Inferentialism as an Alternative to Expressivism

Matthew Chrisman
University of Edinburgh

1. Introduction

Normative discourse is about what people ought to do, think, and feel. It includes statements which appear to be truth-apt expressions of normative beliefs, some of which count as knowledge. But normative *oughts* do not seem to fit cleanly amongst what is knowable through observation and broadly scientific methods. This leads many philosophers to adopt the expressivist position that normative statements get their meaning from how they express attitudes akin desires in their “direction of fit.” Maybe we call these “beliefs,” but only in a minimalist sense. And maybe some kind of deflationary account of truth could help us earn the right to treat normative statements as truth-apt and to recognize some of the “beliefs” they express as knowledge. In the final analysis, however, the expressive role of normative statements is importantly different from the expressive role of descriptive statements. The mental states they express “guild and stain” or “guide and goad” rather than represent normative reality.

The most influential challenge for this position comes from logically complex statements that embed normative elements. The statements “Either Alex is there or Sarah ought to go,” “If James ought to pay, then he will pay,” and “It might be the case that Bob ought to wait,” are clearly meaningful. But it’s difficult to explain their meanings within the expressivist framework while holding onto uniform semantic treatments of logical connectives and operators.¹ This challenge leads some metaethicists go in for hybrid or global or dynamic expressivist semantics which pursue complicated but hopefully empirically adequate and compositionally tractable treatments of logical operators and connectives.² But, seeing these as desperate moves in the philosophy of language motivated mainly by a desire to save expressivism, many more metaethicists seem resigned to the view that normative language must be representational after all.

¹ To see this, ask yourself: What exactly is the mental state expressed by each of these complex statements? And how exactly does the content of these mental states relate to the meanings of those statements? Critics of expressivism have easy answers to these questions, expressivists do not. See Schroeder (2008, 2010).

Perhaps one of these responses is correct. In this paper, however, I want to put another option on the table. My basic idea is that a general and familiar idea in the philosophy of language can help to develop a nonrepresentationalist but also clearly nonexpressivist version of the view that normative discourse performs a fundamentally different communicative function from descriptive discourse. In a nutshell, the idea is to center the commissive role of language rather than any descriptive or expressive roles. I’ll argue this emphasis opens up ample space to defend the view that normative language is nonrepresentational without any appeal to the idea that normative statements mean what they do because of how they express desire-like attitudes. Rather than guild and stain or guide and goad, I want to suggest that normative vocabulary is most fundamentally for structuring and governing a rational practice of reasoning with our interpersonal discursive commitments about what to do, think, and feel.3

Philosophers of language and metaethicists familiar with inferentialist approaches to meaning will find the brief story told here familiar. But I want to tell in a way that is purpose-built to help metaethicists to glimpse the terrain of possible accounts of the meaning of normative language that are neither representationalist nor expressivist. My main goal is to encourage readers to shed some baggage from twentieth century metaethics in order to explore it more.4 To that end, most of this paper is devoted to explaining the basic conceptual gear we’ll need and suggesting a few paths into this terrain. But I’ll conclude by explaining how I think following these paths promises to open up fruitful new perspectives on three further issues in metaethics that aren’t specifically about normative language.

2. Explaining Meaning

Semanticsists often begin their accounts of meaning with the idea that the basic function of language is to pick out, refer to, or represent bits of reality. The thought is very intuitive: the noun ‘the apple’ picks out a fruit, the name ‘Jane’ refers to a person, the verb ‘is eating’ represents an action, and when we put these together, we can make a statement “Jane is eating the apple” that represents a way reality could be. This is so obvious that one might suppose all other semantic phenomena are just elaborations of this basic referential-cum-representational function of language.

I mention this Fregean approach first because I want to suggest that two other approaches that I’m going to call Lockeian and Peircean are not in tension with the core Fregean idea but rather operate at a more fundamental level of explanation. That is to say that they can accept that a basic function of language is to pick out, refer to, or represent bits of reality, but they will go on to ask two further questions:

- What it is in virtue of which statements such as “Jane is eating an apple” represent reality?
- Does our answer to the first question commit us to thinking that all statements represent reality?

3 Some versions of constructivism might also be seen alternatives to expressivism. See Street (2010) and Bagnoli (2021) for discussion. The view I sketch here is not intended as a version of constructivism, but it might be consistent with some forms of constructivism.

4 I tell a different story with a similar conclusion in Chrisman (2010, 2018) exploring other areas of the metaethical terrain.
Painting quickly with a broad brush, I classify views as Lockean that answer the first question by focusing on the phenomenon of a speaker intentionally expressing ideas, attitudes, or mental states to an audience by uttering words whose meaning depends in part on their conventionalized usability to get an audience to understand what the speaker is thinking or feeling. On this approach, the fundamental function of meaningful speech is expressing what is antecedently in one’s mind as part of communicating one’s attitudes to others. Thus, in order to explain why “Jane is eating an apple” represents the fact that Jane is eating an apple, Lockeans will say something that boils down to the idea that the *expressive* function of this statement is convey the belief that reality contains the fact that Jane is eating an apple. The aspiration to explain meaningfulness this way motivates various developments within both Gricean\(^5\) and Davidsonian\(^6\) programs of accounting for linguistic meaning by appeal to recursively processed compositionally structured linguistic tools for conveying ideas, attitudes, or mental states under the assumption that our interlocutors are more-or-less rational.

In contrast, I classify views as Peircean which answer the first question by focusing on the phenomenon of a speaker vouching for some addition to what is mutually accepted in an ongoing conversation. On this approach, the fundamental function of meaningful speech is to propose something for collective acceptance and thereby undertake a distinctive kind of commitment as part of discussing things with others. Thus, in order to explain why “Jane is eating an apple” represents the fact that Jane is eating an apple, Peirceans will say something that boils down to the idea that the *commissive* function of this statement is to undertake a commitment to reality’s containing the fact that Jane is eating an apple. The aspiration to explain meaningfulness in this way motivates various developments within both the Sellarsian\(^7\) and Lewisian\(^8\) programs of accounting for linguistic meaning by appeal to conventionalized and functionally characterizable tools for updating the ongoing conversational score of discursive commitments amongst interlocutors.

In sum, Lockeans see *potential to express our minds to others* as the key to understanding linguistic meaning whereas Peirceans see *potential to undertake mutually recognized discursive commitments* as the key to understanding linguistic meaning. I’m not here to argue that one approach is better than the other.\(^9\) They both have lots to teach us about meaning.

The important point for what follows is that they are both consistent with a *negative* answer to the second question above. As far as the story has been told, Lockeans and Peirceans can both

---

\(^5\) Grice (1989), Schiffer (1972) Bennett (1976), and Davis (2003).

\(^6\) Davidson (1984), Harman (1972), and Lycan (1984). Griceans and Davidsonians sometimes disagree about whether semantic content and mental content are explanatorily equally fundamental or whether the former can be explained in terms of the latter, but they agree that it is in the nexus between these that the explanation of meaning is to be found.


\(^8\) Lewis (1969), Stalnaker (1978). These philosophers characterize conversational updates in terms of coordinating individual beliefs, which provides a bridge to more Lockean sorts of views. However, as long as update potentials are normatively characterized and thought to explain mental content, rather than vice versa, I would characterize these positions as Peircean.

\(^9\) Certain causal theories of reference might be thought to avoid recourse to the Lockean idea of expressing mental states and the Peircean idea of undertaking discursive commitments. However, these won’t have anything to say about what constitutes the meaning of non-referring terms such as ‘ought.’
maintain that not all statements represent reality. We clearly use language for far more than just describing things thought to be antecedently “out there.” There are statements such as ‘That’s disgusting’ and ‘The Backstreet Boys are like wow’ whose function is clearly expressive; and performatives such as ‘I promise’ and ‘We shall overcome’ have a commissive function that is not easily assimilated to description. So, granting that some language represents reality, we still need to make room for these nondescriptive functions. Herein lies the prospects for expressivist and inferentialist treatments of normative language.

3. Expressivism and Inferentialism

It should be relatively clear where expressivism fits into the picture I have drawn so far. Expressivists are Lockeans about meaning who are convinced that the expressive role of normative language is not to convey a belief about how reality is but rather to express an attitude that has a desire-like direction of fit. Hence, the interesting question for what follows is what a Peircean alternative to this might look like, and whether it carries any distinctive advantages or disadvantages compared to expressivism.

I think the best place to go for a Peircean alternative is the inferentialist accounts of conceptual content developed by philosophers such as Sellars (1957, 1974), Rosenberg (1974, 1976), Brandom (1994, 2008), and Williams (2004, 2010). However, the historical and dialectical contexts of their work is rarely metaethical. So, my goal in the next three sections is first to constructively excavate a basic version of metaethical inferentialism from these philosophers and then to offer some elaborations and defenses for the purposes of metaethics. This will put us in a position at the end to appreciate new perspectives on three other issues in metaethics.

As I understand the label, these Peircean philosophers of language are “inferentialists” because they understand the meaning-determining commissive function of statements in terms of inferential responsibilities and entitlements. The basic thought is that, when a person makes a statement in ordinary conversation, they incur a defeasible responsibility for backing it up with reasons (the sorts of further statements from which one could infer the truth of the target statement) should it later come into question. Additionally, they are tacitly understood to have acknowledged a defeasible collective entitlement to rely on the truth of the target statement (as a consideration from which one could make further inferences) in reasoning about what to do, think, and feel.10

By itself, that doesn’t entail any kind of nonrepresentationalism about normative discourse. We get that from a further claim that normative vocabulary is not primarily for undertaking commitments to reality being a particular way but rather for undertaking commitments about how to act and react in light of various ways that reality is or might be. The rough idea is that the inferential responsibilities and entitlements that normative vocabulary functions to register on the conversational scoreboard are different from the inferential responsibilities and entitlements that descriptive vocabulary functions to register on the conversational scoreboard. Can we make that more precise?

---

10 The label “inferentialism” is potentially misleading, as it shouldn’t be thought to commit these philosophers to the idea that all rationally defensible attitudes and actions are always the result of an inference, conceived as a reason-guided movement of the mind from premises to conclusion. The relevant sort of “inferential support” may be available only post hoc or to third-personal observers of the action/attitude, and it would still count as “inferential” on this broad way of thinking of reasoned support for actions and attitudes.
Yes. The first step is to note that inferentialists do not conceive of meaning-determining discursive commitments as a special kind of mental state. A speaker might count as insincere or hypocritical if they lack certain mental states when undertaking a discursive commitment, and the whole practice of recognizing inferential responsibilities and entitlements presupposes that participants have various psychological capacities and tendencies. Nevertheless, the inferentialist’s notion of commitment belongs to the public domain of social standings rather than the psychological domain of particular people’s minds.\textsuperscript{11}

Because of this, inferentialists will not offer a psychological explanation of the difference between normative and nonnormative vocabulary. Indeed, they typically assume that all statements express beliefs in a univocal sense of “belief” which allows them to recognize that logically complex statements express beliefs in logically complex propositions, whether or not normative vocabulary is involved.\textsuperscript{12} (This is the first step to avoiding some of the specific the specific trouble expressivists run into articulating the expressive role of complex sentences that embed normative elements.) But, if it’s not in psychological terms, then how would an inferentialist explain why normative discourse is not representational of reality?

To begin to answer this question, focus first on the commissive function of different kinds of vocabulary.\textsuperscript{13} Even within the uniform terrain of statements whose meaning is explicable in terms of inferential commitments, there are interesting ways to distinguish the kinds of inferential commitments different vocabularies allow us to undertake. Inferentialist arguments for this kind of distinction sometime take the form of a speculative genealogy. One makes suggestions about how a particular kind of vocabulary entered the conceptual-linguistic repertoire of humans in order to motivate a claim about what it is for. Other times (and I think more plausibly) inferentialist arguments takes the form of a holistic reflection on all the things we do with language, how they relate to each other, and the things we couldn’t do (or do nearly as easily) if we lacked specific vocabularies. Either way we argue for it, the approach is easiest to appreciate by first considering some nonnormative examples.

We might identify empirical vocabulary as those terms/concepts that are primarily for tracking objects in our environment (e.g., tables, chairs, electrons, quarks) and properties of those objects (e.g., being flat, being red, being small). The basic idea in calling this vocabulary “empirical” is that these things are presumed to be knowable through observation and scientific explanation. When someone makes a statement classifying things with empirical vocabulary, this commits them to reality being a particular way. The inferentialist will cash out this commitment in terms of responsibilities for giving reasons involving our ability to observe,

\textsuperscript{11} In contrast to Blackburn (1988) and Horgan and Timmons (2006) who treat commitments these as kinds of mental states.

\textsuperscript{12} Brown (2022) argues that anti-representationalists should embrace a unified and ontologically non-committal account of belief and proposition.

\textsuperscript{13} This isn’t mandatory for being an inferentialist. For example, inspired by Hare (1952), an inferentialist might argue that normative statements are proposals to add universalized prescriptions (“Everyone, see to it that p!”) to the common ground of a conversation, where these are understood as commitments which could be justified by other things someone might say and which could be appealed to in order to justify other prescriptions but also intentions and actions. As long as prescriptions are not conceived as private psychological states but understood in terms of mutually acknowledged upstream responsibilities and downstream entitlements to and from which one can infer, I think this would count as forms of metaethical inferentialism rather than metaethical expressivism. However, I don’t think Harean inferentialism is the most plausible version.
track, or think about how reality is and entitlements for relying on the statement as a premise in further collective reasoning that depends on how reality is.

An inferentialist might then distinguish empirical vocabulary from logical vocabulary (e.g., if-then, not, and, or) understood as those terms/concepts which are domain-neutral scaffolding for articulating and undertaking commitment to logico-semantic connections between other more basic sorts of statements. The idea in classifying logical vocabulary as separate from empirical vocabulary is that we don’t typically think of logical terms/concepts as earning their keep by picking out and describing objects in reality. Of course, they can be used to formulate complex statements carrying descriptive content about objects in our environment (e.g. “Grass is green and snow is white”) but they can just as well be used to state semantic facts (“If someone is a bachelor, then they are unmarried”), explicate reasoning about subjective matters of taste (“These olives are disgusting or they don’t go well with this wine”), issue complex prescriptions “Sally, you are not to leave the house, but if you have to, then take your brother with you”) and to formulate many other kinds of statements that are not so naturally assimilated to descriptions of reality.

The point is not that logically complex statements cannot represent reality; rather the point is that the logical vocabulary allows us to undertake a different kind of commitment. It’s not a commitment directly to reality’s being a particular way but rather a commitment to reason in particular ways, sometimes with regards to embedded statements about how reality is, but sometimes with regards to other sorts of statements. In either case, in line with the general program, inferentialists will identify this complex sort of commitment in terms of responsibilities for giving reasons that support the specifically logical connection asserted in the statement and entitlements to rely on this logical connection in making further inferences about reality, but also about other things.

A slightly more controversial example is probabilistic vocabulary. Many philosophers and linguists think words such as ‘might’ and ‘likely’ function primarily as ways to affect the level of confidence one commits to when making a statement. Instead of saying “It will rain,” one says “It might rain” or “It’ll likely rain.” These statements don’t undertake full commitment to rain, rather they modulate such a commitment. Accordingly, we could think of the words ‘might’ and ‘likely’ as playing a modulating rather than representing role in ordinary discourse. This means that although they can be caught up in discussions about how reality is, probabilistic terms themselves should not be thought to refer to or represent any particular bit of reality: mightness and likelihood are not things we refer to in describing reality.

Then, the inferentialist thesis would be that probabilistic vocabulary is not for undertaking commitments about intrinsically probabilistic features of reality but instead for undertaking commitments about how to rationally manage varying confidence levels in the truth of other statements (both descriptive and nondescriptive) in our collective discussion. For example, someone who says “It might rain,” could be interpreted as defeasibly disallowing any further statement, action, or attitude recognized to be inferentially incompatible with the truth of “It will rain.” And they could be interpreted as defeasibly taking responsibility for providing reasons to think rain is not ruled out (even if it is not confirmed) by propositions taken to be common ground in the conversational context.

4. A Sketch of a Theory

We’ve already said that a Peircean alternative to expressivism would emphasize the commissive function of normative vocabulary over the expressive function. And we’ve
explored some examples of how an inferentialist account of the commissive function of other vocabularies might make space for the idea that not all declarative language is straightforwardly descriptive. The next step is to flesh out an inferentialist treatment of normative vocabulary. To do this, we’re going to need a more detailed account of the discursive responsibilities and entitlesments that normative vocabulary functions to register on the conversational scoreboard, and an explanation of how these are different from the discursive responsibilities and entitlements that descriptive vocabulary functions to register on the conversational scoreboard.

Let’s start with a toy example. Sometimes we make statements such as “White always moves first,” to describe an empirical regularity (imagine, e.g., a novice observing a chess tournament). Similarly, sometimes we say “This is a game where white moves first,” to describe the fact that certain rules are recognized by others to govern a practice (imagine, e.g., someone reporting on the rules of a game they’ve never played). But we also use such statements for another purpose which looks nondescriptive. When inventing the game of chess, one might say “White always moves first.” Or when teaching the game to someone, one might say “This is a game where white moves first.” It’s plausible to think these statements institute or acknowledge a rule rather than describe its adoption or represent empirical regularities due to its being in force.

More specifically, in the latter contexts but not in the former, these statements affirm the support some facts give for acting or reacting in a particular way. Even more specifically, we could say that both of these statements undertake a commitment to the goodness of a practical inference from certain contextually salient facts, e.g., the fact that we’re playing chess, to an action, e.g., playing such that white moves first. So, if it’s agreed that we’re playing chess (descriptive claim), then with one of these statements, the speaker can license playing such that white moves first (action). As a result, if you’re white and the speaker goes on to criticize or sanction you for moving first, then you’d be justified in saying: “Wait a second, you just committed yourself to the validity of this rule! Are taking that back?”

This example involves explicit rules for how to play a game, and the conversational context rather than specific vocabulary is used to make clear that the relevant statements are statements of a rule rather than descriptions of reality. By contrast, the target of metaethical inquiry is normative vocabulary. And it usually pertains to less explicit norms that cover attitudes as well as actions. Still, I think we can lean on the model of nondescriptive statements of a rule of a game for beginning to understand the inferential responsibilities and entitlements determining the meaning of normative vocabulary.

To see how this could work for a particular word, notice that we can easily replace the statements above about the rules of chess with a statement using ‘must,’ e.g., “White must move first.” As before, this statement has both descriptive and nondescriptive uses, but the nondescriptive use is (the inferentialist claims) more fundamental to explaining the meaning of deontic ‘must.’

Relatedly, ‘ought’—the normative term par excellence—is typically used to discuss something more like strong recommendations than strict rules. But we can get the same distinction between using ‘ought’ descriptively and using it nondescriptively, and the normative usage again is plausibly claimed to be more basic to an account of its meaning. For example, imagine the chess teacher explaining to the novice, “If the opponent is behind in development, then you ought to open some lines.” In this case, it’s plausible to interpret the teacher as articulating a norm (a cousin of stating a rule) for the purpose of affirming the rational support between
certain facts and an action. More specifically, we might view the teacher as having undertaken a commitment to the goodness of a practical inference: something like an inference from the fact that the opponent is behind in development and one wants to win to the action of opening some lines. This is why the novice could reasonably ask the teacher something like, “Wait, could you explain why opening some lines at this point will help me win the game?”

When sincere, the teacher can be said to believe this fact justifies that action. But, in a Peircean spirit, the inferentialist will want to insist that the commitment undertaken by the statement is not a private mental state, but rather a public discursive status. It should be represented on any conversational scoreboard that keeps track of who is responsible for which claims. It is part of the complex web of expectations and entitlements to reason together in various ways based on common knowledge that we constantly negotiate with each other via conversation. All of this will be true whether or not the commitment has psychological reality as a belief for the teacher. This is also why an inferentialist will insist on differentiating affirming rational support for from being disposed to infer from.

So far, that’s only the beginning. We clearly use normative vocabulary not only to state rules/norms but also to make statements understood as normative consequences of rules/norms that have been previously instituted or accepted, whether explicitly or tacitly. For example, we could imagine the chess novice reasoning from the instructor’s statement “It’s wrong to move the rook diagonally” to the conclusion “So, I cannot move my rook from a1 to b2.” Although the instructor’s statement can be interpreted as the statement of a rule, the student’s statement seems instead to draw a consequence from a rule. Similarly, if the player concedes her queen early, the teacher might say, “You shouldn’t worry as that just leads to defensive play” or “It’d be wrong to think victory is unlikely.” These statements aren’t plausibly interpreted as stating rules/norms, but they are plausibly interpreted as communicating normative consequences of rules/norms. They license thoughts and attitudes, which means that the player is thereby shielded from certain sorts of sanction or criticism from the teacher if they think and feel certain ways. And they make the teacher defeasibly open to certain sorts of requests for further justification of the rational support between facts and thoughts/attitudes they have committed to in making these statements. This is why we could imagine the player saying, “Wait, I’ve conceded my queen (descriptive claim), can you explain why that doesn’t rationally support worrying and playing more defensively? Isn’t it reasonable for me now to think victory is unlikely?”

In all of these cases, the statements plausibly involve the use of normative vocabulary to affirm the rational support between some facts and some action or attitude. On the Peircean approach, this affirmation is understood as the undertaking of a specific kind of discursive commitment. Like all discursive commitments, the commitments undertaken by normative statements are constituted by the responsibilities and entitlements that are created in ordinary conversation. They are conceived as elements of public discursive space rather than private psychological reality. But unlike commitments to reality’s being a particular way (what is characteristically undertaken using empirical vocabulary to describe something), a metaethical inferentialist will see affirming rational support for as different from describing some normative feature of reality. Broadly speaking, it’s about how to reason with our other commitments rather than about how reality is. In this sense, deontic ‘must’ and ‘ought’ can be seen as a kind of logical

---

14 See Thomasson (2020), ch. 2 for helpful discussion.
vocabulary; they facilitate collective reasoning about how to act and react rather than tracking things in our environment.

Of course, one could be a realist about rational support relations, and one could think normative statements represent these normative bits of reality, in which case one might embrace a kind of Peircean approach that classifies normative statements as descriptive along with statements such as ‘Jane is eating an apple.’ However, it is by no means obvious that the Fregean starting point about linguistic meaning should be extended to cover terms such as ‘must’ and ‘ought.’ Semantically, they look much more like ‘might’ and ‘likely’ where the default view is surely to treat them as nonreferential. And the history of 20th century metaethics teaches us the perils of treating normative vocabulary as referential but not empirical.

Generalizing from these examples, then, my suggestion for a foundational thesis for metaethical inferentialism is that the meaning-determining function of normative vocabulary is to undertake commitments to rational support relations between facts and actions/reactions. These are not commitments to reality being a particular way. They’re commitments to conducting discursive-inferential practice in a particular way. They’re like practically instituting or acknowledging the rules of a game rather than describing those rules or the empirical realities that result from their adoption. On this view, when someone uses a sentence of the form “S ought to do/think/feel x” to make a normative statement, they are implicitly committing to there being certain facts that would justify S’s doing/thinking/feeling x, and in doing so they are licensing S to do/think/feel x whenever it is mutually agreed that the relevant facts obtain. This means that S is immune to certain sorts of sanction or criticism from the speaker when they do/think/feel x. And the speaker is potentially responsible for backing up the license they have granted by explaining why the relevant facts would justify those actions, thoughts, or attitudes.

5. Elaborations and Clarifications

At this stage, I imagine many readers will wonder what the inferentialist is going to say about logically complex statements with normative elements. Can we get an analog of the expressivists’ embedding problem?

I don’t think so. To be sure, normative vocabulary sometimes figures in logically complex statements that embed normative elements. But inferentialists work with a general account of the meaning-determining function of all statements as devices for undertaking discursive commitments. Any statement, however logically complex, will thus be treated as a proposed addition to the conversational common ground, which registers on the conversational scoreboard as a defeasible responsibility for the speaker to back it up with reasons and a defeasible entitlement for everyone to appeal to it as a reason in defending other things. Because of this, inferentialists can blandly say that logically complex statements commit the speaker to the truth of a logically complex proposition. Importantly, however, that’s not yet meant to imply anything about the relationship between those statements and reality. For inferentialists, the issue of which statements carry representational content is theoretically downstream from their uniform treatment of the commissive function of all statements.

Still, the question remains: What according to the inferentialist is the commitment undertaken by logically complex statements embedding normative elements, and what ensures that it doesn’t involve any representation of normative bits of reality?
There are perhaps several ways for inferentialists to answer this\textsuperscript{15}, but the one I find most plausible leans on the inferentialist treatment of logical and probabilistic vocabulary as undertaking commitments to reason together in various ways rather than to the reality of logical or probabilistic relations. As we saw above, inferentialists generally maintain that logically complex statements don’t directly represent reality. In saying “Either it’s raining or the sun is shining” or “If it’s raining, then sun will be behind clouds,” one doesn’t commit to reality containing disjunctivity or conditionality. Rather, the idea is that one commits to a pattern of collective reasoning regarding the embedded parts of the statements.

These statements might still be said to carry conditional descriptive content. In the first case, if we were also to reject the proposition that it is raining, then we’d be defeasibly entitled to conclude that the sun is shining. And this conclusion can be viewed as a representation of reality. So, we can view this representation as descriptive content conditional carried by the original disjunctive statement. Similarly in the second case. If we were also to accept the proposition that it’s raining, then we’d be defeasibly entitled to conclude that the sun will be behind the clouds. This entitlement to a representational proposition can then be viewed as descriptive content conditionally carried by the original conditional statement.

Parallel story for disjunctive statements with a normative disjunct, such as “Either Alex is there or Sarah ought to go.” The inferentialist will again insist that this doesn’t commit us to reality containing disjunctivity. Rather, it commits us to reason in certain ways regarding the two disjuncts. As before, if we also reject the proposition that Sarah ought to go, then we’re entitled to say Alex is there. This statement can be viewed as a commitment to reality being a particular way. So, the original disjunctive statement can be said to carry conditional descriptive content. However, it also carries conditional normative content. If we deny that reality contains the fact that Alex is there, then we’re entitled to the normative proposition that Sarah ought to go. Accepting this is not to represent reality but to acknowledge the rational support the action of Sarah’s going has from other propositions we’ve agreed to.

Similarly, in saying “If James ought to pay, then he will pay,” one doesn’t directly commit to reality containing conditionality; one commits to reason in certain ways with others regarding the antecedent and consequent. So, again, the statement can easily be understood to carry conditional descriptive content. If we accept the proposition that James ought to pay, then we’re entitled to say that he will be there, which (modulo questions about temporal modality and the open future) can be viewed as a commitment to reality being a particular way.

The story is slightly different for statements embedding normative vocabulary under probabilistic operators such as ‘might.’ But only because probabilistic vocabulary plays a different kind of role in the sort of collective reasoning that goes on in ordinary conversations. There are complexities I have to gloss over, but a starting suggestion is that ‘might’ functions

---

\textsuperscript{15} Warren (2015) argues that morally ‘ought’ has the meaning-determining function of facilitating practical reasoning, which it retains in embedded contexts. I worry that this doesn’t cover all uses of normative ‘ought’ that we should want to be nonrepresentationalist about. Incurvati and Schlöder (2022) motivate a distinction between the attitudes of assent and dissent and then use this to explain the meaning of the logical connectives ‘not’ and ‘or’ in terms of which attitudes of assent or dissent are inferable from any use of these connectives (including embedded uses). This combines an expressivist semantics of normative terms with an inferentialist account of the logical connectives in a way that solves Unwin’s (1999) negation problem. But I’m inclined to think this distinction should be mobilized in the first instance to characterize the social discursive space of the inferential commitments undertaken by linguistic acts, and only derivatively to characterize the private psychological space of individual’s mental states.
to modulate the epistemic strength of the commitments one undertakes. And, for inferentialism, this will be reflected in the kinds of discursive responsibilities and entitlements that get registered by use of an ordinary might-statement. This is true whether we’re talking about the weather or what people ought to do. So, inferentialists will view someone who says “It might be the case that Bob ought to wait,” as defeasibly disallowing any further statement, action, or attitude recognized to be inferentially incompatible with the truth of “Bob ought to wait.” And they could be interpreted as defeasibly taking responsibility for providing reasons to think this normative truth rain is not ruled out (even if it is not confirmed) by propositions taken to be common ground in the conversational context.

This explanation of the meaning of statements that embed normative elements under various logical operators will remind many readers of Blackburn’s (1988) suggestion that a speaker who makes a logically complex statement is thereby “tied to a tree.” His idea was that a disjunctive statement with evaluative and descriptive disjuncts does not directly express an attitude of booing/hooraying or a belief about reality. Rather, it expresses a commitment “to accepting the one branch should the other prove untenable” (1988, 512). That sounds pretty similar to the inferentialist treatment just sketched. So, why do I think expressivists like Blackburn have a problem with embedding, and how is the inferentialist proposal any better?

I think Blackburn’s suggestion nicely highlights the possibility of a nonrepresentationalist treatment of logically complex statements. But I think the distinction between Lockean and Peircean frameworks is important for making this possibility work. Insofar as expressivists want to appeal to the idea of being tied to a tree to explain the meaning of logically complex statements, they’ll need to tell us enough about the mental state expressed by the statement which ties its author to a tree. That is, they’ll need to show how the statement gets its logico-semantic properties from the mental state it expresses. This is just the embedding challenge we encountered at the beginning: if one maintains that descriptive statements express beliefs about reality and normative statements express desire-like attitudes, one owes an account of mental state expressed by a disjunctive statement embedding descriptive and normative elements.

For his part, Blackburn suggests that the commitment these statements express is a third kind mental state, the state of being tied to a tree, which something like a higher-order state taken towards combinations of first-order states of believing and booing/hooraying. However, as Schroeder (2010, 139–141) has argued, this threatens to take expressivists down the slippery slope of needing to posit an infinite hierarchy of kinds of mental states to handle the more and more complex statements that can be formed by logical embedding. Maybe this is a theoretical cost worth paying. But if expressivists aspire to elucidate features of normative language by appeal to its role in expressing our mental states, then they should prefer a more empirically tractable and less controversial account of the nature and contents of those mental states.

Alternatively, an expressivist might argue that the notion of being tied to a tree is not about a distinctive kind of higher-order mental state but instead a description of the logico-semantic properties that the nonrepresentational states expressed by the relevant statements must have. But then Schroeder’s complaint is that expressivists are helping themselves to what their account of the meaning of logically complex statements is meant to explain. We want to know why the statement “Either Alex is there or Sarah ought to go” means something disjunctive, and they tell us it’s because the statement expresses a commitment to one of the disjuncts if the other one gets rejected. But when asked what this commitment is, we’re told something like: it is a mental state characterized by the fact that it “disagrees” with any other mental state that disagrees with both disjuncts. This way of going opens expressivists up to Schroeder’s (2010, ch. 7) argument that they haven’t explained why the mental state has this property. Insofar as
expressivists hope to explain the meaning of the statement in terms of the content of the mental state it expresses, they seem to owe such an explanation.

Inferentialists, in contrast, do not seek to explain the meaning of statements in terms of the mental states they express. If anything, they’ll go the other direction, where the explanation of mental (conceptual) content appeals to the semantic properties of language. This is why it’s important that the inferentialist notion of commitment is not of a kind of mental state (something in psychological space) but instead of a kind of discursive status (something in social space). The basic Peircean inferentialist account is one where the logico-semantic properties of a statement are explained by the inferential structure of the mutually recognized discursive responsibilities and entitlements the statement can be used to undertake. For example, on this view, part of the role of ‘or’ is to let us entitle each other to affirm one of the disjuncts if the other is or becomes rejected in our ongoing conversation. This is a feature of ordinary discursive practice regardless of whether the disjuncts are descriptive, probabilistic, modal, normative or even imperatival. And we know this by observing what ordinary speakers do with ‘or’ in normal conversation. Similarly, part of the role of ‘might’ is to let us disentitle each other from affirming something inconsistent with the proposition it embeds. Again, this is true no matter what kind of proposition is embedded. And we know this by observing normal probabilistic conversation.

Still, I can imagine expressivists and their critics responding: “Ok, I get that inferentialists want to explain meaning in terms of social discursive statuses rather than private psychological states. But they still need some moral psychology, don’t they? What do they think the mental state expressed by logically complex statements with normative parts is like? Is it a cognitive representation of reality, a desire-like motivation to action, or something else?”

Response: Inferentialists can happily endorse a moral psychology allowing for logically complex beliefs that have normative parts. They might say the belief that either Alex is there or Sarah ought to go, and the belief that it might be the case that Sarah ought to go, are appropriately attributed to someone when doing so is part of the best rationalization of what that person does, thinks, and feels. At least this is a plausible starting point for general theory of belief that fits well with inferentialism. Furthermore, it’s plausible to think that someone who believes a proposition (including a logically complex proposition) will be defeasibly disposed to undertake commitment to its truth in normal conversation, when this truth is relevant. As before, and quite generally, inferentialism provides an account of this commitment in terms of the intersection of the speaker’s defeasible responsibility to justify acceptance of the proposition and the defeasible entitlement for all participants to the conversation to treat the proposition as a premise in further collective reasoning.

So, when it comes to the moral-psychological project of explaining how beliefs with particular contents are linked to dispositions to do, think, and feel particular things, I think inferentialists can draw on investigation of the (likely quite complex and socially-infused) psychological underpinnings of discursive responsibilities and entitlements. They can also appeal to rationalizing explanations of why people do, think, and feel whatever they do, think, and feel. As long as these sorts of explanatory projects allow for the positing of logically complex beliefs with normative elements, there’s no theoretical pressure on inferentialists to say that logically complex statements embedding normative elements express some desire-like state, which then leads to the primary challenge facing expressivism.

Of course, that’s not to say they don’t face other challenges.
6. Objections and Replies

Because the inferentialists I mentioned above analyze meanings primarily in terms of inferential responsibilities and entitlements, they are sometimes interpreted as holding that all vocabularies—including empirical vocabulary—are “fraught with ought.” Accordingly, one might worry that inferentialism treats all vocabulary as normative and so collapses the distinction between normative and descriptive discourse. Indeed, this interpretation might encourage the impression that inferentialism is covertly some kind of global expressivism.\(^{16}\)

While recognizing this is a route from inferentialism back to a kind of expressivism, I don’t think it is the way most inferentialists understand their view. Insofar as they embrace a Peircean rather than Lockean explanation of linguistic meaning, they will hold that meaning bottoms out in the public social space of mutually recognized discursive commitments rather than in the private psychological space of individuals’ mental states. Moreover, in the context of metaethics, because we want to explain the differences between normative vocabulary and empirical vocabulary, we shouldn’t develop an inferentialist view of normative vocabulary into a global form of expressivism.\(^{17}\)

Moreover, if the kind of metaethical inferentialism sketched above is a cogent way to argue that normative vocabulary is not representational of normative aspects of reality, it has two important attractions that would be forfeited by turning it into a global form of expressivism.

First, the inferentialist can say that statements deploying normative concepts express beliefs just like any other statement, and the logico-semantic properties of these beliefs can be explained by their propositional contents rather than their role in motivating action or manifesting affect. This means that the inferentialist avoids appealing to a meaning-determining psychofunctional difference between normative beliefs and representational beliefs, which is precisely what opened the expressivist up to the problem of logically mixed statements.\(^ {18}\)

Second, unlike contemporary expressivists, inferentialists don’t have to adopt a minimalist account of belief or a deflationary account of truth in order to “earn the right” to call normative attitudes beliefs, treat normative statements as truth-apt, or recognize some of the beliefs they express as knowledge. To be sure, inferentialists will need to adopt a theory of truth consistent with the idea that there are some true statements that do not represent reality. There are many richly-developed and well-defended pluralist and coherentist alternatives to the correspondence theory of truth which are helpful here. But, from the beginning, the inferentialist will adopt a uniform understanding of the relationship between a statement, its propositional content, and the mental state it expresses, rather than starting out with a representationalist conception of belief and seeking to argue that normative attitudes are enough like these beliefs that they too can be called “beliefs” even though they are really desire-like. Similarly, the inferentialist starts out from the idea that there are lots of different kinds of

\(^{16}\) For discussion, see Price (2019) and Simpson (2020).

\(^{17}\) For a similar point, see Brandom (2013).

\(^{18}\) Commitment to a strong form of motivational internalism and the “Humean” theory of motivation might lead one to see this as a problem for the view. However, I think inferentialism actually benefits in comparison to expressivism from not needing to appeal to implausibly strong forms of motivational internalism to motivate the claim that normative vocabulary is nonrepresentational. For further discussion, see Chrisman (2022a), 180–183.
concepts, and so lots of different kinds of knowledge. Only some of this (empirical knowledge) involves reliably tracking reality, but that doesn’t impugn the more general project of accounting for the discursive commitments undertaken by someone who makes a knowledge attribution in terms of truth, belief, reason, and connections between these—arguably, this is precisely what philosophical work on the theory of knowledge is trying to figure out.

Perhaps some staunchly representationalist readers will think: “Perhaps metaethical inferentialism avoids some of the problems of expressivism while retaining the intuitive idea that normative language plays a different role than descriptive language. But to maintain that all statements express beliefs in a univocal sense of ‘belief, while insisting that statements deploying normative vocabulary do not represent reality, one must abandon the idea that part of what it is to be a belief is to represent reality. And to maintain that logically complex sentences with normative elements have ordinary truth-apt propositions as their contents, while insisting that statements made with these sentences aren’t straightforward representations of reality, one must abandon the idea that part of what it is for a proposition to be truth apt is for it to represent a way reality could be.”

If one regards these as costs, then yes these are costs of the view. But I think these are theoretical commitments that anyone motivated by the distinctive function of normative vocabulary to develop a nonrepresentationalist account of normative language should already embrace. One of my main suggestions in this paper is that there is an attractive way to embrace these commitments without maintaining that normative statements get their meaning from expressing desire-like attitudes.

Another worry some might have about inferentialism in this context is that its overall explanation of meaning uses normative language (e.g., what one is discursively committed to, what responsibilities and entitlements are carried by adding something to the common ground, etc.), and so it cannot provide a reductive, non-normative, or “sideways on” description of how language gets meaning. Thus, if one wanted to use inferentialism to make sense of the idea that normative vocabulary doesn’t represent reality, it may seem that one is thereby committed to irreducibly normative descriptions of that part of reality that involves language having meaning.

But the availability of reductive explanations is a notoriously bad test for whether or not a concept picks out some feature of reality. We cannot reduce logical and probabilistic concepts to non-logical and non-probabilistic concepts, but that doesn’t immediately commit us to the idea that reality includes disjunctivity and mightness. Furthermore, the inferentialist’s core claims can be interpreted like any other statements involving normative vocabulary. There is a perfectly coherent use of the relevant sentences to describe the fact that some community accepts a norm or to report the observation of empirical regularities due to this acceptance. And sometimes, as a bit of linguistic science, we adopt this sideways on stance when explaining the semantic norms that characterize a particular fragment of language. But there’s also a use of the relevant sentences to affirm or reason with various semantic norms from within the practice of speaking the language. And inferentialists are going to insist that this normative use is the more fundamental to constituting the meaning of the relevant normative vocabulary. Norm-accepting language users could exist even if there were no descriptive semanticists, but we couldn’t do descriptivist semanticists without first being norm-accepting language users.
7. Conclusion

By accounting for the meaning of statements primarily in commissive rather than expressive terms, the Peircean approach to explaining meaningfulness encourages us to focus on the social space of discursive responsibilities and entitlements rather than the psychological space of different kinds of attitudes. And by appealing to the inferential role of different vocabularies rather than to the psychological role of different kinds of mental states the inferentialist version of this approach offers an account of the meaning of normative vocabulary that connects it to rational discussion about what to do, think, and feel without committing to the problematic idea that normative vocabulary functions to express a fundamentally different kind of mental state than descriptive vocabulary. In this way, I hope to have offered the beginnings of an inferentialist alternative to expressivism.

I wouldn’t have tried to do this if I didn’t think the resulting view is more attractive than extant versions of expressivism. However, I don’t imagine that counting up plausibility points is really going to convince anyone that inferentialism is the right account of normative discourse. All seriously defended metaethical views are by now highly sophisticated and rooted in rich theoretical traditions, which makes me doubtful of the probative value of our intuitions about whether this or that tenet of a view is a theoretical cost or benefit. Moreover, my own reasons for liking the application of inferentialism to metaethics have much more to do with the underexplored questions this approach encourages us to ask than with the comparative plausibility of any particular theoretical claims. So, I want to conclude the paper by briefly discussing the new perspective inferentialism offers on three further metaethical issues.

One issue is about the nature of normative concepts, which in turn raises important questions about the nature of concepts in general. Debate between expressivists and their critics has tended to focus on normative language and move quickly to questions about the kind and content of mental states expressed by normative statements. This way of approaching the issue has encouraged some philosophers to deny the existence or even possible coherence of genuinely normative concepts. And I suspect it often encourages the assumption that concepts are best understood as elements of individual’s psychological economy that are about bits of reality, something like ways a person thinks about (or “conceptualizes”) a property or relation.

Maybe that’s the right view, but inferentialist philosophers tend to work with more social and less psychological pictures of concepts, which has room for the idea that the content of concepts is determined by inferential roles, which itself may be something that is not fixed by the psychology of any one person. Moreover, this idea is founded on the thought that vocabulary such as ‘or,’ ‘if-then,’ and ‘might’ express concepts but are not ways of conceptualizing real properties or relations. Accordingly, I’m inclined to think that normative concepts have content more like these non-referential logical concepts (especially if we treat the modal concept ‘ought’ as a paradigmatic and central normative concept). As a result, I think it’d be useful for us to think more about the possibility that the content of normative concepts is socially extended/constructed in ways that cannot be reduced to the way any particular individuals conceptualize bits of reality.


20 For helpful discussion, see Millgram (2015) ch. 6., which I compare and contrast with the view implied by Chrisman (2016) in Chrisman (2019).
Another issue highlighted by inferentialism is about what it takes to back up a normative statement if challenged and how entitlement to rely on someone else’s normative statement in further reasoning is transmitted. In our traditional focus on the question of what normative statements represent or express, I think metaethicists have tended not to think hard enough about the rational structure of the defense of our first-order normative views, and why this matters for their justification. Some recent discussions of the comparative oddness of forming a normative belief based solely on the testimony of another person touches upon this question. However, there is a tendency in these discussions to view normative testimony in implicitly Lockean terms. That is, testimony is often assumed to be the communicative process whereby one person communicates what they believe to another person, who then must decide whether to accept this belief.

Inferentialism encourages us instead to view all testimony as proposed additions to the common ground, whereby one says something whose communicative function is to acquire an inferential responsibility and to grant inferential entitlements. This suggests an approach to thinking about normative testimony focused not so much on the reliability of the psychological process by which a testifier forms normative beliefs but more on the availability to collective discursive practice of considerations that would count in a particular conversational context as backing up a normative statement in the face of challenge. If we think normative statements affirm rational support relations between various facts and certain actions or attitudes, then it shouldn’t be surprising that the defensibility of such statements is different from the defensibility of statements about reality. And it is an interesting challenge, highlighted by inferentialism, to explain this difference and bring it to bear on the question of how and when we are entitled to rely on someone’s normative testimony.

A final issue is about what it takes to be a good normative interlocutor. In the public domain, there is heavy focus on sincerity and authenticity, conceived as something like really believing what one says and not acting in ways that hypocritically contradict the attitudes revealed when one speaks. This is an important expressive virtue, but there are other things we should want from our interlocutors in public political settings. It is usually important that participants not only state their sincere views but also seek to collectively reason with others about what to do, think, and feel. In thinking about communicative virtues, inferentialism encourages us to look not mainly to the psychology of sincerity and authenticity but to participants’ commitment to giving reasons for what they say and to modulating what has been said in the face of countervailing reasons given by others.

---

21 See especially Hills (2009), and Fletcher (2016).

22 For further discussion, see Chrisman (2022b).

23 For helpful discussion, I am grateful to Mahrad Almotahari, Berislav Marušić, Graham Hubbs, Sebastian Köhler, Barry Maguire, Eduardo Pérez Navarro, Makan Nojoumian, Michael Ridge, Andrés Soria Ruiz, Stefan Sciaraffa, two anonymous referees and participants of the 2021 Madison Metaethics Workshop.
References


