



Expressivism, but at a Whole Other Level

Sebastian Köhler¹ 

Received: 23 June 2022 / Accepted: 19 May 2023
© The Author(s) 2023

Abstract

A core commitment of meta-ethical expressivism is that ordinary descriptive judgements are representational states, while normative judgements are non-representational directive states. Traditionally, this commitment has been understood as a psychological thesis about the nature of normative judgements, as the view that normative judgements consist in certain sorts of conative propositional attitudes. This paper's aim is to challenge this reading and to show that changing our view on how this commitment is to be understood opens up space for attractive forms of expressivism. The paper argues this by example: by presenting one of the possible expressivist views that can be developed on this reading and by arguing that this view has distinctive advantages over forms of expressivism that endorse the traditional interpretation of expressivism.

1 Introduction

Meta-ethical expressivism is characterized by two commitments, one in the philosophy of language and one in the philosophy of mind. The first is that the meaning of sentences is to be explained in terms of the judgements their uses express. The second is that there is a distinctive difference between ordinary descriptive and normative judgements: while ordinary descriptive judgements are *representational* states, normative judgements are *non-representational directive* states.

The focus of this paper will be on expressivism's *second* commitment. On the *traditional* reading, this commitment has been understood as the view that normative judgements consist in certain sorts of conative propositional attitudes, such as plans, intentions, desires, or emotions (see e.g., Blackburn, 1998, Gibbard, 1992, 2003, Ridge, 2014, or Schroeder, 2008). Such views cash this commitment out in what we might call “first order” psychological terms, by appealing to certain

✉ Sebastian Köhler
s.koehler@fs.de

¹ Department of Computational Science & Philosophy, Frankfurt School of Finance & Management, Frankfurt, Germany

kinds of conative propositional attitudes with certain kinds of contents to explain the nature of normative judgements. Simon Blackburn, for example, claims (1998: 67/68):

“[I]f we imagine the general field of an agent’s concerns, his or her values might be regarded as those concerns that he is also concerned to preserve: the ones by which he stands. [T]o hold a value is to have a relatively fixed attitude to some aspect of things, an attitude with which one identifies in the sense of being set to resist change, or set to feel pain when concerns are not met.”

And, Mark Schroeder (2008: 58/59) suggests:

“[T]hey will have to say that [normative judgement] is some kind of very general *noncognitive* attitude. Let’s give it a name, and call it *being for*. [T]he assumption that being for is inconsistency-transmitting is simply the assumption that it has a property that is familiar from other sorts of mental state, including belief and intention.”

These authors, clearly, offer an interpretation of expressivism’s second commitment on which it is cashed out in first-order psychological terms, by appealing to certain conative propositional attitudes, their contents, and structural and phenomenological features. This is the case even for those who hold that expressivists will offer an account that does not proceed from familiar pre-theoretical conative propositional attitudes, but via specifically introduced technical attitudes.

Let us call views that accept the traditional reading of expressivism’s *second* commitment, forms of “traditional expressivism.” This paper’s aim is to challenge traditional expressivism and to show that changing our view on how to understand expressivism’s second commitment opens space for novel and attractive forms of expressivism.

More concretely, the paper’s aims are as follows: First, to argue that there is an alternative reading of expressivism’s second commitment as a commitment in what we might call the “foundational theory of mind.” Second, to argue that this reading is feasible and attractive, by presenting one of the expressivist views that can be developed on this reading—a view I call “conceptual role expressivism”¹—and by arguing that this view has distinctive advantages over traditional expressivism. Note that the main point of the paper is not to argue for, flesh out, or defend conceptual role expressivism—though the paper also pushes that program further. Rather, the paper’s main point is to support a systematic and programmatic general point by drawing out the benefits of expressivist approaches that discard the traditional reading for an alternative way of fleshing out expressivism’s commitments using the resources provided by the foundational theory of mind.

The paper proceeds as follows: Sect. 2 introduces the traditional and my alternative reading of expressivism’s second commitment in more detail. Section 3 argues that the alternative reading is, indeed, viable and attractive, because it opens space for attractive expressivist views. It does so via proof of concept, namely by

¹ I’ve also suggested or further developed this view elsewhere. See Köhler (2017, 2018, and forthcoming).

presenting one way how we might develop an expressivist view so understood, *conceptual role expressivism*. Sections 4 and 5 argue that conceptual role expressivism has distinctive advantages over forms of traditional expressivism. Section 4 argues that conceptual role expressivism offers a more plausible version of judgement internalism and a more attractive picture of what unifies normative beliefs than traditional expressivism. Section 5 argues that conceptual role expressivism has distinctive advantages when it comes to dealing with (a variant of) the Frege-Geach Problem.

2 Expressivism: Toward the Foundational Theory of Mind

Expressivism's second commitment is a commitment about the nature of normative judgements. It is the commitment that these judgements are not representational states, but instead are directive states of some sort. How should we understand this commitment, though?

Above I've said that *traditional* expressivism understands it in first-order psychological terms. Here is what I mean by this: these accounts take a certain set of folk-psychological conative states, such as desires, intentions, plans, or emotions as given. They then try to account for the nature of normative judgements by either identifying normative judgements with (certain combinations of) such states or by introducing a new kind of conative state whose features can be explained by using these kinds of states as a model. So, these accounts aim to account for the nature of normative judgements, only by appealing to the features of certain *other* kinds of propositional (or otherwise contentful; in what follows I only focus on propositional attitudes for simplicity's sake) attitudes and their contents—specifically, by using paradigm kinds of non-representational, directive states from folk-psychology as a model.

It should be clear that this explanatory approach is *restricted* in certain ways: to shoulder explanatory burdens, it only has those tools available that these relevant kinds of states offer, such as their structure, phenomenology, rational relations, and so on. However, there is a different way how we can account for the nature of psychological states, such as normative judgements. Specifically, we can do this at the level of the foundational theory of mind. As I understand it, the *foundational theory of mind* is the philosophical enterprise of giving an account of the foundations of psychology. In the foundational theory of mind, we also give an account of the psychological attitudes. However, our aim is to explain *in virtue of what* certain kinds of psychological states are the states they are or in virtue of what they have their contents. So, for example, it will explain in virtue of what *beliefs* are the states they are and in virtue of what they have their contents—and similarly for the other propositional attitudes, of course. Examples of views in the foundational theory of mind are e.g., functionalism (e.g., Fodor, 1968, Lewis, 1966, or Putnam, 1967), teleofunctionalism (e.g., Dretske, 1988, Millikan, 1984, or Papineau, 1984), or interpretationalism (e.g., Davidson, 1973, Dennett, 1987, or Lewis, 1974). These explain *in virtue of what* a particular state is a belief, desire, etc. and *in virtue of what* it has its content.

My suggestion is that expressivism's second commitment can and should be understood as a commitment in the foundational theory of mind, namely as a view in virtue of what certain sorts of mental states—*normative beliefs*—are *beliefs* and have their specific *contents*. This possibility opens and offers significant conceptual space for unexplored expressivist views.² For example, one could develop views that account for normative judgement solely in terms of the biological *function* of beliefs in general and of normative beliefs in particular. Or one could develop views, as I will do below, that account for normative judgement solely in terms of the *causal-functional role* of beliefs in general and normative beliefs in particular.

Of course, in one sense traditional expressivism is also a view in virtue of what normative beliefs are beliefs and have their contents. Almost all contemporary expressivists are *quasi-realists* and hold that their theory explains why normative judgements are beliefs and have the contents they do. However, these explanations proceed by identifying normative beliefs with suitable conative propositional attitudes (see e.g., Blackburn, 1998, Ridge, 2014, or Schroeder, 2008: 136–141). So, these are accounts of normative belief that account for it in terms of psychological *reduction*.

One main advantage of locating expressivism's second commitment in the foundational theory of mind is that one opens *additional* theoretical resources to account for the nature of normative judgement, *without* incurring commitment to such a reduction—a reduction that is traditional expressivism's main source of problems. If an expressivist takes this route, she is no longer restricted to the explanatory resources first-order psychology offers. She does not have to appeal to certain conative attitudes, their structure, phenomenology, or contents as models. Instead, she can hold that the nature of normative judgements can be explained in other terms, e.g., by appealing only to their distinctive causal-functional role or biological function *independently of whether there are any conative attitudes that share this role or function*.

Of course, a reduction of normative beliefs to suitable (conative) propositional attitudes might fall out of an account that cashes out expressivism's second commitment in the foundational theory of mind. But such views are not committed to this, and the relocation opens conceptual space for expressivist views that explicitly resist such a commitment. On such views, normative judgement is, *in first-order psychological terms*, a *belief* and nothing else, while the interesting theoretical claim made by expressivism's second commitment is purely located at the level of the foundational theory of mind. Hence, these views can, potentially, side-step many of the problems faced by traditional expressivism, while keeping the theoretical advantages of expressivism's second commitment.

Of course, so far, I have only pointed out possibilities. The remaining paper aims to substantiate these claims. It does so by example: it fleshes out what a view of this kind looks like—conceptual role expressivism—and some advantages it has. In doing so, my aim is to *first* substantiate and illustrate the suggestion that we can locate expressivism's second commitment in the foundational theory of mind and to argue *second* that this has distinctive advantages over traditional expressivism.

² Of course, it is quite possible that some authors already, implicitly, understand their view this way or that some views in the literature should already be understood in this way. However, so far no one has made the possibility of this interpretation explicit.

3 Conceptual Role Expressivism

Expressivists who hold that their position is a view in the foundational theory of mind that answers the question in virtue of what normative beliefs are *beliefs* and *have their contents*, face two primary explanatory challenges. First, they need an account in virtue of what mental states are beliefs that is compatible with a reading of expressivism's second commitment on which normative judgements are beliefs. Second, they need an account that explains in virtue of what such beliefs have their contents that is compatible with the view that the kinds of states that expressivists hold normative judgements to be, are beliefs with such contents. Let's start with the first challenge.

3.1 Normative Belief for Expressivists

The most natural option for expressivists here is, probably, functionalism about the propositional attitudes. *Functionalism* is the view that mental states are characterized by their *causal-functional role* within cognitive systems. This causal-functional role can be usefully distinguished into three parts.³ First, "mind-entry" conditions, which specify the state's role in the procession of sensory stimuli and other *inputs* to the mind. Second, "mind-to-mind" conditions, which specify the state's role in relation to *other* mental states. This, for example, includes its role in processes such as deliberation, reasoning, or inference. Third, "mind-exit" conditions, which specify the state's role in the production of bodily behavior.

For functionalists about the propositional attitudes, mental states belong to propositional attitude types, such as belief, desire, etc., in virtue of the *types* of mind-entry, mind-to-mind, and mind-exit conditions that characterize them. For example, on this view states are beliefs in virtue of the ways in which they figure in deliberation and inference, combine with desires to produce action, and so on. States are desires, on the other hand, in virtue of, for example, their connection to states of anticipation and pleasure and how they combine with beliefs to produce actions.

Functionalism characterizes mental states by their *relational* properties, by their *role*, not by their intrinsic nature. And it is this feature that makes functionalism a likely place to locate expressivism's psychological commitment. After all, if it is a state's *role* that makes it a belief, there is no *in principle* reason to assume that the role expressivists' associate with normative judgement could not be the role of a kind of belief. Of course, whether this is so, depends on what causal-functional role characterizes beliefs.

Naturally, expressivists should deviate as little as possible from standard accounts of the functional role of belief. On such accounts, I take it, beliefs are characterized by distinctive mind-to-mind conditions, namely a certain role in processes of deliberation and reasoning, and by distinctive mind-exit conditions, namely distinctive

³ See Braddon-Mitchell & Jackson 2007: 47. This kind of systematic distinction for role accounts was first introduced by Wilfrid Sellars (1954).

ways of combining with desires to produce action. Some beliefs are also characterized by distinctive mind-entry conditions, e.g., they are formed in response to sensory inputs, but other beliefs (such as, for example, beliefs with logically complex or modal contents) are unlikely to be characterized by such conditions and are more likely to be formed by inference from other beliefs.

The crucial question here is what sort of causal-functional role would make beliefs robustly representational, i.e., would rule out the kind of expressivism under consideration. A plausible answer that seems to emerge in the literature concerned with this issue (e.g., Price, 2013, Williams 2013, or Dreier, 2004) is that beliefs are robustly representational, if and only if it is part of the explanation of their content that they (in some relevant sense) *track* certain properties. On a functionalist reading, this would suggest that a belief is robustly representational, if and only if the account of their functional role must refer to certain properties in a certain way.

Of course, it is a tricky question *how* the relevant properties must figure in the account of a belief's content for it to be robustly representational. Typically, the stance expressivists (and like-minded theorists (e.g., Williams 2013 or Price, 2013)) take is that tracking must be understood *causally*, e.g., such that the property *causes* the belief in certain conditions. It should be noted, though, that it could be possible to understand tracking non-causally, e.g., to accommodate that mathematical beliefs are representational. For example, Michael Williams (2013: 140) suggests that one way for a belief to be representational is when the account of its content must refer to a reliable discriminatory disposition to form the belief in the presence of certain things. While Williams understands this in causal terms, note that thinkers could have such a reliable disposition due to some third-factor explanation, meaning the property tracked itself is causally inert (Enoch, 2011: 177–184). Unfortunately, the issue to what extent “tracking” could be understood in non-causal terms is underexplored and no fleshed proposal is available. While going further into this issue is beyond the scope of this paper, it should be flagged as an important area for future research.⁴

⁴ One worry that might be raised here is that expressivists cannot really appeal to non-causal tracking to distinguish robustly representational beliefs from beliefs that are not robustly representational. This is so, because all contemporary expressivists are *quasi-realists* who want to accommodate ordinary claims about truth, knowledge, or objectivity. As such, they must concede that normative beliefs *do* track normative properties in some non-causal sense. This might suggest that a causal notion of tracking is the only notion expressivists can appeal to as the mark of robustly representational belief. While I concede that this is a worry to take seriously, let me provide some license for optimism that expressivists could get around this. First, and most importantly, what matters on the account suggested is that tracking is part of the *explanation of the content of the belief*. So, what would be the mark of robustly representational belief for the quasi-realist is not whether belief non-causally tracks certain properties, but whether that tracking must be invoked to explain the content of the belief. But, while quasi-realists might concede that normative beliefs track properties, they will deny that this claim is part of their meta-ethical account of the contents of normative beliefs. In this way, the worry is avoided. Second, it might be possible to make the distinction in terms of a theoretically or metaphysically robust notion of tracking for which quasi-realists would not accept that normative beliefs track properties in this sense. For example, maybe it is possible to understand “tracking” by invoking reference magnetism and the metaphysical notion of “eliteness” (see e.g., Suikkanen 2017), for which it is unlikely that quasi-realists will accept that normative beliefs track normative properties in this sense. Thanks to an anonymous referee for raising this worry.

In any case, expressivists who take this route of cashing out belief should accept that belief is characterized by a distinctive role in processes of inference, deliberation, and reasoning, and a distinctive way of combining with desires to produce action but resist the idea that belief's functional role must refer to properties in the relevant way. Expressivists can, of course, concede that such reference marks certain *kinds* of beliefs. All they must deny is that such conditions are part of what makes states *beliefs*.

This expressivist-friendly account of *belief* is not unmotivated. First, it is plausible to think that certain beliefs are not characterized by a tracking causal-functional role anyways. This is true, for example, for (the above mentioned) logically complex or modal beliefs. Such beliefs are, plausibly, primarily characterized by their role in inference and deliberation, not tracking conditions. Second, there has been some work on the compatibility of traditional expressivism with the idea that normative judgements are beliefs (e.g., Blackburn, 1998, Gibbard, 2003, Lenman, 2003, Ridge, 2006, or Sinclair, 2006). What has emerged here is that there are plausible accounts of belief that do not *require* belief to be robustly representational. Rather, it seems that, indeed, belief's role in inference and deliberation, as well as how it combines with desire is primary for our understanding of belief. We should now turn to the second issue: In virtue of what do beliefs have their contents?

3.2 Normative Contents for Expressivists

This question concerns the *foundational theory of mental content*, which in contemporary philosophy of mind has been primarily pursued as part of the project of *naturalizing mental content*.⁵ Prominent views here are e.g., causal-informational theories (e.g., Fodor, 1979 or Dretske, 1981), interpretative theories (e.g., Davidson, 1973, Dennett, 1987, or Lewis, 1974), teleo-functional theories (e.g. Millikan, 1984; Neander, 2017; Papineau, 1984), and conceptual role theories (e.g., Block, 1986; Field, 1978 and Harman, 1999).

These views all attempt to explain in virtue of what mental states have their contents. And, except for purely causal-informational accounts, there is scope for developing an expressivist account within any of these camps.⁶ Here, I will focus on the views that I think lend themselves most naturally for an expressivist-friendly framework. These are *conceptual role* views about mental content.⁷

⁵ Of course, expressivists won't just need an expressivist friendly account in virtue of what normative beliefs have their contents, but also an account of the nature of these contents. Expressivists have a variety of options that would fit the framework suggested here (e.g., Brown 2019, Field 2001, Ridge 2014, and Köhler 2018, 2017. Even Gibbardian contents might work (Gibbard 1992, 2003), though as noted below, on this kind of account they'd play a different explanatory role than on Gibbard's account.

⁶ James Brown (2022), for example, has recently suggested an expressivist view that is located in an interpretative foundational theory of mind.

⁷ I've suggested an expressivist view in these terms in Köhler 2017, 2018. Others have also suggested versions of expressivism cashed out within (or combined with) conceptual role semantics (e.g., Båve 2013, Brown 2019, Chrisman 2017, Horwich 2010, or Sinclair 2018). Note that Ralph Wedgwood (2008, 2001) was the first author who explicitly defended a conceptual role view in meta-ethics. However, his

According to conceptual role views mental states *also* have their *contents* in virtue of their causal-functional role.⁸ So, for conceptual role views it's not just the difference between *types* of propositional attitudes that is accounted for in terms of causal-functional role. Rather, the difference between states of the same kind with different contents (and the similarity between states of different kind with the same content), i.e., the difference between e.g., the belief *that there is beer in the fridge* and the belief *that the cat is outside*, is *also* explained in this fashion. Specifically, the difference between states of the same kind with different contents depends on concrete differences in the mind-entry, mind-to-mind, and mind-exit conditions that characterize them. So, for example, all beliefs are characterized by how they figure in deliberation and inference and combine with desires to produce action. However, the difference between the belief *that there is beer in the fridge* and the belief *that the cat is outside* depends on the concrete mind-to-mind, mind-exit, and mind-entry conditions that characterize them. For example, the belief *that there is beer in fridge*, but not the belief *that the cat is outside* is formed in response to seeing beer in the fridge. According to conceptual role views, mental states have their content in virtue of these *concrete* ways in which the causal role of a mental state operates. Let us call these features of mental states' functional role their *conceptual role*.

What is important about such views is that they do not impose strong restrictions on the kinds of conceptual roles in virtue of which beliefs have their contents. Of course, there will be *some* restrictions on what conceptual roles could possibly account for the contents of beliefs. At the very least, it needs to have a conceptual role that allows it to play belief's distinctive role in deliberation and the production of actions. At least *in principle* though, different types of conceptual roles that differ significantly among each other in other ways should be compatible with that requirement. In particular, it allows the space to draw the distinction characteristic of expressivism, but completely on the level of the theory in virtue of what normative beliefs have their content. Here is, for illustration, one way this could go⁹:

First, the requirement is compatible with there being beliefs characterized by the following kind of conceptual role: It is characterized by significantly robust mind-entry conditions, which are such that we are required to expand our ontology or show that those conditions function to track something already part of our ontology, unless we want these beliefs to be misfiring systematically. When it comes to mind-exit conditions, on the other hand, these beliefs will be characterized by conditions according to which they play—at most—a contributory role in the production of action in combination with desires. It seems very plausible that beliefs with this type

Footnote 7 (continued)

view is *not* a form of expressivism and Wedgwood did not discuss the possibility that conceptual role semantics could be used by expressivists to their advantage.

⁸ It is an open question whether conceptual roles must be understood, at least partially, in normative terms (e.g., in terms of their role in *rational* agents). I will assume that conceptual roles can be fully understood in non-normative terms but remain open to a view on which conceptual roles are normative in some minimal sense (a view I do not take to conflict with expressivism (see e.g., Toppinen 2015).

⁹ Here, I am drawing on Köhler 2017: 204. Note that in that paper I still say that the expressivist will characterize the relevant state to be a kind of conative attitude. I no longer think that needs to be true.

of conceptual role have the causal-functional profile of representational states. And expressivists can hold that descriptive beliefs are beliefs of exactly this kind.

Second, the restriction is *also* compatible with there being beliefs with a very different conceptual role. If this conceptual role is characterized by mind-entry conditions at all, these conditions will be such that they do not expand our ontology.¹⁰ This could be the case, for example, because those mind-entry conditions only specify restricted conditions which *rule out* the belief characterized by them. But the conceptual role would be characterized by robust mind-exit conditions which provide these beliefs with a motivating role in the production of action (*in addition* to their supporting role in producing action with certain sorts of desires). It seems very plausible, however, that beliefs with a conceptual role of this type will have a causal-functional profile very similar to that of directive mental states. And expressivists can hold that normative beliefs are exactly of this kind.

Note that this is exactly the kind of distinction expressivists traditionally want to draw between normative and descriptive judgements. However, the distinction is completely drawn at the level of the foundational theory of mind, not in first-order psychology. Hence, with the resources developed, we can offer a form of expressivism that fully locates its second commitment in the foundational theory of mind. I will call any view that draws the distinction in this way by appealing to conceptual role semantics, a form of *conceptual role expressivism*.

Note furthermore that the distinction is drawn only in terms of the explanation of how different kinds of *beliefs* get their contents. So, on this view it turns out that for the expressivist normative judgements are simply *beliefs*: they do not belong to any other propositional attitude type, even if they might have some *similarities* to other types. In fact, on this view, neither are normative judgements any sort of conative attitude, nor do they belong to a *third* type of state between belief and desire. What is distinctive about normative judgements is not that they are fundamentally a different type of state compared to other types of beliefs, but in virtue of what they have their contents. That this is possible is a distinctive advantage of understanding expressivism as a view in the foundational theory of mind.

One important further thing to notice about conceptual role expressivism is that there is space for several different views of this kind, depending on how one cashes out the conceptual role in virtue of which normative beliefs have their contents. While expressivism's second commitment is not compatible with *every* non-representational role normative judgement might play, there is *some* space for different types of roles normative judgements could play. Importantly, the *action-guiding* role sketched above is just one of many possible ways to go.

With this view on the table, my thesis that we can relocate expressivism's second commitment to the foundational theory of mind is well supported. This clears up the possible conceptual space in which expressivist views can be located and enriches expressivists' conceptual and theoretical resources. However, I will now

¹⁰ For an illustration what such a picture of mind-entry conditions this might look like, see Christine Tiefensee's (2023) recent account of thick concepts for inferentialists, which should work *mutatis mutandis* for conceptual role expressivism too.

also argue that conceptual role expressivism has distinctive advantages over forms of traditional expressivism, *due* to its location in the foundational theory of mind. This, then, supports and illustrates that locating expressivism's commitment in the foundational theory of mind has distinctive theoretical and explanatory benefits for expressivists.

4 Conceptual Role Expressivism and the Unity of Normative Belief

A first advantage of conceptual role expressivism is that it opens space for a more plausible picture of normative belief. Historically, meta-ethical theories have been only concerned with the *moral* domain, which was seen as having certain features that were regarded as particularly philosophically problematic. However, the growing consensus in meta-ethics is that the moral domain shares these features with others, which can all be grouped together into the *normative* domain and that it is *this* domain that should concern meta-ethics. This domain includes, for example, the prudential and epistemic domains. A plausible meta-ethical theory should provide an account of all these beliefs and what they have in common.

For traditional expressivism normative judgements consist in some conative propositional attitude, similar to desires, intentions, or emotions. This view, hence, comes with a distinctive form of *judgement internalism*, on which normative beliefs (tend to) *motivate* (rational) agents to perform actions in the way such attitudes do. Specifically, it comes with a commitment to the idea that normative beliefs stand in a particular causal relation—call it “motivation”—to certain responses, in the sense that we can offer an *intentional* explanation of the response by appealing (at least in part) to the normative belief. Given that meta-ethical theory aims to provide a unified account of normative judgement, this should be traditional expressivists' view about *all* normative judgements. However, this picture is problematic.¹¹

The main problem is that not all normative judgements motivate us (e.g., Chrisman, 2016 or Gregory, 2017). While desires, intentions, and emotions motivate *actions*, not all normative judgements can do so. For example, Matthew Chrisman (2016: 180–1) points out that judgements about the past, judgements about fundamentally different kinds of agents from ourselves, or certain non-agentive normative judgements such as

- (1) I ought to have begun my higher education at a different college.
- (2) NATO ought not to expand any farther.¹²
- (3) The victim's identity ought to be protected by the court.

are not capable of motivating action. Similarly, Alex Gregory (2017: 36–39) argues that *other*-regarding normative judgements, such as

¹¹ My argumentation here draws on Boulton and Köhler 2020.

¹² See Ridge 2017 for a way how traditional expressivists might deal with (2).

(4) Paul ought to buy a present for his wife.

do not plausibly motivate agents to act. And, of course, these worries are much more pronounced when we consider *epistemic* judgements, such as

(5) (Based on the evidence) I ought to believe that climate change is real.

where it is unclear what the connection to *action* would be.¹³

A tempting response is that in such cases normative judgement does not motivate action, but other responses, such as reactive attitudes or beliefs (see e.g., Gibbard (2003, 2012) or Ridge (2007)). On this view, expressivists would hold that normative judgement consists in conative attitudes characterized by a motivational connection not just to action, but to all kinds of (using Scanlon's (1998: 20) term) *judgement-sensitive* attitudes, i.e., intention, belief, desire. However, on closer inspection this view is either not plausible or effectively gives up on the idea that normative judgements are conative attitudes without providing a theoretical alternative.

Assume we take the suggested view seriously and hold that all normative judgements are *conative* states, but that they motivate judgement-sensitive attitudes. So, for example, epistemic judgements would be like *plans* to form beliefs. This is not particularly plausible: First, it is unclear whether conative attitudes *can* motivate these sorts of judgement-sensitive responses in the systematic fashion presupposed by this view. Second, it is implausible to think of the connection between normative judgements and how we respond to them as *motivation* unless our response is an intention or action. Believing in response to epistemic judgement, for example, is not plausibly an instance of being *motivated* to believe.

What if we hold that the connection in these cases is not motivation, but some other way in which normative judgement is connected to judgement-sensitive attitudes? The problem is that it is now unclear why we should assume that normative judgements are *conative* attitudes, as they then behave quite differently from the familiar conative attitudes. Of course, you might hold that they are not, in fact, conative attitudes, but similar sorts of attitudes. One problem with this response, though, is that from a psychological perspective, we now no longer have any sort of model for accounting for these attitudes. Furthermore, this raises the problem to provide a unifying *psychological* account of what holds these different attitude types—some of which are desire-like, some of which are not—together.

Note that the problem here does not simply boil down to problems that affect judgement internalism generally. After all, the problems do not affect *weaker* forms of judgement internalism. For example, one could accept that normative judgement *is* necessarily connected to certain responses but deny that the connection is *motivation*. For example, we could accept that (in rational agents) making an epistemic judgement necessarily comes with forming the corresponding belief but deny that

¹³ Klemens Kappel and Frederick Moeller (2014) and Michael Ridge (2018) have suggested forms of traditional expressivism that provide such a connection. However, Ridge's view only works with a strong form of pragmatic encroachment, making expressivism robustly hostage to fortune. For problems with Kappel and Moeller's view see Boulton & Köhler (2020): 742–46. For general problems with the idea the epistemic judgements could be motivating, see Boulton & Köhler (2020).

this can plausibly be understood in terms of *motivation*. The problem is that it does not seem as if traditional expressivism can really endorse such a weaker form of internalism, as its model for the connection *is* motivation.

Conceptual role expressivism does not face these worries because it, at no point, needs to say that normative judgements are conative attitudes. For conceptual role expressivists, what holds normative belief together is a specific type of *conceptual role*. For example, they can hold that normative beliefs are characterized by *directive* conceptual roles, but where this is not to be understood along a desire-like model, but in terms of a distinctive structuring, organizing, regulating, and controlling *causal-functional* role in our cognitive economy vis-à-vis other mental states. So, for example, on this view epistemic judgements structure, organize, regulate, and control our *beliefs* in distinctive ways, moral judgements structure, organize, regulate, and control our *intentions* and *reactive* attitudes in distinctive ways, and so on.

This view would not face the challenges Chrisman or Gregory raise, because it allows for a wider range of ways in which normative judgements can be directive, without having to understand *being directive* analogous to motivation. For example, we might hold that first-person normative judgements like

(1) I ought to have begun my higher education at a different college.

do not motivate action, but constrain future intentions by e.g., ruling out intentions formed based on similar reasoning as the one that led one to study at one's actual undergraduate institution. (1) also might lead me to form regret. Judgements such as

(2) NATO ought not to expand any farther.

or

(4) Paul ought to buy a present for his wife.

on the other hand, do not motivate actions, but structure, organize, regulate, and control our reactive attitudes in ways aimed at facilitating e.g., cooperation.

Obviously, this picture requires fleshing out. For example, a full account needs to explain what characterizes causal-functional roles that structure, organize, regulate, and control other mental states. Furthermore, such an account needs to explain what kinds of mental states are structured, organized, regulated, and controlled by what normative judgements and in what circumstances. However, even this rough picture is sufficient to make my point.

This is that such conceptual roles offer expressivists a unified picture of normative beliefs and the role they play in our mental economy, without having to appeal to conative attitudes in their explanation. Conceptual role expressivists can appeal to something broader, of which, for example, those normative judgements that *do* motivate action are a special case, such as a structuring, organizing, regulating, and controlling conceptual role. Hence, contrary to traditional expressivism, conceptual role expressivism *can* hold on to the idea that normative judgements are directive, while giving up on the idea that all normative judgements are conative attitudes, because it *does* provide a general theoretical account for all these judgements—an account

we could not give with only the resources first-order psychology offers. This way, conceptual role expressivism allows expressivists to adopt a more plausible version of judgement internalism on which normative judgements are not connected to *motivation*, but instead to other mental states that they regulate. This, in turn, highlights that when we understand expressivism's commitments as commitments in the foundational theory of mind, we get theoretical resources that allow the expressivist to develop their account in new directions that have distinctive advantages over the more traditional versions.

Of course, it might turn out that some conative attitudes have *similar* structuring, organizing, regulating, and controlling *causal-functional* roles to some normative beliefs. In this case, there will be a respect in which the conceptual roles of those conative attitudes and certain normative beliefs will be similar. Note, though, that this holds only to some extent, because the role of normative beliefs will be broader than this—for example, normative beliefs will have a conceptual role *also* characterized by the kind of role in inference and deliberation characteristic for beliefs. Hence, it is not the case that those attitudes *will be* normative judgements, unless they also play this broader role. Note, furthermore that this does *not* mean that normative beliefs should be accounted for by using these conative attitudes as a model. Instead, it just means that on a more fundamental level, these psychological attitudes will to some extent receive similar treatments when we explain their contents.

5 Conceptual Role Expressivism and Frege-Geach

Another advantage of conceptual role expressivism is its relation to the Frege-Geach Problem, or more specifically, to a similar challenge that arises at the meta-semantic level. The *Frege-Geach Problem* is the challenge to expressivists to provide an account of the *semantics* of normative sentences that can account systematically for the contents of complex sentences in which normative sentences figure. It arises on the understanding of expressivism's *first* commitment as a semantic view that forces expressivists to reject orthodox approaches to semantics such as truth-conditional semantics and so, requires them to offer an alternative that does the same explanatory work.¹⁴

I will assume here that we can avoid *this* challenge by adopting a *meta-semantic* form of expressivism. This is a recent reinterpretation of expressivism's first commitment as a *meta-semantic* claim (e.g., Chrisman, 2016, Ridge, 2014, Silk, 2013, or Sinclair, 2021). The difference between semantics and meta-semantics as I will understand this distinction here, is, roughly, this (for a similar distinction, see Lewis, 1970 or Speaks, 2019): *Semantics* attempts to develop a theory of the literal meanings (the semantic contents) of items in natural languages in a way that, for example,

¹⁴ The most prominent alternatives have been developed by Blackburn (1993), Gibbard (2003), and Schroeder (2008). For problems with these views, see e.g., Schroeder, 2008: 39-55 and 164-187. Of course, the Frege-Geach Problem is still a live avenue of research (see e.g., Baker & Woods 2015, Charlow 2014, 2015, Ridge 2014, or Sinclair 2021).

explains their compositionality. *Meta-semantics*, on the other hand, aims to explain in virtue of what natural language sentences have their literal meanings.

Understanding expressivism's first commitment as a meta-semantic claim has the benefit of making expressivism to some extent *semantically neutral*. In particular, it allows expressivists to concede the legitimacy of orthodox semantic approaches and to incur no burden of developing an alternative semantics. Of course, expressivists still need interpretations of core semantic terms like "truth," "reference," "proposition," etc. on which their employment in semantics is compatible with expressivism. But this can be achieved, e.g., by establishing that expressivism's commitments are compatible with e.g., deflationary (or otherwise expressivist-friendly) readings of these terms, which are themselves compatible with orthodox semantic approaches (see e.g., Burgess, 2011, Field, 1994, Köhler, 2017, Sinclair, 2021, and Williams, 1999)).

However, even if this removes the semantic challenge, meta-semantic expressivism still faces a similar challenge in meta-semantics and the foundational theory of mind. After all, meta-semantic expressivists must show that they can provide a plausible story that tells us in *virtue of what* normative sentences and beliefs have the kinds of contents required for orthodox semantics: they must *earn the right* to employing contents in semantics, by showing that their meta-semantic story allows normative sentences to have contents with the right sorts of features. However, when it comes to *this* challenge, conceptual role expressivism has distinctive advantages over traditional expressivism, simply in virtue of the foundational level at which the former locates this challenge. To explain why this is so, let me first clarify what specific challenge conceptual role expressivists face. The starting challenge is to explain how, if meta-semantic expressivism is true, normative sentences have the kinds of contents orthodox semantic approaches associate with them. This challenge is easily met for conceptual role expressivists: they will hold that normative sentences have these contents in virtue of being associated with beliefs that have the appropriate sorts of contents. This is, probably, the response that other meta-semantic expressivists will give as well.

However, if we take this response, we now require an account in virtue of what, on an *expressivist* view, normative beliefs have contents of the right kind. What are contents *of the right kind*? Minimally, I'd assume that they must be contents that have the right *inferential* features to plausibly function as the contents of meaningful declarative sentences: they must be contents that can connect to the contents of other beliefs in just the way in which normative sentences can meaningfully combine with other sentences in ever more complex ways. Hence, the challenge that meta-semantic forms of expressivism face is this: in virtue of what do normative beliefs have contents with the right inferential features?

This is, plausibly, a variant of the Frege-Geach Problem. And the advantage of conceptual role expressivism—as opposed to meta-semantic forms of traditional expressivism—is that for *conceptual role expressivism*, the challenge is not how we can explain that there are beliefs with such contents by providing an explanation in terms of conative attitudes, their contents, and their structural features. Rather, the challenge is how to explain the contents of these beliefs at the level in which we explain the contents of *any* sort of mental state. And as it turns out,

this turns the challenge just into a particular instance of a general challenge in the philosophy of mind, namely the challenge to explain, given that our mental states have contents of a certain kind, in virtue of what features our mental states have these contents.

This is good news for conceptual role expressivists for two reasons. First, there are good reasons to assume that expressivists have no special problem with dealing the challenge under discussion at this level compared to representationalists (see, e.g., Pérez Carballo, 2019). Second, at this level, conceptual role expressivists have resources available that would not help traditional expressivists. While I cannot, here, offer a full account, I will make substantial suggestions at how such an account might proceed and why it would be in line with commitments we already take on when we adopt conceptual role semantics. What I will suggest is that conceptual role expressivists can fruitfully draw on resources offered by Blackburn (1993) and Gibbard (2003) to address the Frege-Geach Problem in semantics, but without the problems these responses generate there. Let me briefly describe how the response works and what worries have been raised about it, to then turn to my explanation of how conceptual role expressivists might appropriate that response. Doing so will also highlight why the challenge at the meta-semantic level won't be as easy to deal with for forms of traditional expressivism.

What I take to be the core of Blackburn's and Gibbard's response is to offer a formal model that generates the inferential structure characteristic of declarative contents, without making commitments that require the model's components to be robustly representational. Specifically, they offer formal models that generate the inferential structure characteristic of declarative contents, but where the "contents" of these declarative sentences might be non-cognitive states. They do this, by taking inferential relations between mental states as explanatorily basic. This allows us to offer a compositional semantics that is neutral on whether the mental states expressed by declarative sentences are non-cognitive attitudes or ordinary beliefs—as long as we can support that these inference patterns can be instantiated among such a variety of mental states.

Let's give a more concrete example, using Blackburn's (1993) version for simplicity's sake. His idea is that we can enrich any language by introducing two operators $H!$ and $T!$. These are supposed to be related such that $H!p$ is equivalent to $\neg T!p$ and $\neg H!p$ equivalent to $T!p$. And, in particular, $H!p$ and $T!p$ are supposed to be *inconsistent*. Blackburn then works out how taking inferential transitions, rather than truth, as basic, we can generate inferential tree structures within the language that includes the operators that satisfy the normal rules of compositionality, as well as validity. But, if we take $H!$ to stand for an attitude of approval and $T!$ for an attitude of tolerance and assume that they satisfy the stipulated rules, then, so Blackburn argues, we have a formal model for an expressivist-friendly language that satisfies compositionality.

The biggest problem with this sort of proposal has been highlighted by Mark Schroeder (e.g., 2008: 48/49): as Schroeder argues the only uncontroversial form of inconsistency expressivists can appeal to is the inconsistency that we find in attitudes of the same kind with inconsistent contents—inconsistency in virtue of attitudes' abilities to "transmit" the inconsistency of their contents. This sort of inconsistency,

so Schroeder points out, is something *anyone* has to appeal to in their explanation, as it is the inconsistency that we find in belief, for example. However, the problem is that the attitudes assumed to be inconsistent in the formal model above (and similarly for Gibbard's formal model) are not inconsistent in that sense.

For example, take the above suggestion that it is the ordinary psychological attitudes of disapproval and tolerance as the attitudes signified by the H! and T! operators. Then H! and T! will not stand for attitudes of the same kind that are inconsistent by appealing to their inconsistent contents. After all, tolerance and disapproval are attitudes of *different* kinds and for Blackburn's model to work they need to be inconsistent when held toward *consistent* contents. So, for Blackburn's model to work we must assume that there is an additional type of inconsistency between attitudes of *distinct* types that holds when their contents are consistent. And, as Schroeder goes on to argue, we must assume that there is an infinite hierarchy of such distinct non-cognitive attitude types that can stand in such relations. So, all the model offers, is a list of things that must be the case for expressivist semantics to be true. And, what we find, once we check whether our psychology offers attitudes that satisfy this list, is that we must postulate infinitely many distinct types of conative attitudes that all stands in inexplicable inconsistency relations to each other—not in virtue of their inconsistent contents, but in virtue of something else. This simply is not a good psychological assumption to make.

The main problem with Blackburn's (and similarly with Gibbard's) response to the Frege-Geach Problem is, therefore, that our psychology does not seem have enough non-cognitive attitudes standing in the right sorts of relations to each other to make their suggestions work. This is why the challenge located at the meta-semantic level is not going to be easier to meet for traditional expressivism than the Frege-Geach Problem is for semantic expressivism. After all, to provide an appropriate account of the contents of normative beliefs, these expressivists still must identify suitable conative attitudes that stand in the right sorts of relations such that the relevant beliefs can be reduced to those attitudes—which runs them straight back into the problem just discussed.¹⁵ However, these problems will not emerge when we move to the foundational theory of mind, as I will now demonstrate using conceptual role expressivism. In fact, moving to this level opens a way to still use a model like Blackburn's (or Gibbard's) to deal with the challenge at hand. Let me explain.

First, it is important to emphasize again that conceptual role expressivists undertake no commitment to the idea that to account for the inferential role of normative beliefs we need to identify certain conative attitudes those beliefs are identical to that stand in the right sorts of inferential relations to each other. All conceptual role expressivists are committed to is that normative judgements are beliefs with

¹⁵ Of course, there has been some discussion of Schroeder's problem and how to deal with it as a traditional expressivist (see e.g., Baker & Woods 2015 or Sinclair 2021: 137–142). I am not going to go into a discussion of such responses, because my main point here is that expressivism as a view in the foundational theory of mind simply avoids the issue altogether. Note also that hybrid forms of traditional expressivism (e.g., Ridge 2014) might have an advantage here over non-hybrid forms of expressivism. However, conceptual role expressivism still fares better than these when it comes to offering a plausible form of judgement internalism across all normative domains.

certain propositional contents. So, they are not tied to any unusual views in psychology about the relationship between certain sorts of mental states. Rather, what conceptual role expressivists offer is a particular sort of account in virtue of what these beliefs have their contents.

How, then, do they account for the contents of logically complex beliefs like.

(6) If eating meat causes pain, then eating meat is wrong.

or

(7) It is not wrong to eat meat.

as well as other logically complex beliefs with normative components? And how do they account for the inconsistency of (7) with

(8) It is wrong to eat meat.

For expressivists who locate their commitment in the foundational theory of mind, this challenge just is the challenge of offering an account that explains in virtue of what *beliefs* have contents that can embed in logical complex constructions, and which is compatible with their distinctive story about *normative* beliefs. So, all they need to address that challenge are two things: first, a general account of beliefs that have the relevant kinds of contents. Second, they need to show that beliefs with an expressivist conceptual role can fall under this account. If this can be shown, there is no need to bring in other attitudes or stipulate that normative beliefs can be identified with some further propositional attitude type beyond belief. Normative judgements would just be beliefs with the relevant kinds of contents. Can conceptual role expressivists shoulder this burden? Yes, they can.

Let's start with the question what sort of general account conceptual role expressivists can give of beliefs that have the relevant sorts of contents. Here, it seems that conceptual role expressivists should just follow whatever account is already given by conceptual role semantics. I take it that conceptual role views take a similar stance on this as Blackburn: they take role in deliberation, reasoning, and the like as explanatorily basic (see e.g., Harman, 1999, Block, 1986, or Field, 1978). However, where it seems that Blackburn's suggestion should be understood as taking inferential operations on *contents* as basic, conceptual role theorists hold that the fact that beliefs *have* contents that stand in inferential relations to each other, is something that is *explained* in terms of their causal-functional role in processes of reasoning, deliberation, etc. that are cashed out without mentioning contents themselves. Hence, on this view deliberation, reasoning, etc. are explanatory fundamental in an even stronger sense: beliefs have the kinds of contents that can embed in logically complex constructions in virtue of the causal-functional role those beliefs play in the mental processes of reasoning and deliberation. Let's take a brief look at what that account might say more specifically, to allow us to better see whether it is compatible with conceptual role expressivists commitments about normative beliefs.

Let's start with "atomic" beliefs, i.e., beliefs with contents that are logically atomic. On the conceptual role view I will presuppose here, such beliefs have such contents *because* they, at least partially, have their contents in virtue of their

causal-functional roles in processes such as inference, deliberation, and so on, and how these beliefs interact with logically complex beliefs in such processes (here I follow the account that Block (1986) and Field (1978) seem to endorse). For example, a crucial part of the conceptual role of any belief that F will be that it interacts in specific ways in processes of reasoning with a belief of the kind.

(9) If F, then G.

For example, a mental state is the belief that F in virtue of tending to (causally) move the thinker in processes of reasoning to the belief that G, if combined with (9), and so on. That is part of *what it is* to be the belief that F. And note that this holds for any type of atomic belief—independently of what F is supposed to be.

Similarly, though, for beliefs with logically complex contents: these beliefs have the kinds of contents that they do in virtue of how they interact with the relevant beliefs with atomic contents in processes of inference and deliberation. For example, any belief of the type given by (9) has its content in virtue of, among other things, moving someone who believes that F to also believe that G, moving them to abandoning the belief that F, upon coming to believe that not G, and so on. That is, this belief has its content in virtue of being a dispositional state characterized by these sorts of mental transitions. This holds for any instance of this type of belief irrespective of the content of “F” and “G” and the distinct instances of this type will only be distinguished by the beliefs they will (causally) move thinkers into. Hence, the difference between a belief in

(10) If it rains, then the streets are wet.

and a belief in

(11) If lying is wrong, so is getting your little brother to lie.

consists in things such as, for example, that a belief in (10) moves the thinker into the belief that the streets are wet, when she comes to believe that it rains, while a belief in (11) moves the thinker into the belief that getting your little brother to lie is wrong upon coming to believe that lying is wrong.

Similar things will hold for beliefs in negated contents, beliefs in disjunctions, beliefs in conjunctions, and so on, which are all also characterized by the causal-functional role they play in processes of reasoning, deliberation, and the like. And, of course, playing a certain sort of role in processes of reasoning, deliberation, etc. in relation to these types of beliefs with logically complex contents will be part of the conceptual role of the relevant atomic beliefs. For example, a crucial part of the conceptual role of the belief that F is that thinkers who are in that state tend to not move into the belief that not F, tend to abandon the belief that F if they come to believe that not F, and so on.

This provides a picture of how a conceptual role view would account for beliefs with logically complex contents and in virtue of what beliefs have contents that can stand in these sorts of complex inferential relations. On this view, it is their role in inference and deliberation in virtue of which these beliefs have these contents. The

crucial question, of course, is whether and how conceptual role expressivists can make the case that this account is also legitimately applicable to normative beliefs.

I think it is already relatively transparent that it is, because there is no reason to hold that the kind of role just sketched is incompatible with the conceptual role that expressivists identify for normative beliefs. However, we can make the case even clearer and more systematically by putting a different theoretical spin on Blackburn's (and Gibbard's) models to account for beliefs with logically complex content with normative components. Basically, we interpret these models as homing in, *not* on distinct conative attitudes, but on the causal-functional structure of states with certain *conceptual roles*. Specifically, we can see these models as identifying an inferential structure amongst formal objects that is structurally isomorphic to the causal-functional structure of certain conceptual roles, namely the causal-functional structure of the conceptual roles in virtue of which beliefs have contents that stand in inferential relations to each other. However, what these models specifically allow us to do, is to pick out this causal-functional structure in a way that leaves room for some states with those conceptual roles to play a wide variety of *other* mind-input, mind-to-mind, and mind-output roles. That is, these models show us that playing the kind of causal-functional role in inference and deliberation spelled out above, which is characteristic for beliefs with the relevant kinds of contents is in principle *compatible* with the conceptual role of some beliefs also being partially characterized by the sort of *directive* role required by expressivism's second commitment. Hence, they show us that there can, in principle, be *beliefs* with the conceptual role that expressivists suggest.

Assume, for example, that in Blackburn's model "disapproval" and "tolerance" do not actually stand for the psychological attitudes of disapproval and tolerance, but for conceptual roles that are partially characterized by mind-output conditions. In this case, we can take Blackburn's formal model as a kind of *possibility proof* that states that are characterized by a conceptual role with the kinds of mind-output conditions required by expressivism's second commitment could still also be characterized by the relevant role in reasoning and deliberation.¹⁶

Furthermore, the model provides us (at least partially) an account of what the conceptual role of beliefs with logically complex contents that have normative components would be like. The conceptual role of such beliefs would just be an instance of the role in inference and deliberation that is characteristic of all beliefs of that type, an instance that is individuated by what other beliefs it is related to in what way in its conceptual role—similarly to how we individuate (10) and distinguish it from (11) above. So, what the model would show us is that the account we would give for the content of *normative* beliefs with logically complex contents would be no different to the account we would give for the content of any other kind of belief with logically complex contents, namely in terms of a type of causal-functional role in processes such as reasoning and deliberation.

¹⁶ Of course, such a possibility proof itself is not sufficient to *earn the right* to saying that expressivist normative concepts legitimately have the relevant kinds of conceptual roles. This requires a further explanatory step. See Köhler (forthcoming) for an account that deals with this issue.

Note that on this view, conceptual role expressivists do not need to postulate a distinctive form of inconsistency to explain the inconsistency of inconsistent normative beliefs: *all* beliefs would be inconsistent with other beliefs in virtue of certain sorts of causal-functional phenomena, specifically the causal-functional relations between inconsistent beliefs. In this respect, the H! and T! beliefs will be no different from ordinary descriptive beliefs. However, this is the sort of phenomenon *any* naturalistic theory of content will likely have to appeal to, to explain why beliefs with inconsistent contents are inconsistent. Expressivists will also not have to appeal to an infinite hierarchy of logically distinct, but somehow inconsistent *sui generis* psychological attitudes to account for the infinite number of possible logically complex normative judgments and sentences. They only must appeal to belief-types that can be realized by infinitely many different causal-functional roles (that are instances of a finite number of causal-functional role *types*). But any functionalist must accept this.

To summarize, expressivists who take their view just to be a distinctive claim in the foundational theory of mind how certain types of beliefs get their content can just take over a general account provided by conceptual role semantics to explain in virtue of what beliefs have contents that can be logically complex. They can then argue that on their account normative beliefs can legitimately be assumed to have such contents, using e.g., a reinterpreted version of Blackburn's account.

6 Conclusion

This paper has argued for two claims: First, that expressivism's commitment about the nature of normative judgements can and should be located in the foundational theory of mind. Second, that doing so opens conceptual space for novel and attractive expressivist views that have distinctive advantages of their more traditional competitors. I've argued for these two claims via an expressivist view—*conceptual role expressivism*—that has its theoretical commitments fully located at the level of the foundational theory of mind and by arguing that this view side-steps certain problems faced by traditional expressivism. While my discussion was mostly programmatic, the paper has also moved conceptual role expressivism itself forward, by highlighting unappreciated benefits it has and flagging potential for further development.

Acknowledgements For helpful and constructive feedback and discussion, I would like to thank Emad Atiq, James Brown, Matthew Chrisman, Jessica Isserow, Silvia Jonas, Nick Laskowski, Michael Ridge, Christine Tiefensee, Teemu Toppinen, Herman Veluwenkamp, as well as two anonymous referees for *Erkenntnis*.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the

material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baker, D., & Woods, J. (2015). How expressivists can and should explain inconsistency. *Ethics*, 125(2), 391–424. <https://doi.org/10.1086/678371>
- Båve, A. (2013). Compositional semantics for expressivists. *The Philosophical Quarterly*, 63(253), 633–659.
- Blackburn, S. (1993). Attitudes and contents. In *Essays in quasi-realism* (pp. 182–197). Oxford University Press.
- Blackburn, S. (1998). *Ruling Passions*. Oxford University Press.
- Block, N. (1986). Advertisement for a semantics for psychology. *Midwest Studies in Philosophy*, 10, 615–678.
- Boult, C., & Köhler, S. (2020). Epistemic judgement and motivation. *The Philosophical Quarterly*, 70(281), 738–758. <https://doi.org/10.1093/pq/pqaa007>
- Braddon-Mitchell, D., & Jackson, F. (2007). *The philosophy of mind and cognition*. Blackwell.
- Brown, J. (2019). Expressivism and cognitive propositions. *Journal of the American Philosophical Association*, 5(3), 371–387.
- Brown, J. (2022). Interpretative expressivism: A theory of normative belief. *Philosophical Studies*, 179, 1–20.
- Burgess, A. (2011). Mainstream semantics + deflationary truth. *Linguistics and Philosophy*, 34(5), 397–410.
- Charlow, N. (2014). The problem with the Frege-Geach problem. *Philosophical Studies*, 167(3), 635–665.
- Charlow, N. (2015). Prospects for an expressivist theory of meaning. *Philosophers' Imprint*, 15(23), 1–43.
- Chrisman, M. (2016). *The meaning of "ought": Beyond descriptivism and expressivism in metaethics*. Oxford University Press.
- Chrisman, M. (2017). Conceptual role accounts of meaning in metaethics. In T. McPherson & D. Plunkett (Eds.), *The Routledge handbook of metaethics* (pp. 260–274). Routledge.
- Davidson, D. (1973). Radical interpretation. *Dialectica*, 27, 314–328.
- Dennett, D. (1987). *The Intentional Stance*. MIT Press.
- Dreier, J. (2004). Meta-ethics and the problem of creeping minimalism. In *Philosophical Perspectives*, 18, 23–44.
- Dretske, F. (1981). *Knowledge and the flow of information*. MIT Press.
- Dretske, F. (1988). *Explaining Behaviour: Reasons in a World of Causes*. MIT Press.
- Enoch, D. (2011). *Taking morality seriously, a defense of robust realism*. Oxford University Press.
- Field, H. (1978). Mental representation. *Erkenntnis*, 13, 9–61.
- Field, H. (1994). Deflationist views of meaning and content. *Mind*, 103, 249–285.
- Field, H. (2001). Attributions of meaning and content. In H. Field (Ed.), *Truth and the absence of fact* (pp. 157–174). Oxford: Oxford University Press.
- Fodor, J. (1968). *Psychological Explanations*. Random House.
- Fodor, J. (1979). *The Language of Thought*. Harvard University Press.
- Gibbard, A. (1992). *Wise Choices, Apt Feelings*. Harvard University Press.
- Gibbard, A. (2003). *Thinking How to Live*. Harvard University Press.
- Gibbard, A. (2012). *Meaning and Normativity*. Oxford University Press.
- Gregory, A. (2017). Are all normative judgments desire-like? *Journal of Ethics and Social Philosophy*, 12(1), 29–55.
- Harman, G. (1999). (Nonsolipsistic) conceptual role semantics. In G. Harman (Ed.), *Reasoning, meaning, and mind* (pp. 207–232). Oxford University Press.
- Horwich, P. (2010). The Motive Power of Evaluative Concepts. In *Truth-meaning-reality* (pp. 167–196). Oxford: Clarendon Press.

- Kappel, K., & Moeller, E. F. L. (2014). Epistemic expressivism and the argument from motivation. *Synthese*, 191(7), 1529–1547.
- Köhler, S. (forthcoming). Saving conceptual role expressivism from defect. *The Journal of Philosophy*.
- Köhler, S. (2017). Expressivism, belief, and all that. *Journal of Philosophy*, 114(4), 189–207. <https://doi.org/10.5840/jphil2017114416>
- Köhler, S. (2018). Expressivism, meaning, and all that. *Canadian Journal of Philosophy*, 48(3–4), 337–356. <https://doi.org/10.1080/00455091.2017.1408278>
- Lenman, J. (2003). Disciplined syntacticism and moral expressivism. *Philosophy and Phenomenological Research*, 66(1), 32–57.
- Lewis, D. (1966). An argument for the identity theory. *Journal of Philosophy*, 63, 17–25.
- Lewis, D. (1970). General semantics. *Synthese*, 22, 18–67.
- Lewis, D. (1974). Radical interpretation. *Synthese*, 27, 331–344.
- Millikan, R. (1984). *Language, Thought, and Other Biological Categories*. MIT Press.
- Neander, K. (2017). *A mark of the mental, in defense of informational teleosemantics*. Cambridge, MA: MIT Press.
- Papineau, D. (1984). Representation and explanation. *Philosophy of Science*, 51, 550–572.
- Pérez Carballo, A. (2019). Negation, expressivism, and intentionality. *The Philosophical Quarterly*, 70(279), 246–267. <https://doi.org/10.1093/pq/pqz047>
- Price, H. (2013). *Expressivism, Pragmatism and Representationalism*. Cambridge University Press.
- Putnam, H. (1967). The nature of mental states. In *Mind, language, and reality* (pp. 429–440). Cambridge University Press.
- Ridge, M. (2006). Saving the ethical appearances. *Mind*, 115(July), 633–649.
- Ridge, M. (2007). Expressivism and epistemology: epistemology for ecumenical expressivists. In *Aristotelian Society Supplementary*, 81, 83–108.
- Ridge, M. (2014). *Impassioned Belief*. Oxford University Press.
- Ridge, M. (2017). Expressivism and collectives. *Mind*, 127(507), 833–861.
- Ridge, M. (2018). How to be an epistemic expressivist. In C. McHugh, J. Way, & D. Whiting (Eds.), *Metaepistemology* (pp. 141–158). Oxford University Press.
- Scanlon, T. M. (1998). *What we owe to each other*. Harvard University Press.
- Schroeder, M. (2008). *Being For: Evaluating the Semantic Program of Expressivism*. Oxford University Press.
- Sellars, W. (1954). Some reflections on language games. *Philosophy of Science*, 21, 204–228.
- Silk, A. (2013). Truth conditions and the meanings of ethical terms. In R. Shafer-Landau (Ed.), *Oxford studies in metaethics* (pp. 195–222). Oxford: Oxford University Press.
- Sinclair, N. (2006). The moral belief problem. *Ratio*, 19(June), 249–260.
- Sinclair, N. (2018). Conceptual role semantics and the reference of moral concepts. *European Journal of Philosophy*, 26, 95–121.
- Sinclair, N. (2021). *Practical Expressivism*. Oxford University Press.
- Speaks, J. (2019). Theories of meaning. In Edward N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/spr2021/entries/meaning/>. Accessed 11 May 2023.
- Suikkanen, J. (2017). Non-naturalism and a reference. *Journal of Ethics and Social Philosophy*, 11(2), 1–25.
- Tiefensee, C. (2023). Through tick and thin: Seamless metaconceptualism. *Synthese*. <https://doi.org/10.1007/s11229-022-03913-6>
- Toppinen, T. (2015). Expressivism and the normativity of attitudes. *Pacific Philosophical Quarterly*, 96(2), 233–255.
- Wedgwood, R. (2001). Conceptual role semantics for moral terms. *The Philosophical Review*, 110, 1–30.
- Wedgwood, R. (2008). *The Nature of Normativity*. Oxford University Press.
- Williams, M. (1999). Meaning and deflationary truth. *Journal of Philosophy*, 96, 545–564.
- Williams, M. (2013). How pragmatists can be local expressivists. In H. Price (Ed.), *Expressivism, pragmatism and representationalism* (pp. 128–144). Cambridge University Press.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.