
Article:
Brown, J.L.D. orcid.org/0000-0003-3875-592X (2025) Moral inferentialism and moral psychology. Synthese, 206 (1). 26. ISSN 0039-7857

https://doi.org/10.1007/s11229-025-05102-7

Synthese (2025) 206:26
https://doi.org/10.1007/s11229-025-05102-7

ORIGINAL RESEARCH

Moral inferentialism and moral psychology

James L. D. Brown[1]

Received: 30 August 2024 / Accepted: 30 May 2025

Abstract
This paper raises a challenge for moral inferentialism. Moral inferentialism explains moral discourse in terms of the distinctive kinds of discursive commitments we acknowledge and undertake in making moral claims. However, like any metaethical theory, inferentialism owes us an account not only of what it is to make moral claims, but of what it is to think moral thoughts. The paper argues that what inferentialists have said about moral thought is unsatisfactory. While more satisfactory accounts are available, adopting such accounts robs inferentialism of certain supposed advantages that it enjoys over its main competitor, moral expressivism. Insofar as inferentialism is motivated as a non-representationalist alternative to expressivism, the challenge therefore undermines a central motivation for inferentialism.

Keywords Moral inferentialism · Inferentialism · Expressivism · Metaethics · Moral judgment · Moral belief

1 Introduction

Are moral statements about how things stand in the world? Or are they about something else, perhaps nothing at all? Many view this as the central question of metaethics. In one way or another, our answer will either vindicate or debunk the moral practices that are so central to our lives. Those who answer 'yes' to the first question are *representationalists* about the moral domain. They come in many different stripes. Some believe that moral statements are made true by some natural part of the world, others by some non-natural part, and still others by no part at all, leaving moral practice in systematic error. Thus, if we are representationalists, the vindication of morality will depend on whether reality supplies the right kind of object for our practices

✉ James L. D. Brown
j.ld.brown@sheffield.ac.uk

[1] University of Sheffield, Sheffield, UK

Published online: 01 July 2025

to be responsive to. By contrast, those who answer 'no' to the first question reject the idea that the vindication of morality depends on any robust moral metaphysics. For they believe that morality is not in the business of representing moral reality in the first place. These people are *non-representationalists*. They think that all we need is some account of what we are *doing* when engage in moral practice. If this practice is in good order, morality is thereby vindicated.

For a long time, if you were a non-representationalist, you were probably an *expressivist*. Expressivists aim to explain moral practice in terms of the practical attitudes and stances we express when making moral claims. Throughout the history of metaethics, expressivism and its forebears have enjoyed a dominance—one might say hegemony—over the non-representationalist theoretical landscape. There is, however, a contender beginning to emerge that aims to unsettle expressivism's seat at the head of the table. The contender is *moral inferentialism*. It aims to explain moral practice by applying a certain kind of inferentialist theory of meaning from elsewhere in philosophy to the moral domain. Like expressivists, inferentialists deny that moral statements represent moral reality, at least in any robust sense. Unlike expressivists, however, they deny that moral statements express practical attitudes and stances. Instead, they maintain, moral statements codify certain kinds of inferential commitments. The hope is that this retains all the non-representationalist benefits of expressivism while at the same time avoiding its most serious pitfalls. Thus, if we are going to be non-representationalists, the argument goes, we would do better being inferentialists rather than expressivists.

While interest in moral inferentialism has been steadily growing, most of the literature has been concerned with articulating the view as a distinctive metaethical position.[1] It therefore now seems like a good time to take stock of the emerging landscape. To this end, the paper has two central aims. The first aim is to provide an overview of the current landscape in its essentials (Sect. 2). The second aim is to examine whether non-representationalists really should be inferentialists rather than expressivists (Sect. 3). I will raise some doubts concerning three supposed advantages of inferentialism over expressivism, but the main challenge I will focus on relates to *moral psychology*. Like any other metaethical theory, moral inferentialism owes us an account not only of moral language but also of moral *thought*—what it is to think moral thoughts, to have moral attitudes, and so on. On this front, inferentialists have less to say than in relation to moral language, which is unsurprising given the linguistic focus of inferentialism more broadly. However, I will argue that extant inferentialist accounts of moral thought are unsatisfactory. Moreover, I will argue that the most plausible way of providing such an account leads us back to expressivism—or at least, to something sufficiently close to it to raise doubts that inferentialism is preferable to expressivism. I will conclude by discussing whether this is a genuine problem for moral inferentialism, or whether it is simply a problem for the idea of inferentialism as an alternative to expressivism (Sect. 4).

[1] See, for example, Brandom (2000, 2008), Chrisman (2008, 2010, 2012, 2013, 2016, 2018, 2023), Warren (2015); Tiefensee (2016, 2019, 2021, 2023), Karimi (2021), Incurvati and Schlöder (2021, 2023).

2 Moral inferentialism

In this section, I explain the basic tenets and motivations of moral inferentialism. After outlining these programmatically (Sect. 2.1), I then distinguish between two dominant strands of moral inferentialism that take their lead from *logical expressivism* (Sect. 2.2) and *logical bilateralism* (Sect. 2.3) respectively. Since the overall aim of this section is simply to provide an outline of the landscape, familiar travellers may wish to skip to the next section.

2.1 Basics

The key idea behind moral inferentialism is to apply a more general inferentialist approach to theorizing about language and meaning to the moral domain. Inferentialism as a theory of meaning aims to explain linguistic meaning in terms of *inferential role*. The inferential role of an expression consists in a set of rules governing its use in inference. The inferential relations specified by an inferential role codify *commitments* acknowledged and undertaken by the speaker. Such commitments are not a kind of psychological state; rather, they are *socially* instituted in a public linguistic practice. In this way, inferentialism explains meaning without recourse to representational notions like truth, reference, satisfaction, and so on. While some expressions might be taken to represent reality in virtue of the specific commitments encoded by their inferential roles, meaning itself is not inherently representational. And it is precisely this feature of the view that makes inferentialism an attractive framework for developing a non-representationalist account of moral discourse. For it allows us to explain of the meaning of moral terms in terms of their inferential roles, which encode non-representational, publicly instituted, inferential commitments.

The key question for any version of moral inferentialism is precisely how to understand the inferential roles for moral vocabulary. But the above sketch already suffices to understand the motivations for accepting inferentialism over expressivism. The first motivation is that moral inferentialism retains all the attractions of non-representationalism enjoyed by moral expressivism. Like inferentialism, expressivism avoids explaining moral discourse using representational notions. Instead, expressivism explains the meaning of moral terms in terms of the distinctive practical attitudes we conventionally express in using such terms. Thus, while inferentialism and expressivism offer very different accounts of moral discourse, both theories offer vindicatory explanations of moral discourse without positing any putatively problematic domain of moral entities that moral discourse is about.

The second motivation is that moral inferentialism does not face the Frege-Geach problem, or the problem of explaining the meaning of embedded moral terms. [2] The problem arises for expressivists because the meaning-constituting attitudes purportedly expressed by moral terms are typically only expressed in unembedded contexts. Thus, when a moral term is embedded in a conditional, attitude ascription, question, etc., it does not express the attitude that purportedly explained its meaning in unembedded contexts. By contrast, because inferentialists explain logical vocabulary as

[2] For an overview of this problem, see Schroeder (2008), Woods (2017).

operating on commitments (more on which below), no special story is needed to explain how moral commitments embed, since inferential meaning is stable across embedded and unembedded contexts.

The third motivation is that moral inferentialism offers a more straightforward moral psychology. Since expressivism aims to offer psychologistic explanations of moral practice, its account of moral thought must satisfy a number of desiderata. Specifically, its account of moral thought must explain moral meaning, the role of moral thought in our cognitive economy, how moral disagreement arises, how moral thought relates to other kinds of conation, why moral thought is non-representational, and so on. Needless to say, constructing a moral psychology that satisfies these desiderata is no simple task, and any attempt is bound to be controversial. By contrast, since moral inferentialism provides a non-psychologistic explanation of moral meaning in terms of inferential commitments, it incurs no such explanatory burdens in relation to what it says about moral thought. Indeed, since inferentialism denies that there is anything special about moral assertion, it can maintain that moral claims are ordinary assertions that express beliefs just like any other assertion.

I will shortly go on to describe moral inferentialism in a little more detail, focusing on two distinct strands that model moral meaning on two different kinds of inferentialist views in the philosophy of logic. Before proceeding, however, some clarifications regarding the above characterization of moral inferentialism. First, there may be other approaches to explaining meaning that might appropriately be called 'inferentialist' that do not meet the description given above. Second, and relatedly, there may be other views in metaethics that deserve the name 'moral inferentialism' that do not meet the description of this view given above. In line with those I am discussing, however, I will reserve the terms for only those views with the above commitments.[3] Third, I have so far loosely talked of 'explaining meaning'. This might be taken either as a *semantic* thesis—explaining what an expression's meaning *is* or what it *consists in*—or as a *meta-semantic* thesis—explaining what *determines* an expression's meaning or explaining *that in virtue of which* an expression means what it does. While moral inferentialism is sometimes described as a purely meta-semantic view, this cannot be right, since this is also compatible with a representationalist semantics.[4] However, this distinction will not be too important in what follows, so I will often continue to talk loosely without specifying whether the relevant explanations are best thought of as semantic or meta-semantic. Fourth, while the primary focus of this paper concerns *moral* thought and language, it is widely thought that the proper object of study for metaethics is not morality narrowly construed but *normativity* more broadly. However, since most of the views I discuss are formulated in specifically moral terms, I will gloss over this distinction in most of what follows.

[3] For examples of other approaches to explaining meaning that might be described as inferentialist, see, e.g., Block (1986), Peacocke (1992), Chalmers (2021). For examples of other approaches to explaining moral discourse that might be described as inferentialist, see, e.g., Wedgwood (2001, 2007). A more accurate name for the approach discussed in this paper might be 'non-representational moral inferentialism'.

[4] See the citations contained in the previous footnote; compare also the discussion in Baker (2020).

2.2 Model one: logical expressivism

Our first strand of moral inferentialism takes it leads from Brandom's logical expressivism.[5] The starting point for Brandom is a particular picture of what we are doing when we engage in discursive practice. Meaning is then explained in terms of this practice. Here is the practice:

> Saying or thinking *that* things are thus-and-so is undertaking a distinctive kind of *inferentially* articulated commitment: putting it forward as a fit premise for further inferences, that is, *authorizing* its use as such a premise, and undertaking *responsibility* to entitle oneself to that commitment, to vindicate one's authority, under suitable circumstances, paradigmatically by exhibiting it as the conclusion of an inference from other such commitments to which one is or can become entitled. Grasping the *concept* that is applied in such a making explicit is mastering its *inferential* use: knowing (in the practical sense of being able to distinguish, a kind of knowing *how*) what else one would be committing oneself to by applying the concept, what would entitle one to do so, and what would preclude such entitlement. (2000: 11)

Thus, to engage in linguistic practice is to engage in such a practice of giving and asking for reasons. The meaning of a content or claim consists in its inferential role, understood in terms of the conditions under which one is justified in asserting the claim—its 'upstream commitments'—as well as the consequences that claim licenses—its 'downstream commitments'.

Two sets of distinctions will be vital for understanding the first strand of moral inferentialism. The first is the distinction due to Sellars (1974a) between *intra-linguistic*, *language-entry*, and *language-exit* inferential roles. Intra-linguistic inferential roles encode commitments between claims (e.g., "Sarah is James's mother" → "James is Sarah's son"). Language-entry inferential roles encode upstream commitments from perception or observation (e.g., [observing a red post-box] → "This post-box is red"). Language-exit inferential roles encode downstream commitments to action (e.g., "I shall kick the ball" → [kicking the ball]). While these latter transitions are *non*inferential (Brandom, 2000: 28), it is nonetheless essential that the relevant claims *also* involve various inferential transitions.

The second is the distinction between *first-order conceptual* vocabularies and *second-order meta-conceptual* vocabularies. The former encode commitments such as those given in the paragraph above; they "enable us to make claims about the world and how to act in it." (Tiefensee, 2023: 7) The latter make explicit the inferential commitments implicit in our use of first-order vocabularies; they enable us to undertake and acknowledge commitments to reason in particular ways. The paradigm here is *logical* vocabulary, in particular the conditional:

[5] The notion of expression central to Brandom's view is the idea of *making something explicit* rather than *expressing one's attitudes*, which is the central notion for metaethical expressivism. On the differences and similarities between these two different kinds of 'expressivism', see Price (2011).

> Prior to the introduction of such a conditional locution, one could *do* something, one could treat a judgment as having a certain content (implicitly attribute that content to it) by endorsing various inferences involving it and rejecting others. After conditional locutions have been introduced, one can *say*, as part of the content of a claim (something that can serve as a premise and conclusion in inference), *that* a certain inference is acceptable. One is able to make explicit material inferential relations between an antecedent or premise and a consequent or conclusion. (Brandom, 2000: 60)

Thus, whereas we might infer "Leo is a mammal" from "Leo is a lion", we can articulate this inference explicitly with the conditional "If Leo is a lion, then Leo is a mammal".[6] Central here is the idea, due to Sellars (1954), that the primary notion of inference is not *deductive* but rather *material*, where material inferences are those inferences licensed by the *content* of a claim (such as inferring that Leo is a mammal from Leo's being a lion).

With these distinctions in place, moral inferentialism can be understood as the view that moral expressions—and normative expressions more generally—are metaconceptual devices that allow us to articulate and undertake commitments to *practically* reason in certain ways; such vocabulary "plays the same expressive role on the *practical* side that *conditionals* do on the *theoretical* side." (Brandom, 2000: 89) Thus, where one might (materially) infer "I shall not repeat the gossip" from "repeating the gossip would gratuitously harm someone", we can acknowledge and make this inference explicit with a moral claim like "It is wrong to inflict gratuitous harm". We might then add detail to this basic idea in the following ways:

> 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. (Chrisman, 2023: 9)

> [T]he metaconceptual function of moral evaluative terms concerns the systematisation of legitimate language exit transitions, where based on the method of reflective equilibrium, these systematisations ground explanations of why tran-

[6] Whereas moral inferentialists like Chrisman and Tiefensee talk in terms of the conceptual and metaconceptual, Brandom talks in terms of *consciousness* and *self-consciousness*: "In the first application, we get an account of *consciousness*—for example, *that* Leo is a lion. In the second application we get an account of a kind of semantic *self*-consciousness. For in this way we begin to *say* what we are *doing* in *saying* that Leo is a lion. For instance, we make explicit (in the form of a claimable, and so propositional content) that we are committing ourselves thereby to his being a mammal by saying *that* if something is a lion, then it is a mammal." (2000: 20).

sitions such as 'Charles needs help with the children, so I shall help Charles' are legitimate. (Tiefensee, 2023: 8)

U's assertion "one ought to ϕ" licenses the following inferences

R1: That *U* has attitudes in favor of ϕ-ing, including perhaps motivation to ϕ, a commitment to approve of one who ϕ's, and the second-order belief "one ought to be motivated to ϕ". These attitudes, commitments, and dispositions reinforce the practical clout of the moral 'ought'; their practical significance is not contingent on the desires or goals of particular agents.

R2: That *U* has similar attitudes and commitments against not ϕ-ing, including perhaps a commitment to feel guilt upon failing to ϕ, disapproval or disgust towards not ϕ-ing, perhaps to the point of effecting punishment and even pressure on third parties to react punitively to those who fail to ϕ, and the second-order belief "one ought to disapprove of not ϕ-ing.

R3: If *U* has a disposition or commitment to approve of *S*'s ψ-ing, to disapprove or feel disgust at *S* not ψ-ing, to punish *S* for not ψ-ing, to pressure third parties to react punitively to *S*'s failure to ψ, etc., and if the practical significance of these commitments is not contingent on the desires or goals of particular agents, then *U* is *defeasibly* licensed to assert that *S* ought to ψ." (Warren, 2015: 2871-2).

In each case, the meaning of moral discourse is explained in terms of commitments to practically reason in certain ways. This concludes the presentation of our first strand of moral inferentialism.[7]

2.3 Model two: logical bilateralism

Our second strand of moral inferentialism argues that inferentialists ought to incorporate the notion of attitude expression within their theoretical framework rather than dispense with it entirely. For this reason, Incurvati and Schlöder, the main proponents of this strand of inferentialism, dub the view *inferential expressivism*. Here is how they sum up the main idea: "Traditional expressivism holds that the meaning of certain linguistic items is given by the attitudes they are used to express. Inferentialism holds that the meaning of certain linguistic items is given by the inferences they license. Inferential expressivism holds that the meaning of certain linguistic items

[7] A complication: For Brandom, "there are two species of discursive commitment: the cognitive (or doxastic) and the practical. The latter are commitments to *act*. Acknowledgments of the first sort of commitment correspond to *beliefs*; acknowledgments of the second sort of commitment correspond to *intentions*." (2000: 83) However, as Chrisman (2023: 9) makes clear, normative discourse involves not only commitments to act, but also to think and feel certain ways. This is a problem because we cannot intend to believe a proposition or feel a certain way, since they are outside direct volitional control. So commitments to think and feel cannot be intentions, leaving such commitments outside the scope of Brandom's framework.

is given by the inferences to attitude expressions they license." (2023: 63). Since the view still aims to provide a non-representationalist alternative to expressivism by applying a more general inferentialist theory of meaning to the moral domain, for present purposes I will understand inferential expressivism as a variety of moral inferentialism.[8] However, the labels themselves are less important than the substantive views that accompany them.

Incurvati and Schlöder take their lead from *bilateralism* about logic. This starts with the idea that the logical connectives can be defined in terms of their introduction and elimination rules. These provide the meaning-constituting conditions under which we are entitled to affirm a claim involving that connective, as well as those conditions under which we are entitled to affirm a claim that follows from some set of claims involving that connective. For present purposes, we can think of these rules as specifying inferential roles. For example, for conjunction, we have the following:

I1. Assert A and Assert B → Assert A&B

E1. Assert A&B → Assert A

E2. Assert A&B → Assert B

This conception of the logical concepts is *unilateralist*, insofar as the meanings of sentences containing the logical connectives are explained entirely in terms of a single speech act: assertion.

A *bilateralist* conception, by contrast, explains the meanings of logical constants in terms of two kinds of speech acts: assertion and rejection. For example, negation can be explained in terms of the following introduction and elimination rules:

I1. Reject A → Assert not-A

I2. Assert A → Reject not-A

E1. Assert not-A → Reject A

E2. Reject not-A → Assert A

In this way, the meaning of negation is explained entirely in terms of its inferential connection to the speech act of rejection. Crucially for bilateralists, assertion and rejection are distinct activities where neither is more fundamental than the other; they are "on all fours with each other." (Smiley, 1996: 1). Introduction and elimination rules can then be given in terms of asserting and rejecting claims involving the logical constants.[9]

To illustrate, consider first:

[8] It also falls under the general characterization of moral inferentialism given at the start of this section.

[9] For details, see Smiley (1996); Rumfitt (2000).

(1) "Tibbles is on the mat."
(2) "Yes." [In answer to the question "Is Tibbles on the mat?"].

In uttering either (1) or (2), one asserts *that Tibbles is on the mat*, where this speech act expresses a belief with the same content. By contrast, consider next:

(3) "Tibbles is not on the mat."
(4) "No." [In answer to the question "Is Tibbles on the mat?"].

In uttering (3), one asserts that *Tibbles is not on the mat*. However, in uttering (4), one does not assert anything; rather, one *rejects* the content *Tibbles is on the mat*, where Incurvati and Schlöder suggest that this speech act expresses a corresponding attitude of *disbelief* toward the same content. So (3) and (4) express distinct attitudes towards distinct contents. Nonetheless, the meaning of (3) is explained entirely in terms of its inferential connection to (4).[10]

Incurvati and Schlöder (2023) then extend this basic idea to explain other kinds of vocabulary, such as epistemic modals, conditionals, probability operators, the truth predicate, and moral terms. Specifically, they develop a *multilateral* approach in which different kinds of vocabulary can be explained in terms of further kinds of speech acts.[11] In relation to the moral domain, Incurvati and Schlöder (2023: 150-6) posit the additional primitive speech acts of *approval* and *disapproval*. These speech acts—realized in ordinary language utterances like "Yay" and "Boo" as in "Charity? Yay!" and "Stealing? Boo!"—express corresponding attitudes of approval and disapproval, much as assertion and rejection express belief and disbelief.[12] We can then explain the meanings of 'right' and 'wrong' in terms of the following elimination rules:

E1(right). *Assert* right(a) → *Approve a*

I1(right). *Approve a* → *Assert* right(a)

E1(wrong). *Assert* wrong(a) → *Disapprove a*

I1(wrong). *Disapprove a* → *Assert* wrong(a)

[10] I'm sceptical that there is a genuinely distinct attitude of disbelief, even assuming that rejection is a distinct speech act (neither Smiley's nor Rumfitt's bilateralism posits any such attitude). However, exploring this would take us too far afield from the main concerns of the paper, so I will do no more than simply register my scepticism (see Price 1990 for a defence of disbelief). Incurvati and Schlöder (2023: 69–70) also explain the speech acts of assertion and rejection in terms of their updating a positive and negative conversational common ground. To keep things simple, I'll focus exclusively on the attitude expression component.

[11] An important part of Incurvati and Schlöder's (2023) view is to distinguish between two kinds of assertion/rejection: strong assertion/rejection and weak assertion/rejection. This is required to explain assertions and rejections that commit one to *refraining* from believing certain contents. To keep things simple, I'll continue to talk as if there is only one kind of assertion/rejection.

[12] They also update corresponding conversational common grounds relating to what is (dis)approved.

In this way, the meanings of moral terms can be explained solely in terms of their inferential connection to the speech acts of approval and disapproval, where this is cashed out in terms of the commitments one undertakes and is entitled to in public discursive practice. This concludes the presentation of our second strand of moral inferentialism.

3 The challenge from moral psychology

3.1 Revisiting the three main benefits

In the previous section, we saw how inferentialism aims to enjoy three central advantages over expressivism. First, it retains all the benefits of non-representationalism, such as not being hostage to ontological and epistemological scepticism about morality. Second, it avoids the Frege-Geach problem by locating moral embedding within a wider, non-representationalist theory of meaning. Third, it avoids any controversial commitments within moral psychology regarding the nature of moral belief, since moral claims are ordinary assertions that express ordinary beliefs. While my primary focus in the remainder of the paper will be to raise a challenge for inferentialism relating to its commitments within moral psychology, it is worth briefly saying something in relation to the first two advantages.[13]

Regarding the first, there is a complication due to the presence of normative explainers within the general inferentialist theory of meaning (e.g. *entitlement*, *authority*, *license*, etc.). To those who are attracted to non-representationalism in virtue of its straightforward compatibility with a naturalistic worldview, any unreduced explanatory appeal to normativity will be problematic. Several responses invite themselves, such as reducing the relevant notions to something naturalistically acceptable, applying our non-representational theory of meaning to itself, or rejecting the naturalistic motivations upon which the objection is premised. While inferentialists have had much to say about this question, I won't pursue the issue any further here. The point is simply that it is *not obvious* that inferentialism retains the same benefits with respect to its non-representationalist commitments as expressivism.[14]

Regarding the second putative advantage, many now accept that expressivists are not committed to a psychologistic semantics for moral terms. Instead, expressivists can adopt a variant of possible worlds semantics in which the extensions of moral terms are fixed relative to an additional, non-factual parameter as well the world.

[13] This is not to deny that there are other motivations for preferring inferentialism to expressivism. For example, Incurvati and Schlöder (2023: 159−66) argue that moral inferentialism provides a better solution the wishful thinking problem and the problem of Moorean moral sentences.

[14] For discussion, see Brandom (1994: 618−50); Chrisman (2016: 204-9); and Tiefensee (2021: 128−30). Of course, some (though only some) expressivist theories face the same problem insofar as they think that the relevant attitudes expressed by moral claims must be partly characterized in normative terms (e.g., Gibbard, 2012). It seems to me, however, that the simplest response for the expressivist is just to deny that moral attitudes must be characterized in this way. Perhaps propositional attitudes are essentially rational attitudes, but the relevant sense of rationality should not be understood as robustly normative (see Ridge, 2014). Insofar as inferentialists typically think that talk of entitlements and the like is irreducibly normative, playing an expressive rather than descriptive role, it is unclear whether this option is open to them.

For example, we might understand this parameter in terms of *hyperplans*, *To-Do Lists*, *orderings*, *standards*, and so on.[15] Given that possible worlds semantics already posits additional parameters within its models to account for other kinds of linguistic expressions (e.g. de se expressions, imperatives), this sort of move need not be especially ad hoc or controversial.[16] While any such approach owes us an adequate account of the semantics of moral terms within this framework, it allows the expressivist to explain embedding and compositionality in standard model-theoretic ways. This does not obviate the need for an account of moral attitudes, one that is likely to still play a central role in the expressivist's meta-semantics. And no doubt such an account will still be controversial. But we are a long way off from the classic problems of psychologistic expressivist semantics that inferentialists often cite as motivation for preferring their view.[17] In any case, we will return to the Frege-Geach problem below, so I will not pursue this line of thought any further.

Regarding the third putative benefit, we can grant that since moral inferentialism eschews any kind of psychologistic semantics or meta-semantics, its theory of meaning does not require positing any particular moral psychology. However, this does not discharge inferentialism from the burden of explaining moral thought. Thinking moral thoughts, making moral judgments, engaging in moral deliberation, and so on, are all central aspects of moral practice—aspects, moreover, that we should expect any metaethical theory to explain. So moral inferentialists still owe us an account of our moral psychology.[18] We see, however, that what inferentialists say about moral thought is unsatisfactory. While more plausible accounts of moral thought are available to inferentialists, these lead us back to expressivism about moral thought. The remainder of the paper is dedicated to arguing for these claims.

3.2 Moral psychology for inferentialists

We have seen that moral inferentialism explains the meaning of moral expressions in terms of the commitments we make in using moral expressions. We acknowledge and undertake these commitments by asserting moral claims. And such assertions, we are told, express moral beliefs. So the first answer to our question, 'what is it to think a moral thought?' is that it consists in having a belief with a moral content. But what is it to have a belief with a moral content? In what follows, I examine different ways in which the inferentialist might answer this question, beginning with the least committal answers and building up to more constructive ones.

The first answer is to appeal to minimalism or deflationism about belief. According to minimalism, a mental state is a belief just in case it is expressible by sincerely

[15] Yalcin, 2012a); Charlow (2014a); Silk (2015); Ridge (2014) respectively. While Gibbard (1990, 2003) is often the starting point for such views, the crucial point is that while Gibbard provides a psychologistic interpretation of his formal apparatus, this is not mandatory; see Yalcin (2018, 2022) for discussion.

[16] On de se expressions, see Liao (2012); on imperatives, see Charlow (2014b).

[17] For instance, Incurvati and Schlöder (2023: 175-6) reject Gibbard's hyperplan semantics largely due to problems that arise from Gibbard's psychologistic interpretation of the semantics. Many subsequent theorists who employ Gibbard's semantics, however, argue that we should adopt a non-psychologistic interpretation; see, e.g., Yalcin (2012b), 2018), 2022).

[18] See Chrisman (2023: 63) for an articulation of this challenge.

asserting a sentence that is syntactically disciplined (Sinclair, 2006: 253).[19] Since whatever mental states are expressed by moral claims clearly meet these criteria, we can maintain that moral claims express moral beliefs. The problem, however, is that minimalism is not an account of what any mental state consists in; rather, it is an account of what conditions must be met in order for a mental state to count as a minimal belief.[20] This is brought out by the fact that most contemporary expressivists think that moral claims express minimal beliefs. Yet they deny that minimalism provides an account of what moral belief consists in. So minimalism does not provide the right sort of answer to our question.

A second answer is to observe that inferentialism is compatible with many robust conceptions of belief. For instance, Chrisman (2023: 63) endorses an interpretationist account of belief, according to which one believes a moral content just in case this best rationalizes one's behaviour. Further, Incurvati and Schlöder (2023: 189) suggest that moral inferentialists might adopt Fodor's representational theory of belief, according to which one believes a moral content just in case one has a concrete representation of a moral sentence stored in one's 'belief box'. And similarly for many other substantive conceptions of belief.

By itself, however, this kind of compatibility is not sufficient to explain moral belief. For it is a core feature of such theories that they conceive of belief as a representational state. It is because of this that representationalist metaethical theories can simply help themselves to a general notion of belief and with it their preferred substantive theory and then explain moral beliefs specifically in terms of what those beliefs represent. By contrast, insofar as inferentialism denies that beliefs in general or moral beliefs in particular are representational, it cannot explain moral belief in the same way. So really, inferentialism's compatibility with theories like interpretationism and the representational theory of the mind is orthogonal to the issue under discussion, since rejecting the representationalist picture implicit within such theories leaves us without an account of how to individuate moral beliefs within those theories. What we need is an account of the essential or characteristic dispositional profile of moral belief that explains why moral beliefs realize the belief-role in our cognitive economy without thereby representing moral reality.

Given that the contents of moral beliefs are identified with the contents of the moral claims that express them, and given that the contents of moral claims are explained in terms of acknowledging and undertaking commitments in a public discursive practice, a third answer is to say that moral belief consists in being disposed to sincerely assert its content in the appropriate circumstances.[21] This might be understood as part of a more general non-representationalist account that explains belief in general this way. The problem, however, is that being disposed to sincerely assert p (and thus acknowledge and undertake the commitment to p) is neither necessary nor sufficient for believing p. This point is quite general and not specific to moral belief, as

[19] See also Incurvati and Schlöder (2023: 185).

[20] See Brown (2022: 14).

[21] See Incurvati and Schlöder (2023: 185-6) and Chrisman (2023: 63). Chrisman also suggests here that we might explain moral belief by examining the "psychological underpinnings" of such dispositions.

is highlighted by the following kinds of example given by Marcus (1990).[22] First, consider cases in which our actions belie our sincerely reported beliefs. While this might be due to conflicting beliefs or extraneous features of the situation, sometimes we are simply wrong about what we believe—whether this is due to self-deception, false consciousness, unthinkingly repeating what we have heard elsewhere, or simply being mistaken. If this is possible, then being disposed to sincerely assert *p* cannot be sufficient for believing *p*, since we might have the disposition without the corresponding belief. Second, consider cases in which having been asked why we acted as we did, we must reflect to discover what we believe before reporting our belief. This might happen on the therapist's chair, but this also happens in day-to-day life. If this is possible, being disposed to sincerely assert *p* cannot be necessary for believing *p* either, since we might lack this disposition before our discovery.[23]

An important feature of these examples is that the relevant beliefs manifest in ways other than in what the individual *says*; rather, they manifest in what the individual *does*. This raises the independently plausible thought that the dispositional profile of belief is defined in terms of its role in the explanation of action.[24] In light of this, a fourth answer might therefore appeal to the role of moral belief in action explanation. Clearly, this role cannot be the kind of representational role possessed by ordinary factual beliefs. So we need to identify some kind of non-representational role that moral belief plays in action explanation. The most obvious candidate is that moral beliefs have a distinctive non-representational role in action explanation in virtue of the fact that moral beliefs can directly motivate actions. This is in contrast to non-moral beliefs, which can only motivate in conjunction with desires.

While a good place to start, motivational internalism is too simplistic a model to explain the non-representational role of moral thought in general. For there are many moral beliefs do not motivate action in accordance with the simple internalist model. These include beliefs about what we should feel and believe, beliefs about what others should do, and logically complex moral beliefs. However, while I lack the space to fully develop this point here, it is nonetheless plausible that such beliefs play some kind of directive role, broadly conceived, in our cognitive economy. For instance, that an agent feels guilty in some circumstance might be explained by the fact that they believe that they have done something morally wrong. Moreover, that an agent feels resentful or angry might be explained by the fact that they believe that someone else has done something morally wrong. Further, that an agent chooses to stop consuming dairy products might be (partly) explained by the fact that they believe that if it is wrong to kill animals for food, it is wrong to support industries that rely on killing animals for food. Thus, in each case, moral beliefs play a broadly directive role

[22] See Dennett (1978) for similar examples. This leads Dennett to distinguish between *beliefs*, which are involved in the explanation of action, and *opinions*, which involve assenting to certain sentences. See also Dennett (1981: 19n, 112−14).

[23] Note that this is compatible with the claim that language is necessary for moral thought, since the latter is a claim about the enabling conditions for moral thought in general, not what it is to have some particular moral belief or other.

[24] This is accepted by a wide variety of theories of belief. For example, it is central to Dennett's (1981) interpretationism, Stalnaker (1984) and Lewis' (1994) analytic functionalism, Fodor's (1975) representational theory of mind, and Marcus' (1990) pragmatic dispositionalism, to name but a few.

within the agent's cognitive economy in governing and regulating the agent's actions and attitudes, even if this role does not fully align with the simple internalist model.

To develop these observations into a mature theory of moral belief, we would need to provide a constructive and precise account of the directive role of such beliefs in our cognitive economy. However, that there is some account of the directive role of belief is made plausible by theory-neutral observations about the explanatory role of moral belief in ordinary folk-psychological explanation. Moreover, adopting such an account allows us to accommodate the kinds of cases that linguistic accounts of belief failed to explain. For it seems particularly commonplace for people's actions and feelings to belie their professed moral beliefs. If this is right, then we need something like the above account to explain the role of moral belief in our cognitive economy, at least if we are assuming it lacks a representational role.

Suppose, then, that the inferentialist accepts this account of moral belief. While this is perfectly consistent, the problem is that we are now positing exactly the same kind of account of moral belief as that given by expressivists. Indeed, at the fundamental level, to provide an account of the non-representational, directive role of moral belief in this way just is to provide an expressivist account of moral belief. This is not to say that the overall package of views will be identical. Since inferentialism is not in the business of providing psychologistic explanations of moral practice, it will not assign the same explanatory role to its moral psychology that expressivism assigns to its moral psychology. Instead, the inferentialist will want to tell some complicated story about how moral belief somehow derives its features from moral discursive practice—though we have seen that any such story cannot be given purely in terms of our dispositions to undertake and acknowledge moral commitments via moral assertion, since this is neither necessary nor sufficient for moral belief. However, even if the explanatory role assigned to moral belief differs between inferentialism and expressivism, it remains the case on the current proposal that the account of moral belief *itself* is the same for the two approaches. Thus, if the inferentialist needs to adopt the kind of account of moral belief sketched above, then it can no longer be maintained as an advantage of inferentialism that it avoids the kind of sophisticated moral psychology posited by expressivism. In this way, the most plausible way of answering the challenge robs inferentialism of its supposed advantages in relation to moral psychology.

In the remainder of this section, I examine three objections to the argument so far. The first is that the argument assumes something inferentialists already reject, namely the priority of thought. The second is that there remain other important reasons for preferring inferentialism, namely its ability to avoid the Frege-Geach problem. The third is that the argument ignores the possibility of alternative accounts of moral belief not yet considered. As well as clarifying the preceding argument, responding to these objections will also allow us to strengthen the case against inferentialism as a competitor to expressivism.

3.3 Objections and responses

3.3.1 The priority of thought

To begin, one might worry that the challenge from moral psychology presupposes something that inferentialists already reject, namely the priority of thought. Many inferentialists maintain on independent grounds that thought is to be explained in terms of language (e.g., Sellars, 1974b), or that neither thought nor language enjoys explanatory priority over the other (e.g., Brandom, 1994: 150-1, 2000: 5–7). By contrast, since expressivists aim to explain moral discourse in terms of what it is to think moral thoughts, this approach seems to be committed to priority of thought over language more generally. Thus, if the challenge from moral psychology presupposes the priority of thought, then this unfairly stacks the deck in favour of expressivism.

However, the challenge from moral psychology does not presuppose the priority of thought. In developing the challenge, it was argued that a simple linguistic account of moral belief is implausible. Further, it was argued that any plausible account must explain how moral belief manifests in what the agent does, not only in what the agent says. However, these claims were motivated not from a prior commitment to the priority of thought, but rather from plausible, independent observations about moral belief and belief more generally. Further, it was not argued that the resulting picture of moral belief is *incompatible* with inferentialism or with rejecting the priority of thought. Rather, the argument, if successful, simply shows that inferentialists do not enjoy a simpler or less controversial moral psychology than expressivists. Indeed, the views end up looking remarkably similar.

Perhaps at this point the inferentialist might cite their prior rejection of the priority of thought as a reason to prefer their account of moral belief. After all, inferentialists take themselves to have good independent reasons to reject this thesis. Accordingly, they might cite this feature of their moral psychology as an advantage of their view. However, this is now a very different argument from the original argument one that inferentialists can avoid the kind of moral psychology characteristic of expressivism, since the new argument accepts this moral psychology. Moreover, expressivists can equally argue that their view is preferable by citing independent arguments for the priority of thought. In the absence of adjudicating this more general debate, then, this response is therefore dialectically ineffective—though perhaps at the end of the day this is where the debate might have to be decided.

3.3.2 Frege-Geach revisited

The challenge from moral psychology aimed to show that inferentialism does not enjoy an obvious advantage over expressivism regarding its commitments within moral psychology. Granting this, however, if inferentialists have other reasons to prefer their view, then this might not be a huge cost to inferentialism all things considered. That said, given that inferentialists do commonly claim that inferentialism avoids an expressivistic moral psychology, it is still dialectically significant if this is false. Moreover, even if inferentialism is overall the preferable view, it still tells us something interesting about inferentialism that a fully worked-out version of the

view would be in substance much closer to expressivism that has been generally recognized.

Be that as it may, I now want to make a further and stronger claim regarding the implications of the challenge from moral psychology. For arguably, the account of moral belief on offer to inferentialists can be appropriated by the expressivist to solve the Frege-Geach problem. To see how, let us consider again the picture of belief we ended up with at the end of the last section. According to this picture, moral belief (a) has a broadly directive role in the agent's cognitive economy, (b) stands in inference-like relations of consistency, entailment, and so on with other beliefs, and (c) is non-representational. We observed that this account was available both to the inferentialist and the expressivist. Their disagreement will concern the explanatory priority of moral belief in relation to discursive practice. Thus, whereas the inferentialist will claim that features (a)-(c) are (partly) explained in terms of moral assertion, the expressivist will claim that moral assertion takes the form that it does because it expresses mental states with features (a)-(c).

However, to provide an account of moral belief that explains moral assertion in this way *just is* to solve the Frege-Geach problem. Moreover, observe that we know that our account of moral belief will have the right sort of structure, since we are supposing that the very same account is available to the inferentialist. In both cases, we have an account of moral thought that is isomorphic with moral language.[25] The difference concerns whether these features of moral belief are explained as (inter-) dependent on moral language or as basic. And since there is nothing essentially inferentialist about the account of belief, it is open to non-representationalists to take either path.

To be clear, I have not here done the work of constructing a positive account of moral belief with features (a)-(c) that would solve the Frege-Geach problem for expressivism. Equally, I have not done the work of constructing a positive account of moral belief with features (a)-(c) that would solve the challenge from moral psychology for inferentialism. The dialectical point is that if the inferentialist can answer the challenge from moral psychology by positing an account of moral belief with features (a)-(c), then the expressivist is equally entitled to posit such an account within their moral psychology, which in turn will provide the resources to answer the Frege-Geach problem.[26]

It might be responded that whereas the inferentialist is entitled to (c) due to their inferentialist account of content, the expressivist is not entitled to (c) and so is not equally entitled to the account of moral belief. While I lack the space to do justice to this issue, there are several things to say in response. First, as was noted above, there are several non-factualist notions of moral content available to the expressivist, and some expressivists have argued that such contents can play an explanatory role in their theory (e.g., Brown, 2022; Schroeder, 2013; Yalcin, 2022). Second, some expressivists have argued that a satisfactory account of moral belief with features (a)-(c) can be constructed using only a deflationary notion of moral content (e.g., Köhler, 2017). Third, it is widely held across different accounts of mental content that a belief's con-

[25] Or at least homomorphic if not isomorphic—see Matthews (2007).

[26] See Köhler (2025) for a similar way of developing expressivism along these lines.

tent is determined by its dispositional profile, which is explanatorily prior. In other words, a belief has the content that it has in virtue of its dispositional profile. So for a thought-first theory like expressivism, there is a sense in which assigning belief content is explanatorily downstream from the question of the fundamental dispositional profile of belief. So assuming that our account of the dispositional profile of belief has the right sort of structure, this should be sufficient to determine a suitable content for that belief, however exactly we understand that content.

3.3.3 Other options

To recap, the challenge from moral psychology argues that inferentialists owe us an account of what it is to think moral thoughts. It was then argued that the most promising non-representationalist account available ends up looking a lot like expressivism. As such, inferentialism does not seem to enjoy any obvious advantages with respect to its account of moral thought compared with expressivism. Supposing this is correct, one might worry that this argument fails to rule out the possibility of developing some alternative account of belief compatible with inferentialism that is also distinctively non-expressivistic. Insofar as this remains a possibility, the inferentialist can reject the claim that they must develop an expressivistic moral psychology.

While nothing in the above argument rules out this possibility, I hope it is clear by now that the burden of the proof lies with the inferentialist to provide an account of moral belief, or at least to explain how they might go about providing one. This is why moral psychology provides a *challenge* for inferentialists. Moreover, the inferentialist must also explain why any such account is not also available to the expressivist as well as why their account is less controversial than the expressivist's. For if it fails to do these things, then the inferentialist cannot claim to have the advantage with respect to its account of moral thought.

Perhaps one reason for optimism regarding the availability of such an account derives from a worry about the expressivist's moral psychology. Specifically, the worry might be that it posits an implausible bifurcation of mental states into the descriptive and directive. While this might be plausible when comparing moral and ordinary factual beliefs, it is implausible when we observe that there are many other domains besides morality that we might wish to be non-representationalists about. For instance, this might include epistemic modality, alethic modality, mathematical discourse, personal taste, truth discourse, and perhaps other 'flavours' of normativity besides morality. Since it is implausible to think that all these domains can be explained as directive, there is reason to doubt that an expressivistic theory of thought will be able to capture these other domains.[27]

While we might respond by simply denying that these other domains are in fact non-representational, it would be nice if our non-representationalist framework was at least in principle generalizable to other domains. Fortunately, however, the expressivist is not committed to the bifurcated picture. Instead, they can embrace a kind of *functional pluralism* about thought, according to which different kinds of thought can be individuated in terms of their distinctive functions. On this view, representing the

[27] Chrisman (2016: 218) raises a worry along these lines.

world and governing our actions and attitudes are just two kinds of function among many (compare Price, 2013). In this spirit, we see that there are distinctively expressivist approaches to a wide variety of discourses, such as epistemic modality (Yalcin, 2011), alethic modality (Blackburn, 1993), truth discourse (Schroeder, 2010), personal taste (Ninan, 2024), mathematical thought (Peréz Carballo 2014), know-how (Santorio, 2016), and prudential discourse (Brown, 2025), to name but a few.

Finally, even if one is sceptical about this approach, this does not in any way obviate the need for inferentialism to provide a constructive moral psychology—one, moreover, that the expressivist cannot adopt. Not only that, but it must be shown to be less controversial than the expressivist's if it is to vindicate the claim that inferentialism is preferable to expressivism in virtue of its commitments within moral psychology. So even if the prospects for expressivism look dim, this does not thereby mean that the prospects for inferentialism look any better. We still lack an argument that non-representationalists should be inferentialists in virtue of their commitments within moral psychology.

4 Conclusion

To conclude, it is worth emphasizing what the preceding argument does and does not show. Importantly, we need not take the argument to show that moral inferentialism should be rejected outright. Rather, the main target of the argument is the claim that inferentialism provides a more attractive non-representationalist metaethics than expressivism. This is compatible with the thought that the best non-representationalist metaethics will incorporate aspects of both. Indeed, many versions of expressivism already come quite close to inferentialist ideas. For example, some expressivists stress the centrality of conceptual role—understood as the mental analog of inferential role—for understanding moral thought.[28] Other expressivists stress the importance of commitments—though understood as a kind of mental state rather than public status—to noncognitive attitudes as being central to moral thought.[29] Still others argue that moral claims express non-representational beliefs, in a non-deflationary sense.[30] And still others stress an essential discursive aspect of the functional role of moral thought.[31] All of these ideas are very similar in spirit to ideas found in inferentialism.

Given these similarities, one might worry that I have loaded the dice against the inferentialist. If expressivists have already moved closer to inferentialists, then is it any surprise that inferentialists end up saying similar things to expressivists? Again, it is important to keep in mind that the main target has been the idea of inferentialism as an alternative to expressivism. Indeed, I suspect the most plausible version of non-representationalism will incorporate aspects of both views. If this is correct, perhaps the question of whether non-representationalists about morality should be inferentialists or expressivists will hinge on broader questions about the background

[28] See, e.g., Horwich (2010); Köhler (2017, 2025); Sinclair (2018).

[29] See, e.g., Sinclair (2021); Baker and Woods (2022); Lenman (2024).

[30] See, e.g., Horgan and Timmons (2006); Schroeder (2013); Brown (2022).

[31] See, e.g., Gibbard (1990).

assumptions of each theory, such as whether he have general reasons to accept or reject the priority of thought.

That said, let me offer this final observation on our debate. Suppose that the expressivist and the inferentialist accept the same moral psychology. The expressivist uses this account to explain the meanings of moral sentences and the contents of moral beliefs. By contrast, while the inferentialist posits the same account of moral thought, they explain the meanings of moral sentences and the contents of moral beliefs (partly) in terms of an additional system of permissions and entitlements realized in public discursive practice. However, if the account of moral psychology is sufficient to explain moral meaning for the expressivist, then why isn't the same account also sufficient for the inferentialist? If the inferentialist must posit this account anyway, then why the need for the additional system of permissions and entitlements?

I have argued that in order to provide an adequate account of moral thought, inferentialists need to say the same sorts of things that expressivists say about moral thought. But if inferentialists say this, then we already have the resources in hand to provide a psychologistic explanation of the relevant phenomena. So quite apart from any general arguments about the explanatory priority or lack thereof of thought over language, we arguably have reason to favour a thought-first approach to explaining moral discourse. This is all compatible with the inferentialist's description of public linguistic practice in terms of acknowledging and undertaking commitments. And perhaps this will be an important part of the overall picture. But the central argument of this paper suggests that it will not play a fundamental role in explaining moral meaning or moral discourse.

Acknowledgements Many thanks to Sebastian Köhler, Christine Tiefensee, and two anonymous reviewers for *Synthese* for comments and feedback on previous versions of this paper. Thanks also to an audience at Universität Graz. This work was supported by funding from the British Academy (PF21\210089) and the Alexander von Humboldt Foundation.

Funding British Academy (grant no. PF21\210089) and Alexander von Humboldt Foundation.

Declarations

Competing interests The author has no competing interests to declare that are relevant to the content of this article. There are no other relevant ethical standards to declare compliance with.

References

Baker, D. (2020). Quasirealism as semantic dispensability. *Philosophical Studies*, *178*(7), 2313–2333.

Baker, D., & Woods, J. (2022). Handling rejection. *Philosophical Studies*, *180*(1), 159–190.

Blackburn, S. (1993). Morals and modals. In *Essays in quasi-realism*. Oxford University Press.

Block, N. (1986). Advertisement for a semantics for psychology. *Midwest Studies in Philosophy*, *10*(1), 615–678.

Brandom, R. (1994). *Making it explicit*. Harvard University Press.

Brandom, R. (2000). *Articulating reasons: An introduction to inferentialism*. Harvard University Press.

Brandom, R. (2008). *Between saying and doing: Towards an analytic pragmatism*. Oxford University Press.

Brown, J. L. D. (2022). Interpretative expressivism: A theory of normative belief. *Philosophical Studies*, *179*(1), 1–20.

Brown, J. L. D. (2025). How to be a prudential expressivist. *Mind*, *134*(534), 324–346.

Chalmers, D. (2021). Inferentialism, Australian style. *Proceedings and Addresses of the American Philosophical Association*, 92.

Charlow, N. (2014a). The problem with the Frege–Geach problem. *Philosophical Studies*, *167*(3), 635–665.

Charlow, N. (2014b). The meaning of imperatives. *Philosophy Compass*, *9*(8), 540–555.

Chrisman, M. (2008). Expressivism, inferentialism, and saving the debate. *Philosophy and Phenomenological Research*, *77*(2), 334–358.

Chrisman, M. (2010). Expressivism, inferentialism, and the theory of meaning. In M. Brady (Ed.), *New waves in metaethics*. Palgrave-Macmillan.

Chrisman, M. (2012). On the meaning of 'ought'. In *Oxford studies in metaethics volume 7*. Oxford University Press.

Chrisman, M. (2013). Attitudinal expressivism and logical pragmatism. In G. Hubbs & D. Lind (Eds.), *Pragmatism, law, and language*. Routledge.

Chrisman, M. (2016). *The meaning of 'ought': Beyond descriptivism and expressivism in metaethics*. Oxford University Press.

Chrisman, M. (2018). Two nondescriptivist views of normative and evaluative statements. *Canadian Journal of Philosophy*, *48*(3–4), 405–424.

Chrisman, M. (2023). Inferentialism as an alternative to expressivism. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics volume 18*. Oxford University Press.

Dennett, D. C. (1978). How to change your mind. *Brainstorms*. MIT Press.

Dennett, D. C. (1981). *The intentional stance*. MIT Press.

Fodor, J. A. (1975). *The language of thought*. Harvard University Press.

Gibbard, A. (1990). *Wise choices, Apt feelings: A theory of normative judgment*. Harvard University Press.

Gibbard, A. (2003). *Thinking how to live*. Harvard University Press.

Gibbard, A. (2012). *Meaning and normativity*. Oxford University Press.

Horgan, T., & Timmons, M. (2006). Cognitivist expressivism. In T. Horgan & M. Timmons (Eds.), *Metaethics after Moore.* Clarendon.

Horwich, P. (2010). The motive power of evaluative concepts. *Truth-Meaning-Reality*. Clarendon.

Incurvati, L., & Schlöder, J. J. (2021). Inferential expressivism and the negation problem. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (Vol. 16). Oxford University Press.

Incurvati, L., & Schlöder, J. J. (2023). *Reasoning with attitude: Foundations and applications of inferential expressivism*. Oxford University Press.

Karimi, P. (2021). Is relaxed realism a genuinely novel view? *Journal of Philosophy*, *118*(10), 572–584.

Köhler, S. (2017). Expressivism, belief, and all that. *Journal of Philosophy*, *114*(4), 189–207.

Köhler, S. (2025). Expressivism, but at a whole other level. *Erkenntnis*, *90*(1), 367–388.

Lenman, J. (2024). *The possibility of moral community*. Oxford University Press.

Lewis, D. K. (1994). Reduction of mind. In S. D. Guttenplan (Ed.), *A companion to the philosophy of mind*. Blackwell.

Liao, S. (2012). What are centered worlds? *Philosophical Quarterly*, *62*(247), 294–316.

Marcus, R. B. (1990). Some revisionary proposals about belief and believing. *Philosophy and Phenomenological Research*, *50*, 133–153.

Matthews, R. (2007). *The measurement of mind: Propositional attitudes and their attribution*. Clarendon.

Ninan, D. (2024). An expressivist theory of taste predicates. *Philosophers' Imprint*, *24*(1).

Peacocke, C. (1992). *A study of concepts*. MIT Press.

Pérez Carballo, A. (2014). Structuring logical space. *Philosophy and Phenomenological Research*, *92*(2), 460–491.
Price, H. (2011). Expressivism for two voices. In J. Knowles & H. Rydenfelt (Eds.), *Pragmatism, science and naturalism* (pp. 87–113). Peter Lang.
Price, H. (2013). *Expressivism, pragmatism and representationalism*. Cambridge University Press.
Ridge, M. (2014). *Impassioned belief*. Oxford University Press.
Rumfitt, I. (2000). Yes and no. *Mind*, *109*(436), 781–823.
Santorio, P. (2016). Nonfactual know-how and the boundaries of semantics. *Philosophical Review*, *125*(1), 35–82.
Schroeder, M. (2008). What is the Frege-Geach problem? *Philosophy Compass*, *3*(4), 703–720.
Schroeder, M. (2010). How to be an expressivist about truth. In C. Wright & N. J. L. L. Pedersen (Eds.), *New waves in truth* (pp. 282–298). Palgrave-Macmillan.
Schroeder, M. (2013). Two roles for propositions: Cause for divorce? *Noûs*, *47*(3):409–430.
Sellars, W. (1954). Some reflections on language games. *Philosophy of Science*, *21*(3), 204–228.
Sellars, W. (1974a). Meaning as functional classification. *Synthese*, *27*(3–4), 417–437.
Sellars, W. (1974b). Language as thought and communication. In *Essays in philosophy and its history*. D Reidel Publishing Company.
Silk, A. (2015). How to be an ethical expressivist. *Philosophy and Phenomenological Research*, *91*(1), 47–81.
Sinclair, N. (2006). The moral belief problem. *Ratio*, *19*(2), 249–260.
Sinclair, N. (2018). Conceptual role semantics and the reference of moral concepts. *European Journal of Philosophy*, *26*(1), 95–121.
Sinclair, N. (2021). *Practical expressivism*. Oxford University Press.
Smiley, T. (1996). Rejection. *Analysis*, *56*(1), 1–9.
Stalnaker, R. (1984). *Inquiry*. Cambridge University Press.
Tiefensee, C. (2016). Inferentialist metaethics, bifurcations and ontological commitment. *Philosophical Studies*, *173*(9), 2437–2459.
Tiefensee, C. (2019). Saving which differences? Creeping minimalism and disagreement. *Philosophical Studies*, *176*(7), 1905–1921.
Tiefensee, C. (2021). Metasemantics for the relaxed. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics volume 16*. Oxford University Press.
Tiefensee, C. (2023). Through thick and thin: Seamless metaconceptualism. *Synthese*, *201*(2), 1–19.
Warren, M. D. (2015). Moral inferentialism and the Frege-Geach problem. *Philosophical Studies*, *172*(11), 2859–2885.
Wedgwood, R. (2001). Conceptual role semantics for moral terms. *Philosophical Review*, *110*(1), 1–30.
Wedgwood, R. (2007). *The nature of normativity*. Clarendon.
Woods, J. (2017). The Frege-Geach problem. In T. McPherson & D. Plunkett (Eds.), *The Routledge handbook of metaethics*. Taylor and Francis.
Yalcin, S. (2011). Nonfactualism about epistemic modality. In A. Egan & B. Weatherson (Eds.), *Epistemic modality*. Oxford University Press.
Yalcin, S. (2012a). Bayesian expressivism. *Proceedings of the Aristotelian Society*, *112*(2pt2), 123–160.
Yalcin, S. (2012b). Comments on Schroeder, *Being for*. Read at an author-meets-critics session, Central APA.
Yalcin, S. (2018). Expressivism by force. In D. Fogal, D. W. Harris, & M. Moss (Eds.), *New work on speech acts*. Oxford University Press.
Yalcin, S. (2022). Modeling with hyperplans. In B. Dunaway & D. Plunkett (Eds.), *Meaning, decision, and norms: Themes from the work of Allan Gibbard*. Maize Books.