

Negation for Expressivists

Jamie Dreier
Brown University

I. The Negation Problem for Expressivists

- (M) Miss M. thinks one must write thank you notes by hand.
(N₁) Miss M. believes one must not write thank you notes by hand.
(N₂) Miss M. believes \neg (one must write thank you notes by hand).
(N₃) Miss M. does not believe one must hand write thank you notes.

Suppose there is some attitude, A, that Miss M. can have toward writing thank you notes by hand, and (M) says she has it. Then N₁ says she has A to not writing them by hand. N₃ says that she lacks A. What does N₂ say?

II. The Family of Problems

Duals

$$\Box A \cong \neg \Diamond \neg A \quad \neg \Box A \cong \Diamond \neg A \quad \neg \Diamond A \cong \Box \neg A$$

$$\mathcal{R}A \cong \neg \mathcal{P} \neg A \quad \neg \mathcal{R}A \cong \mathcal{P} \neg A \quad \neg \mathcal{P}A \cong \mathcal{R} \neg A$$

Why not introduce two brute attitudes? Because then it would be mysterious how they can be related to one another logically.

Completeness

What makes some collection of normative attitudes complete? Gibbard imagines Hera, whose attitudes are complete: for each state of affairs, Hera permits it or its complement (possibly both).

Compare the ultra-tolerant goddess, Laissezfairea, who accepts the maximally permissive normative system (forbidding nothing), with the perfectly unreflective and indecisive Gnoidea, who accepts no normative system. Laissezfairea's outlook is complete, Gnoidea's is radically incomplete; but how can we distinguish them? It would be nice to add, as an additional fact, Gnoidea's state is incomplete. We could define that fact if we could define external negation for normative judgments: an outlook is complete iff for each P the outlook includes either P or $\neg P$.

Consistency

No mortal is complete; no more are we complete in our beliefs. To be complete in beliefs would be to have a fully detailed view about the world. The

views of doxastic gods could be represented by single possible worlds; ours, though, are (better) represented by sets of worlds. Similarly, our normative states of mind are represented by sets of completed normative systems.

Which set of complete normative systems represents my normative state of mind? The set of all complete normative systems that are consistent with my state of mind. But consistency is tricky.

Think of a complete normative system, say S, that requires handwritten notes, and suppose N₃; now, is Miss Manners' state consistent with S? If Miss M. is undecided, then S is consistent with her state; but if N₂ is true, if Miss M. positively permits writing notes in other ways, then S is not consistent with her state. If only we could say what N₂ adds to N₃, we'd have a way to associate a unique set of completed systems with each person's normative state (and conversely). So consistency joins our family.

III. Thinking How to Live

ThL uses planning states in place of norm acceptance. The most fundamental normative judgments (about what to do) express plans.

To solve the negation problem (among others), Gibbard leans on the intuitive idea that it makes sense to disagree with some kinds of psychological states but not to think of disagreement with other kinds; for example, it makes sense to say you disagree with my baseball loyalties but not to say that you disagree with the throbbing pain in my ankle.

This is indeed intuitive, but it is not clear that the difference between the kinds of states can be made out within an expressivist framework. The worry is that the best account of that difference is that some psychological states have a propositional content and others do not, or that some are representational and others are not.

IV. Planning to the Rescue?

Hyperplans promise a solution because of a special property of planning: the *No Mere Permissions* property.

First Approximation: forced completeness and automatic negation

If I do not plan anything that excludes ϕ , then my plans are consistent with the requirement to ϕ . Permitting ϕ -ing is simply not planning not to ϕ . And if someone doesn't plan to ϕ (in C), and doesn't plan not to ϕ , then her plan is incomplete.

But we have solved a problem only by eliminating the problematic phenomenon, namely, the existence of Mere Permissions. Mere Permissions are things that we are permitted but not required to do. Completeness of plans is purchased at the price of elimination of mere permissions.

Second Approximation: Mere Permissions via Indifference

Think of planning as being like preferring.

One hitch: preferring is inherently comparative, whereas planning talk is not. Is this a real issue? Generally, planning to φ is planning to φ rather than to ψ , where the contrast is given by the conversational context, and in degenerate cases planning to φ is simply planning to φ rather than to not φ . (Cf. preference talk.)

Planning is identifiably complete because you have a plan in any given circumstance just in case you prefer one of the options in that circumstance to the others. This inherent completeness in planning is also responsible for the eradication of mere permissions. (Anything that's permitted is planned, in a complete plan.)

But permissions can be reinstated by adding indifference to our model. You regard both φ and ψ as permissible when you are indifferent between them (and do not prefer anything to them).

Indifference is not at all exotic or unusual: the example of dishwashing liquids, and of any case in which refinement of alternatives is intuitively over-refinement.

This leaves two problems: one is tolerable, the other serious.

V. Two Remaining Problems

The impossibility of supererogation

When you are indifferent between alternatives, you don't regard either as better. But when you regard an act as supererogatory, there is some alternative that you take to be permissible, even though the supererogatory act is better than the merely permissible. *The Example of the Lost Wallet*.

This looks like a problem. Supererogation is very intuitive. Is it ruled out entirely by the Planning model of expressivist semantics?

Yes and no. It is ruled out by the expressivist logic of Being the Thing to Do,

the most basic normative notion. But is this a real problem? Moral supererogation is deeply embedded in moral thinking, but maybe nothing like supererogation really does make sense in the realm of the most fundamental 'ought'. When we think about what to make of thinking about how to live, we might recover a place for supererogation in the special realm of the moral. If there is no place for it on the ground floor, that might be tolerable.

Indifference and preference: another family member?

(N₃) is about what plan or preference Miss Manners lacks: she does not plan or prefer to write thank you notes by hand. It is consistent with Miss Manners' having no view at all about thank you notes, and is not, of course, the same as her being indifferent among the ways of writing them.

But why is it not the same? How can the intuitive difference between dishwashing liquid no-preference (indifference) and coffee no-preference (indecision) be made out in an expressivist framework?

Maybe the best answer is that the absence of preference is indifference just in case the agent believes that the alternatives are equally good, and the absence of preference amounts to indecision just in case the agent has no belief about which alternative is better. I suspect the explanation goes the other way around. I suggest explaining the difference between the kinds of no-preference by way of defined indifference. Say that someone is *defined indifferent* between A and B iff (i) she does not prefer A to B or B to A, and (ii) for any X, she prefers A to X iff she prefers B to X and prefers X to A iff she prefers X to B. Defined indifference has the right logical properties to play the role of indifference in the expressivist account of negation. Almost.

Although defined indifference is (equivalent to) indifference for real people, it comes apart in troublesome ways for possible agents with extremely impoverished fields of preference. *The Thoroughly Indecisive Example; the Example of Momentous and Trivial Choices*.

A few last thoughts...

Fertilizer: artificially enriching the field

Reflections on other partial orderings

Imputed intentional states: when does a disposition to prefer (or believe) become a dispositional preference (or belief)?