibilities, and it is neutral on the relation between preferences conditional on maximal vs. non-maximal possibilities. Since not all circumstances in a world may be relevant to one's preferences, just as the preorders themselves may permit ties (they needn't be antisymmetric), so too preorder functions may map different worlds to the same preorder (they needn't be injective). Representing bodies of conditional preference with preorder functions has the advantage of utilizing existing apparatus from contemporary ordering semantics for modals. However, like in §4.1, alternative implementations will be possible depending on one's broader views about the sorts of issues mentioned there.

A world-preference pair now consists of a possible world w and a body of preferences conditional on w being actual. The singleton $\{\langle w, \lesssim_w \rangle\}$ represents a fully opinionated and decided state of mind; it represents a belief state that is settled on w being the actual world, and a body of conditional preferences $\lesssim$, which together determine the preferences $\lesssim_w$. Generalizing, an arbitrary set of world-preference pairs $\{\langle i, \lesssim_i \rangle, \langle j, \lesssim_j \rangle, \ldots, \langle u, \lesssim_u^* \rangle, \langle v, \lesssim_v^* \rangle, \ldots \}$ represents a belief state that is compatible with $i, j, u, v, \ldots$, and a preference state compatible with $\lesssim$ conditional on $i, j, \ldots$, $\lesssim^*$ conditional on $u, v, \ldots$, and so on; it represents a state of uncertainty and indecision. For example, the set $\{\langle w_1, \lesssim_{w_1}^1 \rangle, \langle w_1, \lesssim_{w_1}^2 \rangle, \langle w_2, \lesssim_{w_2}^1 \rangle\}$ represents a state of mind that is uncertain about whether w_1 or w_2 is actual, and is undecided about whether to have the preferences $\lesssim_{w_1}^1$, $\lesssim_{w_1}^2$, or $\lesssim_{w_2}^1$. There are two important things to notice about this: First, indecision in one's preferences can be the result of uncertainty, but it doesn't have to be. Second, one's preference state needn't be independent of one's beliefs.

Let's consider a concrete example to illustrate how appealing to conditional preferences, and representing them in terms of preorder functions, can help solve the disjunction problem. Let $J, \overline{J}, G$, and $\overline{G}$ be (representatives of equivalence classes of) relevant worlds in which Alice has or doesn't have a Job, and Gives or doesn't give to the poor. And let $\lesssim^1, \ldots, \lesssim^4$ be preorder functions with the following properties:

$$\lesssim^1: \quad G <^1_J \overline{G}, \quad G <^1_{\overline{J}} \overline{G}$$
$$\lesssim^2: \quad G <^2_J \overline{G}, \quad \overline{G} <^2_{\overline{J}} G$$
$$\lesssim^3: \quad \overline{G} <^3_J G, \quad G <^3_{\overline{J}} \overline{G}$$
$$\lesssim^4: \quad \overline{G} <^4_J G, \quad \overline{G} <^4_{\overline{J}} G$$

Intuitively, $\precsim^1$ represents a preference that Alice give to the poor regardless of whether she has a job; $\precsim^2$ represents a preference that Alice give to the poor iff she has a job; $\precsim^3$ represents a preference that Alice give to the poor iff she doesn't have a job; and $\precsim^4$ represents a preference that Alice not give to the poor regardless of whether she has a job. The contents of 'Alice doesn't have a job' ($\neg job$) and 'Alice has to give to the poor' ($M(give)$), then, are as follows:

$$\llbracket \neg job \rrbracket =$$
$$\{\langle w, \precsim_w \rangle : \text{Alice doesn't have a job in } w\} =$$
$$\{\langle \bar{J}, \precsim_{\bar{J}}^1 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^2 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^3 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^4 \rangle\}$$

$$\llbracket M(give) \rrbracket =$$
$$\{\langle w, \precsim_w \rangle : \text{Alice gives to the poor in all the } \precsim_w\text{-best worlds}\} =$$
$$\{\langle J, \precsim_J^1 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^1 \rangle, \langle J, \precsim_J^2 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^3 \rangle\}$$

The content of the mixed disjunction 'Alice doesn't have a job or she has to give to the poor' is the union of these sets:

$$\llbracket \neg job \vee M(give) \rrbracket =$$
$$\llbracket \neg job \rrbracket \cup \llbracket M(give) \rrbracket =$$
$$\{\langle J, \precsim_J^1 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^1 \rangle, \langle J, \precsim_J^2 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^2 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^3 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^4 \rangle\}$$

For concreteness, consider the following subset of this set, $T = \{\langle J, \precsim_J^2 \rangle, \langle \bar{J}, \precsim_{\bar{J}}^2 \rangle\}$. The crucial question is this: What state of mind is represented by T? Intuitively, it is the state of mind of being uncertain about whether Alice has a job, and of preferring that she give to the poor conditional on her having a job and preferring that she not give to the poor conditional on her not having a job. It is a state of uncertainty and indecision (in this case, where the indecision derives from the uncertainty). Since T is a subset of the content of the mixed disjunction but not of the content of either disjunct, if T represents Bert's total state of mind, then Bert counts as believing that Alice doesn't have a job or has to give to the poor, without believing that Alice doesn't have a job or believing that she has to give to the poor. The inference from believing $\phi \vee \psi$ to believing ϕ or believing ψ is blocked.

To be clear, the proposal is *not* that utterances of mixed disjunctions conventionally express simple conditional preferences — e.g., preferences that hold conditional on the negation of a factual antecedent. Saying this would fail to provide a general interpretation of disjunction. For instance, it won't apply to disjunctions in which each disjunct is normative. It also naturally raises a worry that disjunctions in which one disjunct is itself a mixed disjunction will require us to iterate the "go con-

ditional" strategy past the point where we have an independent grasp on the posited attitude (what would conditional conditional preference be?). An agent might count as believing that Alice doesn't have a job or she has to give to the poor in virtue of having a standing conditional preference for Alice give to the poor given that she a job. But having this conditional preference isn't necessary for believing the mixed disjunction. It isn't the state of mind *conventionally* expressed by an utterance of $\neg job \vee M(give)$. The attitude conventionally expressed by an utterance of a sentence 'ϕ' is the *least* committal way of thinking that ϕ.

Rather, the strategy is to give a general recipe for interpreting arbitrary sets of world-preference pairs in terms of ordinary beliefs about ways the world might be and preferences that hold conditional on the world being certain of those ways (see above). This recipe is to apply to any disjunction, whether mixed or non-mixed, simple or complex. Indeed, it is to apply to any sentence (or at least any sentence whose content is a set of world-preference pairs).[28] The strategy is this:

How to solve the disjunction problem:
1. Give disjunction the formal semantics we would normally give it. (E.g., treat the content of '$\phi \vee \psi$' as the union of the content of 'ϕ' and the content of 'ψ'.)
2. Drop the assumption that total belief states must be factorable into purely descriptive and purely practical components.
3. Give a general way of interpreting (i.e., characterizing what states of mind correspond to) arbitrary sets of world-preference pairs.
4. Apply this general strategy to the case of the contents of disjunctive sentences.
5. Note that Schroeder's problematic entailment is blocked.

By giving a general way of characterizing what states of mind are represented by arbitrary sets of world-preference pairs, the ordering expressivist can drop Schroeder's assumption in Step 2 above, and can do so in such a way that explains the logical relations among mixed states of mind and their pure counterparts, and, thus, among

---

[28]Treating disjunctions with normative disjuncts as conventionally expressing attitudes of indecision (perhaps among other things) is compatible with granting that one can accept such disjunctions without being undecided. This is no different from the ordinary factual case. Disjunctions with ordinary factual sentences conventionally express attitudes of uncertainty, though one can accept them while being certain about which disjunct is true.

There may ultimately be reasons for further complicating the proposed representations of semantic contents and states of mind in light of independent issues concerning epistemic vocabulary ('might', 'probably', etc.). See note 1; for specific discussion of disjunction in epistemic expressivist theories, see Rothschild 2012, Swanson 2012.

disjunctive and non-disjunctive sentences. This allows us to block problematic entailments in cases with mixed disjunctions, and distinguish believing $\phi \vee \psi$ from believing ϕ or believing ψ. More generally, in providing such a recipe for interpreting arbitrary sets of world-preference pairs, we can give a systematic expressivist interpretation of normative, descriptive, and mixed language. It is in this way that ordering expressivism makes progress toward a "fast track" solution to the Frege-Geach problem (§3).

6 Conclusion

Despite its checkered past, expressivism has found robust support in a variety of domains. In the case of normative judgment, expressivism promises, among other things, a substantive account of its distinctive practical character, a feature that alternative theories that assimilate normative judgment to ordinary factual judgment often struggle to capture. But it is often thought expressivism has intractable problems with capturing the meanings and logical properties of complex sentences, and so cannot make good on this promise.

I have argued that we can make progress in eliminating expressivism's semantic stumbling block by developing what I call *ordering expressivism*. By interdefining the attitudes intuitively expressed by normative sentences in terms of a single basic practical attitude, the expressivist can capture the logical relations among these attitudes and thus among the sentences that express them. Though I have focused on the particular cases of negation and disjunction, the structure of our solution suggests a more general strategy for implementing an expressivist semantics. Ordering expressivism co-opts developments from linguistic semantics for modals by systematically interpreting the orderings in these analyses in terms of a suitable practical attitude. Though expressivism does make novel claims about semantic explanation and the nature of normative language and judgment, it needn't reinvent the compositional semantic wheel. This, I hope, will prompt renewed interest in other distinctive aspects of expressivism, as in philosophy of mind and psychology.

References

Arrow, Kenneth J. 1959. Rational choice functions and orderings. *Economica 26*, 121–127.

AUMANN, ROBERT J. 1962. Utility theory without the completeness axiom. *Econometrica 30*, 445–462.

BLACKBURN, SIMON. 1984. *Spreading the word*. Oxford: Clarendon Press.

BLACKBURN, SIMON. 1988. Attitudes and contents. In BLACKBURN (1993), pp. 182–197.

BLACKBURN, SIMON. 1993. *Essays in quasi-realism*. New York: Oxford University Press.

BLACKBURN, SIMON. 1998. *Ruling passions*. Oxford: Clarendon Press.

BROOME, JOHN. 1991. Desire, belief, and expectation. *Mind 100*, 265–267.

BROOME, JOHN. 1993. Can a Humean be moderate? In R.G. FREY & CHRISTOPHER MORRIS (Eds.), *Value, welfare, and morality*, pp. 51–73. Cambridge: Cambridge University Press.

BROOME, JOHN. 2004. *Weighing lives*. Oxford: Oxford University Press.

BYRNE, ALEX & ALAN HÁJEK. 1997. David Hume, David Lewis, and decision theory. *Mind 105*, 411–428.

CARIANI, FABRIZIO, STEFAN KAUFMANN, & MAGDALENA SCHWAGER. 2011. Deliberative modality under epistemic uncertainty. MS.

CHANG, RUTH. 2002. The possibility of parity. *Ethics 112*, 659–688.

CHARLOW, NATE. 2011a. What we know and what to do. *Synthese* pp. 1–33.

CHARLOW, NATE. 2013. The problem with the Frege-Geach Problem. *Philosophical Studies* .

CHARLOW, NATHAN A. 2011b. *Practical language: Its meaning and use*. Ph.D. thesis, University of Michigan.

DAVIDSON, DONALD. 1976. Hempel on explaining action. In *Essays on actions and events*, pp. 261–276. 2001. New York: Clarendon Press, 2nd edn.

DAVIDSON, DONALD, JOHN McKINSEY, & PATRICK SUPPES. 1955. Outlines of a formal theory of value. *Philosophy of Science 22*, 140–160.

DE SOUSA, RONALD. 1974. The good and the true. *Mind 83*, 534–551.

DOWELL, J.L. 2013. Flexible contextualism about deontic modals: A puzzle about information-sensitivity. *Inquiry 56*, 149–178.

DREIER, JAMES. 2006. Negation for expressivists: A collection of problems with a suggestion for their solution. In RUSS SHAFER-LANDAU (Ed.), *Oxford studies in metaethics*, vol. 1, pp. 217–233. Oxford University Press.

DREIER, JAMES. 2009. Relativism (and expressivism) and the problem of disagreement. *Philosophical Perspectives 23*, 79–110.

ELIAZ, KFIR & EFE A. OK. 2006. Indifference or indecisiveness? Choice-theoretic foundations of incomplete preferences. *Games and Economic Behavior 56*, 61–86.

FIELD, HARTRY. 2000. A priority as an evaluative notion. In P. BOGHOSSIAN & C. PEACOCKE (Eds.), *New essays on the a priori*, pp. 117–149. Oxford: Oxford University Press.

VAN FRAASSEN, BAS C. 1973. Values and the heart's command. *The Journal of Philosophy 70*, 5–19.

GEACH, PETER. 1960. Ascriptivism. *Philosophical Review 69*, 221–225.

GEACH, PETER. 1965. Assertion. *Philosophical Review 74*, 449–465.

GIBBARD, ALLAN. 1990. *Wise choices, apt feelings: A theory of normative judgment*. Cambridge: Harvard University Press.

GIBBARD, ALLAN. 2001. Explaining with mental content: Prospects for content expressivism. Content Seminar, New York University.

GIBBARD, ALLAN. 2003. *Thinking how to live*. Cambridge: Harvard University Press.

GIBBARD, ALLAN. 2012. *Meaning and normativity*. New York: Oxford University Press.

HALPERN, JOSEPH Y. 1997. Defining relative likelihood in partially-ordered preferential structures. *Journal of Artificial Intelligence Research 7*, 1–24.

HANSSON, BENGT. 1968. Choice structures and preference relations. *Synthese 18*, 443–458.

Hansson, Sven Ove. 2001. Preference logic. In Dov M. Gabbay & Franz Guenthner (Eds.), *Handbook of Philosophical Logic, Volume 4*, pp. 319–393. Dordrecht: Kluwer, 2nd edn.

Hansson, Sven Ove & Till Grüne-Yanoff. 2012. Preferences. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Winter 2012 edn.

Hare, Caspar. 2010. Take the sugar. *Analysis 70*, 237–247.

Hare, R.M. 1952. *The language of morals*. Oxford: Oxford University Press.

Hare, R.M. 1981. *Moral thinking: Its levels, method, and point*. Oxford: Clarendon Press.

Hawthorne, John. 2002. Practical realism? *Philosophy and Phenomenological Research 64*, 169–178.

Hintikka, Jaako. 1969. Semantics for propositional attitudes. In J.W. Davis, D.J. Hockney, & W.K. Wilson (Eds.), *Philosophical logic*, pp. 21–45. Dordrecht: Reidel.

Horgan, Terry & Mark Timmons. 2006. Cognitivist expressivism. In T. Horgan & M. Timmons (Eds.), *Metaethics after Moore*, pp. 255–298. Oxford: Oxford University Press.

Horgan, Terry & Mark Timmons. 2009. Expressivism and contrary-forming negation. *Philosophical Issues 19*, 92–112.

Hsieh, Nien-hê. 2008. Incommensurable values. In Edward N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Fall 2008 edn.

Katz, Graham, Paul Portner, & Aynat Rubinstein. 2012. Ordering combination for modal comparison. In Anca Chereches (Ed.), *Proceedings of SALT 22*, pp. 488–507. Ithaca: CLC Publications.

Kolodny, Niko & John MacFarlane. 2010. Ifs and oughts. *Journal of Philosophy* pp. 115–143.

Kratzer, Angelika. 1977. What 'must' and 'can' must and can mean. *Linguistics and Philosophy 1*, 337–355.

Kratzer, Angelika. 1981. The notional category of modality. In Hans-Jürgen Eikmeyer & Hannes Rieser (Eds.), *Words, worlds, and contexts: New approaches in word semantics*, pp. 38–74. Berlin: de Gruyter.

Kratzer, Angelika. 1991. Modality/Conditionals. In Arnim von Stechow & Dieter Wunderlich (Eds.), *Semantics: An international handbook of contemporary research*, pp. 639–656. New York: de Gruyter.

Kratzer, Angelika. 2012. *Modals and conditionals: New and revised perspectives*. New York: Oxford University Press.

Lassiter, Daniel. 2011. *Measurement and modality: The scalar basis of modal semantics*. Ph.D. thesis, New York University.

Lewis, David. 1973. *Counterfactuals*. Cambridge: Harvard University Press.

Lewis, David. 1981a. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic 10*, 217–234.

Lewis, David. 1981b. What puzzling Pierre does not believe. *Australasian Journal of Philosophy 59*, 283–289.

Lewis, David. 1988. Desire as belief. *Mind 97*, 323–332.

Moss, Sarah. 2012. Epistemology formalized. *Philosophical Review 122*, 1–43.

Nehring, Klaus. 1997. Rational choice and revealed preference without binariness. *Social Choice Welfare 14*, 403–425.

Portner, Paul. 2009. *Modality*. Oxford: Oxford University Press.

Price, Huw. 1983. Does 'probably' modify sense? *Australasian Journal of Philosophy 61*, 396–408.

Ramsey, Frank P. 1928. Truth and probability. In R.B. Braithwaite (Ed.), *The foundations of mathematics and other logical essays*, pp. 156–198. 1931. London: Routledge and Kegan Paul.

Richter, Marcel K. 1966. Revealed preference theory. *Econometrica 34*, 625–645.

Rosen, Gideon. 1998. Blackburn's *Essays in Quasi-Realism*. *Noûs 32*, 386–405.

ROTHSCHILD, DANIEL. 2012. Expressing credences. *Proceedings of the Aristotelian Society 112*, 99–114.

SAMUELSON, P.A. 1938. A note on the pure theory of consumers' behavior. *Economica 5*, 61–71.

SCHROEDER, MARK. 2008a. *Being for: Evaluating the semantic program of expressivism*. Oxford: Oxford University Press.

SCHROEDER, MARK. 2008b. How expressivists can and should solve their problems with negation. *Noûs 42*, 573–599.

SCHROEDER, MARK. 2008c. What is the Frege-Geach problem? *Philosophy Compass 3*, 703–720.

SCHROEDER, MARK. 2011. Attitudes and epistemics. MS, University of Southern California.

SEARLE, JOHN R. 1962. Meaning and speech acts. *Philosophical Review 71*, 423–432.

SEN, AMARTYA. 1970. *Collective choice and social welfare*. San Francisco: Holden Day.

SEN, AMARTYA. 1997. Maximization and the act of choice. *Econometrica 65*, 745–779.

SILK, ALEX. 2013a. Evidence sensitivity in weak necessity deontic modals. *Journal of Philosophical Logic* pp. 1–33.

SILK, ALEX. 2013b. Truth-conditions and the meanings of ethical terms. In RUSS SHAFER-LANDAU (Ed.), *Oxford studies in metaethics*, vol. 8, pp. 195–222. New York: Oxford University Press.

SMITH, MICHAEL. 1987. The Humean theory of motivation. *Mind 96*, 36–61.

STALNAKER, ROBERT. 1984. *Inquiry*. Cambridge: MIT Press.

STALNAKER, ROBERT. 1999. *Context and content: Essays on intentionality in speech and thought*. Oxford: Oxford University Press.

SWANSON, ERIC. 2012. The application of constraint semantics to the language of subjective uncertainty. *Journal of Philosophical Logic* pp. 1–28.

Unwin, Nicholas. 1999. *Quasi*-realism, negation, and the Frege-Geach problem. *Philosophical Quarterly 49*, 337–352.

Unwin, Nicholas. 2001. Norms and negation: A problem for Gibbard's logic. *Philosophical Quarterly 51*, 60–75.

Veltman, Frank. 1976. Prejudices, presuppositions, and the theory of conditionals. In Jeroen Groenendijk & Martin Stokhof (Eds.), *Amsterdam papers in formal grammar*, vol. 1, pp. 248–281. Central Interfaculteit, University of Amsterdam.

Yalcin, Seth. 2007. Epistemic modals. *Mind 116*, 983–1026.

Yalcin, Seth. 2012. Bayesian expressivism. *Proceedings of the Aristotelian Society 112*, 123–160.