

THE MOTIVATION PROBLEM OF EPISTEMIC EXPRESSIVISTS

ALEXANDRE DUVAL

School of Philosophy, Australian National University

CHARLES CÔTÉ-BOUCHARD

Department of Philosophy, Montmorency College

Many philosophers have adopted epistemic expressivism in recent years. The core commitment of epistemic expressivism is that epistemic claims express conative states. This paper assesses the plausibility of this commitment. First, we raise a new type of problem for epistemic expressivism, *the epistemic motivation problem*. The problem arises because epistemic expressivists must provide an account of the motivational force of epistemic judgment (the mental state expressed by an epistemic claim), yet various features of our mental economy seem to show that they can't do so. Second, we develop what we take to be the most promising response to that problem for expressivists. We end by noting that this response faces an important challenge pertaining to the psychology of epistemic criticism and praise.

Keywords: epistemic expressivism; metaepistemology; epistemology; epistemic normativity; knowledge attributions; intentional attitudes

1. Introduction

In recent years, several philosophers have argued that we should be *expressivists* about epistemic discourse.¹ The core commitment of epistemic expressivism is that epistemic claims express conative states, states that aim to change the world rather than to describe it. Various philosophers have argued that this

1. Important works developing epistemic expressivism include Gibbard (1990; 2003; 2012), Blackburn (1996), Chrisman (2007; 2012), Ridge (2007; 2018), Field (2009; 2018), Kappel (2010; 2011), Kyriacou (2012), Kappel and Moeller (2014), Ahlstrom-Vij (2013), and Schafer (2014).

Contact: Alexandre Duval <alexandre.duval@anu.edu.au>
Charles Côté-Bouchard <charles.cote79@gmail.com>

view addresses issues related to the semantics and metaphysics of knowledge attributions and justification attributions. For instance, Chrisman (2007) has argued that it solves semantic problems about the context-sensitivity of knowledge attributions that even epistemic contextualism can't solve. Field (2009) has pointed out that it allows epistemologists to get rid of a problematic, metaphysical account of how justification attributions work. More generally, one of the main attractions of the view is that it purports to account for the normative character of epistemic discourse without positing any metaphysically loaded properties and facts.

So, there are strong motivations to adopt epistemic expressivism. But we are not convinced that it can make good on its core commitment. In this respect, our paper has two main aims. The first is to raise a new type of problem for epistemic expressivism.² This problem—which we call *the epistemic motivation problem*—arises from the fact that it is a constraint on epistemic expressivists that they be able to provide an account of the motivational force of epistemic judgment, yet various features of our mental economy seem to show that they can't do so. The second aim is to develop what we take to be the most promising avenue for expressivists in response to that problem. After spelling it out, we raise an important challenge for that response, a challenge that arises from the psychology of epistemic criticism and praise.

We are ultimately agnostic as to whether the challenge to our response can be met. Either way, we think that the analysis in this paper helps to make significant progress in the debate about the plausibility of epistemic expressivism. If its proponents *can* answer the challenge, then our discussion provides a clear path to develop and defend epistemic expressivism's main commitments in the philosophy of mind and philosophy of action. On the other hand, if its proponents *cannot* answer the challenge, our analysis in Sections 3 to 5 will suggest that epistemic expressivism is untenable.

Here is how we will proceed. In Section 2, we clarify epistemic expressivism and the kind of conative attitude its proponents take epistemic claims to express. In Section 3, we present the epistemic motivation problem. In Section 4, we develop what we take to be the best response to it. Finally, in Section 5, we assess our response by presenting the challenge based on the psychology of epistemic criticism and praise.

2. For other objections against epistemic expressivism specifically (not just expressivism in general), see Kvanvig (2003), Cuneo (2007), and Lynch (2009). There are also some objections to epistemic expressivism that carry over from work on *moral* expressivism, most notably the Frege-Geach problem. See Schroeder (2010) for a presentation of the standard version of the problem, and Kyriacou (2010) for discussion of the Frege-Geach problem as applied to epistemic expressivism.

2. Epistemic Expressivism's Core Commitment

Epistemic expressivism is a view about the type of mental states that *epistemic claims* express. As understood here, epistemic claims are statements that involve epistemic notions (e.g., 'knowledge' and 'justification') and/or that attribute distinctly epistemic obligations, permissions, prohibitions, virtues, and values to an agent or their beliefs. (An attribution is *distinctly epistemic* when it is made from an epistemic perspective, as opposed to a clearly non-epistemic perspective like a moral or prudential one.) Typical epistemic claims include statements of the following forms: 'S knows that P', 'S is (epistemically) justified in believing that P', 'S's belief that P is (epistemically) irrational', 'S (epistemically) ought to believe that P', 'S (epistemically) should not believe that P', and 'S has sufficient overall evidence to believe that P'.

The core commitment of epistemic expressivism is that epistemic claims express *conative states* (possibly combined with other types of states). Conative states, according to a common characterization, are states that do not aim to represent the world, but rather to change it. They contrast with garden-variety beliefs. Typical conative states include desires and intentions.

What kind of conative attitude are epistemic claims supposed to express? The precise answer varies from author to author, and it depends on the general form of expressivism they endorse. Take for example the influential theory developed by Gibbard (2003). It focuses on the notion of *contingency plan*. Contingency plans are conative mental states about what to do or believe in a variety of scenarios that may or may not come to pass. The central tenet of his theory is that claims of the form 'S knows that P' express contingency plans to *rely on S's judgment that P*. For example, on his view, the claim 'Mary knows that it is raining' expresses the attributor's plan to rely on Mary's judgment that it is raining in their own decision-making.

Gibbard's account has the form of what is called a *pure* expressivist theory. Pure theories maintain that epistemic claims express conative states (e.g., contingency plans) and no other types of states apart from that. But there is another category of expressivist theories of epistemic discourse. They are known as *ecumenical* theories. Ecumenical theories hold that epistemic claims express complex mental states that include a conative component *as well as* an ordinary belief component.

Consider for instance the account put forward by Ridge (2007). It proposes that claims of the form 'S knows that P' express complex mental states, which include the following components:

- (I) A conative state of "[e]pistemic endorsement of certain procedures for deciding what to believe" (Ridge 2007: 103).

- (II) “The belief that *S*’s judgement that [P] is causally regulated by either (a) those procedures (anaphoric reference back to those procedures the speaker endorses in [I]) or (b) procedures which are close enough to those procedures, so far as [P] goes or (c) more fully informed successors to those procedures” (Ridge 2007: 103).

Suppose that someone says ‘Mary knows that it is raining’. On Ridge’s account, this claim expresses a complex state which includes the attributor’s endorsement of specific epistemic procedures and the belief that Mary’s judgment that it is raining was causally regulated by procedures that satisfy conditions (a), (b), or (c). What kind of procedures are we talking about? Suppose the attributor saw Mary looking outside and noticing the tell-tale visual signs of rain. It is possible then that the relevant procedures that the attributor endorses are procedures such as *When you see raindrops outside, form the belief that it is raining outside*, or perhaps more generally *When you receive direct perceptual evidence of an event, form the belief that this event is happening*.

Ridge’s theory is ecumenical because it holds that the mental state expressed by an epistemic claim includes an ordinary belief (clause (II) above). In fact, many expressivist theories of epistemic discourse also satisfy this characterization (e.g., Chrisman 2007; Kappel & Moeller 2014).

Let us end with three terminological notes. First, we follow common usage in employing the expression *epistemic judgment* to refer to the full mental state expressed by an epistemic claim. So, for example, within the context of Ridge’s account of knowledge claims, this expression refers to a complex mental state that includes the conative component (I) and the belief component (II). Using this terminology, the core commitment of epistemic expressivism can be reformulated as the thesis that epistemic judgments are conative states or comprise conative states as subparts. Second, we use the phrase *epistemic endorsement* to pick out the purely conative component of an epistemic judgment. (We don’t mean to build any implicit constraints into this phrase, or to rule out by fiat expressivist theories that have a different form than Ridge’s. For instance, the phrase can be meaningfully deployed when talking about Gibbard’s account, and within the context of that theory it refers to a contingency plan of the sort described above.) The introduction of this phrase allows us to highlight one of the main differences between pure and ecumenical theories: whereas pure theories hold that epistemic endorsement constitutes the *whole* of epistemic judgment, ecumenical theories maintain that epistemic endorsement is a *proper* part of epistemic judgment. Third, we treat the phrase *epistemic expressivism* as encompassing both types of expressivist theories of epistemic discourse.

3. The Epistemic Motivation Problem

We are now in a position to formulate the new problem that we raise for epistemic expressivism: *the epistemic motivation problem*. The problem is that a good case can be made for two claims that, when taken together, entail that epistemic expressivism is untenable. The first claim is that epistemic expressivists face an explanatory challenge to provide an account of the motivational force of epistemic judgments. The second claim is that epistemic expressivists cannot adequately address that challenge.

The argument for the first claim can be presented quickly. Consider the mental state of epistemic *endorsement*. Epistemic expressivists take it as a purely conative attitude rather than as a cognitive attitude, as explained in Section 2. Moreover, philosophers of all stripes assume that having motivational force is a central, if not a definitional property of conative states. In fact, it is usually taken as a mark of conative states versus cognitive states, perhaps the only distinguishing mark between the two categories. It follows that, from the perspective of epistemic expressivists, epistemic endorsement *must have* motivational force.³ As a result, the same conclusion applies to epistemic *judgment*. That's because it includes epistemic endorsement as a part (either proper or improper) and mental states inherit the motivational force of their parts. This, in turn, means that epistemic expressivists must answer the following question: What is the *motivational force* of epistemic judgments? This is the explanatory challenge at the heart of the epistemic motivation problem.

3. There might seem to be a way out of accepting this implication. It has become common among *moral* expressivists to endorse minimalism about the notion of *cognitive state* (or equivalent notions like *representational state* or *representation*; see, e.g., Blackburn 2001). Minimalism about *cognitive state* is the view that there is nothing more to say about the nature of cognitive states than the following biconditional: A mental state M is a cognitive state iff M can be expressed by a declarative sentence. For various reasons, we expect that many epistemic expressivists will similarly adopt minimalism about the notion of *cognitive state*. Among those who do, some might then be tempted to claim that the attitude of epistemic endorsement counts as a cognitive state and so does not have any specific motivational force. But we think that making such a claim would be ill-advised. The adoption of minimalism about notions like *cognitive state* in combination with expressivism about *moral* discourse generates a well-known problem, the problem of creeping minimalism (Dreier 2004). And there is widespread agreement in metaethics that the best solution for the problem involves appealing to the specific psychological role—cashed out in terms of motivational force—of the conative component expressed by moral claims to explain the content of these claims (see, e.g., Dreier 2004; 2015). Since the dialectical situation concerning minimalist views in metaepistemology is similar to the one in metaethics, it is safe to assume that a parallel conclusion will be reached in metaepistemology: all epistemic expressivists need to provide an account of the motivational force of epistemic endorsement, including those who adopt minimalism about *cognitive state*.

This leads us to the second claim: epistemic expressivists cannot adequately address that challenge. The argument for that claim requires a much longer discussion, and it constitutes the burden of the rest of this section. Our presentation proceeds by considering various accounts of the motivational force of epistemic judgments that have been developed or hinted at in the metaepistemological literature. For each of them, we argue that it is either implausible or inadequate for the purpose of handling the challenge.

A few notes about what follows. First, the accounts that we consider differ importantly regarding what they say about the *motivational targets* of epistemic judgments, namely *what* such judgments motivate. As we see it, there are six broad options here: Epistemic judgments could motivate (1) doxastic states like beliefs, (2) actions, (3) habits and related entities, (4) affective states, (5) conative states, or (6) a combination of the previous options.⁴ We consider each of these options in the rest of this section (Subsections 3.1–3.6).⁵

Second, the accounts that we discuss below vary significantly with respect to the *type* of epistemic judgments that they take as their central cases. For instance, the proposals that we consider in Section 3.1 were formulated with *first-personal* epistemic judgments in mind, which include judgments like *I ought to believe that P*, *I know that P*, and *I have sufficient overall evidence to believe that P*. On the other hand, the view that we assess in Section 3.2 focuses on knowledge attributions, judgments of the form *S knows that P* (where the agent making the judgment may or may not be identical to the subject S). Given this variation in judgment-type, we present and assess each account based *only* on what it says about the types of epistemic judgments that it takes as its central cases. We do this because we want to avoid implicitly assuming that the motivational force of epistemic judgments has to be globally uniform across all epistemic judgments, that all types of epistemic judgments have to motivate the same type of motivational targets in the same way. We think that it would be unfair to saddle epistemic expressivists with such a constraint. Another reason for proceeding this way is that we want to show that epistemic expressivists cannot avail themselves of these accounts

4. We don't have a knock-down argument to the effect that this list is exhaustive. However, there is a serious challenge for anyone who says that it isn't. Given the types of entities just cited, the remaining candidates for motivational targets are mental states that are at once non-doxastic, non-conative, and non-affective. The challenge is thus to explain what these states are, what kind of dispositional profile they might have, and why some of them have a close enough relationship with epistemic behavior to count as *proper* motivational targets of epistemic judgments. Moreover, seeing as none of the research in metaepistemology has yet found it necessary to posit relations between epistemic judgments and states of that type, we doubt that the challenge can be met. For that reason, we will put this possibility aside in the rest of the paper.

5. See Boulton and Köhler (2020) for a detailed overview of the first two options. What we say in Sections 3.1 and 3.2 below largely follows their way of delineating and formulating the main versions of the claim that epistemic judgments have motivational force, though we adopt a slightly different terminology.

even for the types of judgments for which the accounts were designed, which arguably should be the judgments for which they work best.

3.1. Epistemic Judgments and Doxastic Motivational Force

The first option we assess is that epistemic judgments motivate doxastic states like beliefs. In that case, we talk about the proposed *doxastic motivational force* (DMF) of epistemic judgments. Consider the following version of that first option:

DMF-Strict: When an agent A makes a first-personal epistemic judgment, the judgment *strictly motivates* A, at least to some extent, to revise their beliefs in accordance with the judgment.⁶

We need to say a few words here about the notion of *strict motivation*, which comes from Boulton and Köhler (2020). Here is how they characterize it:

Strict motivation is the kind of relation that holds, perhaps most paradigmatically, between desire-like states—desires, intentions, plans, and the like—and the actions they produce. More precisely, strict motivation is the sort of relation holding between a response and a set of mental states that makes the response a fitting target for a folk-psychological *intentional* explanation. (Boulton & Köhler 2020: 740–41, emphasis in original)

Note that by ‘intentional explanation’ Boulton and Köhler mean *explanation of intentional responses* rather than *explanation that gives an explanatory role to intentional mental states* (as the expression is sometimes used in philosophy of mind). In a helpful clarification further in the text, they write: “As we understand the idea, a response is a fitting target for intentional explanation if that response is the outcome of a connection between a set of mental states and the agent’s *intending to φ*” (2020: 748, our emphasis). Hence, strict motivation is the causal connection that is involved when we correctly explain someone’s action in terms of their *intending to φ*—for example, when someone is properly describing as having moved to their eyes to the left as a direct result of *intending to move their eyes to the left*.

6. This—or something close to it—is defended by Mitova (2011). See Boulton and Köhler (2020: 746) for discussion. We’ve included the clause ‘at least to some extent’ to emphasize the possibility that strict motivation may come in degrees. This might sound like an oxymoron. But the claim that strict motivation comes in degree follows directly from the characterization of strict motivation cited below, and the popular assumption that intrinsically motivational states like desires and intentions come on a continuum of strength.

Is DMF-Strict plausible? We don't think so. It has to do with the fact that DMF-Strict seems to imply *direct doxastic voluntarism*, the view that we can have direct control over our beliefs. Given the characterization of strict motivation provided above, DMF-Strict entails that epistemic judgment can cause belief revision and that, when it does, the causal connection falls under the purview of intentional explanation. From this, it appears to follow that (i) first-personal epistemic judgments like *I ought to believe that P* either constitutively involve or regularly cause a desire-like state to believe that P and that (ii) this desire-like state can, in turn, cause the adoption of the belief that P in a way that involves the subject being properly described as adopting the belief as a direct result of their *intending* to believe that P. However, the claim that such desire-like states to believe that P have this type of intention-involving causal power is a standard construal of the view that we can have direct control over our beliefs (see, e.g., Alston 1988; Chuard & Southwood 2009). And, as Boulton and Köhler (2020: 749) observe, this view — what has become known as 'direct doxastic voluntarism' — is a very costly commitment to take on. For most epistemologists and action theorists, all we have is an indirect form of control over our beliefs. At best, we can intentionally perform certain actions that might affect what we believe.⁷

If DMF-Strict is untenable, is there any other option for those theorists who maintain that epistemic judgments have doxastic motivational force? Taking Beddor (2016) as inspiration, Boulton and Köhler (2020: 750–52) outline one such option. It is to understand motivation in a *loose* sense rather than a *strict* sense. They take the resulting proposal as the most promising account of the motivational force of epistemic judgments proposed in the literature.

DMF-Loose: When an agent A makes a first-personal epistemic judgment, the judgment *loosely motivates* A, at least to some extent, to revise their beliefs in accordance with the judgment.

What is loose motivation? They characterize it as “a causal relation between a mental disposition and its outputs that is significantly broader than strict motivation, and doesn't have to share strict motivation's characteristic features” (Boulton & Köhler 2020: 741). Yet, to avoid implying that just any kind of causal relation will do, they add that “the relevant sort of causal connection [. . .] must be restricted in some way” (Boulton & Köhler 2020: 741). For example, they suggest, as a first pass, that loose motivation may have to be construed as involving a *rational* sort of causal connection, rather than a merely brute causal connection.

7. For a similar argument, see Boulton and Köhler (2000: 747–48). See also Boulton and Köhler (2020: 748–49) for convincing counter-replies to potential responses to the argument.

The upside of adopting DMF-Loose for epistemic expressivists is that it provides the required latitude to escape DMF-Strict's problematic commitment to direct doxastic voluntarism. It does so by allowing them to deny that the motivational role of epistemic judgments is that of typical desire-like states involved in intentional action—that is, anything like desires, intentions, decisions, or plans in the ordinary sense.

This latitude brings with it a major downside, however. It's that DMF-Loose relies on a notion (*loose motivation*) that is characterized almost exclusively in negative terms: loose motivation is "significantly broader than strict motivation", yet it "doesn't have to share strict motivation's characteristic features". And this in turn raises questions as to whether the notion picks out any real kind of mental phenomena. Boulton and Köhler develop a challenge to proponents of DMF-Loose that zeroes in on this issue. As they describe it, the challenge "is to give an account of what 'loose motivation' is, as well as an account of in virtue of what loose motivation holds between two relata" (Boulton & Köhler 2020: 753). A satisfactory response would have to meet several constraints, such as the following: "it must plausibly distinguish loose motivation from other causal connections, and must give an account of the kind of disposition [linking the relata] that is relevant for loose motivation" (Boulton & Köhler 2020: 753).

Boulton and Köhler convincingly argue that the challenge is substantial. They do so by making the case that we should reject a natural way of trying to answer it. The way in question consists in "build[ing] an account of loose motivation by using strict motivation as a model while trying to abstract away from the features of strict motivation that are problematic in, for example, the epistemic case" (Boulton & Köhler 2020: 754). To explain why we should reject this strategy, they take the ordinary state of *planning*, which possesses strict motivational force, as a case study. As they show, abstracting away from the features of planning that give it its strict motivational force to build an account of loose motivation consists in "merely assuming the features we wanted an account of" (Boulton & Köhler 2020: 757). They conclude that appealing to the notion of loose motivation to specify the motivational force of epistemic judgments isn't completely justified yet because loose motivation remains nebulous, and its existence has been posited without any independent constraints to identify and clarify its core features.

Boulton and Köhler themselves remain ultimately agnostic as to whether the challenge they develop for DMF-Loose can be met. We are much more pessimistic. We think that their challenge is likely fatal. The easiest way to see why is precisely to put their challenge in the context of the current inquiry about the plausibility of epistemic expressivism. Epistemic expressivism is not the sole extant position about the nature of epistemic judgments in metaepistemology. Its most direct rival is *epistemic cognitivism*. Epistemic cognitivism is the view that epistemic judgments are beliefs in the most straightforward sense and not conative

states at all. On this view, the judgment *that I ought to believe that it is raining* just consists in the straightforward belief *that I ought to believe that it is raining*.

How does the possibility of endorsing epistemic cognitivism undermine DMF-Loose for epistemic expressivists? Here is the crux. Epistemic expressivists who endorse DMF-Loose are committed to the existence of a *reliable and systematic connection* between first-personal epistemic judgments and belief revision in accordance with these judgments. When the former are made, the latter often happens. It would thus be natural for these epistemic expressivists to try to mimic moral expressivists in arguing that their position should be favored over epistemic cognitivism on account of the fact that it can *best explain* this connection involving first-personal judgments.⁸ But this is one place where the analogy to the metaethical debate surrounding moral expressivism breaks down: Epistemic expressivists are in a much *worse* position than epistemic cognitivists to explain this connection. On the one hand, as Boulton and Köhler's challenge highlights, epistemic expressivists who adopt DMF-Loose arguably cannot explain that connection because they help themselves to a new type of motivation (loose motivation) that seems about as mysterious as this reliable connection to whose existence they are committed. On the other hand, epistemic cognitivists can easily explain the reliable connection between first-personal epistemic judgments and belief revision. In fact, the explanation falls almost directly off the central commitment of epistemic cognitivism. Epistemic cognitivists can simply say that epistemic judgments, because they are straightforward beliefs, reliably cause belief revision through *theoretical inferences*. The notion of theoretical inference—understood as a transition from a set of beliefs as input states to a further belief as output state without the mediation of other types of propositional attitudes—already looms large in philosophy of mind and epistemology, and it seems as well-understood as anything else in these subfields. In particular, epistemic cognitivists can maintain that there are theoretical-inferential rules for the elimination of first-person epistemic operators (e.g., operators like *I ought to believe that* and *I have sufficient overall evidence to believe that*).⁹ For instance, they could hold that possessing the belief *that I ought to believe that P* will cause an

8. For more on this kind of argument in the moral case, see Sinclair (2007) and Schroeder (2010: ch. 1).

9. The expression 'elimination rule' is often used in the context of discussing conceptual role semantics (e.g., Peacocke 1992). In that context, the expression generally carries two conditions for a proposed inferential rule to count as an *elimination rule* with respect to concept C: (i) the premisses of the rule's instances contain C but the conclusion of the rule's instances do not contain C; (ii) it is part of C's possession conditions that someone possessing C will follow that rule. In contrast, we use the expression 'elimination rule' here in a deflationary sense that only requires condition (i). This deflationary sense is compatible with the conceptual role framework as well as other approaches to conceptual content. Similar considerations apply to our use of 'introduction rule' below.

individual to acquire the belief *that P* in the right conditions. This would parallel the case of elimination rules for logical operators. (Consider the *&-elimination* rule: Possessing the belief *that P & Q* will cause an individual to acquire the belief *that P* in the right conditions.) On this picture, we should expect a reliable connection between any given type of first-personal atomic epistemic judgment and belief revision provided that its epistemic operator has an elimination rule that falls into one of two categories: (a) rules that directly give rise to the belief *that P* through a single application; (b) rules that give rise to the belief *that P* in virtue of their operator being linked to the belief *that P* by a sequence of elimination and introduction rules. For example, the elimination rule for *I ought to believe that* proposed above falls under category (a), and as such, it predicts a reliable connection between the judgment *I ought to believe that P* and the belief *that P*. To get an example of (b), consider the following plausible elimination rule for *I have sufficient overall evidence to believe that*: Possessing the belief *that I have sufficient overall evidence to believe that P* (perhaps in combination with other beliefs, like the belief *that I have no countervailing evidence against P* or the belief *that the practical stakes in believing that P are not too high*) will cause an individual to acquire the belief *that I ought to believe that P* in the right conditions. The existence of such a rule would predict a reliable connection between the judgment *I have sufficient overall evidence to believe that P* and the belief *that P* in virtue of the judgment's inferential links with the operator *I ought to believe that*.

So, in light of the fact that epistemic cognitivists can easily explain the connection between first-personal epistemic judgments and belief revision by appealing to well-understood and independently motivated notions from philosophy of mind and epistemology, Boulton and Köhler's challenge to DMF-Loose makes the latter look like an unpromising proposal for epistemic expressivists to endorse. More generally, the considerations reviewed in this section bolster the view that epistemic expressivists shouldn't attribute doxastic motivational force to epistemic judgments, be it in the strict sense or loose sense of motivation.¹⁰

10. Could versions of DMF-Strict and DMF-Loose formulated for *third-personal* epistemic judgments fare better than the first-personal versions of these two principles considered in the main text? We don't think so. Such principles would be susceptible to an even more damaging objection. Take for instance third-personal judgments of the form *From their perspective, S (subjectively) ought to believe that P* or *From their perspective, S is rational/reasonable in believing that P*, where the subject S is distinct from the agent A making the claim. Sometimes, we sincerely make such judgments without believing that P is true or probable. Worse, we sometimes strongly believe that P is false. But we make such judgments precisely because we think that, from S's perspective, all the evidence that S has strongly points toward the truth of P. This may happen in a case where, for instance, we possess specific evidence undermining P that is simply not available to S. It is implausible, for these cases, to claim that making third-personal judgments of the type just proposed would motivate us to form the belief that P or revise our beliefs in a way that is consistent with the truth of P. This, in turn, strongly suggests that DMF-Strict and DMF-Loose would be simply false when formulated to apply to third-personal epistemic judgments.

3.2. Epistemic Judgments and Practical Motivational Force

Let us move to our second option, namely the view that epistemic judgments motivate *actions*. In this context, we talk about the purported *practical motivational force (PMF)* of epistemic judgments. One advantage of this view is that it allows epistemic expressivists to help themselves to the *strict* sense of motivation without facing the objection pertaining to direct doxastic voluntarism raised against DMF-Strict. The main issue with it, though, is to determine a plausible target action. The most central epistemic judgments are not about what to do, but about what to believe. So, it's not obvious what actions such judgments could motivate.

To our knowledge, only one type of answer has been offered to this challenge in the literature, an answer developed by Klemens Kappel and Emil Moeller (2014). Rephrased using Boulton and Köhler's notion of strict motivation, they defend the following account of the motivational force of knowledge attributions:

PMF-Inquiry-termination: When an agent A judges that S knows that P, this judgment strictly motivates A, at least to some extent, to *terminate their inquiry into whether P*.

Unfortunately, we think that this proposal faces a fatal problem. The problem is that there seems to be a much simpler explanation than PMF-Inquiry-termination for why A terminates their inquiry with respect to P after making the judgment that S knows that P. Note that the aim of inquiring into whether P plausibly involves settling the question 'Is P true?'. But if you sincerely judge that someone knows that P, then, since knowledge is factive, you plausibly take the question 'Is P true?' to be settled, and you plausibly believe P yourself. So, a simpler explanation is that A stops inquiring about P after making the knowledge attribution because they take the goal of inquiring into whether P as having been achieved, and the goal then dissipates or loses its motivational power. This explanation invokes a ubiquitous psychological process pertaining to practical reasoning: You stop pursuing a goal because you perceive the goal as having been achieved and the goal then dissipates or loses its motivational power. Given the availability of this simple explanation for why A terminates their inquiry about P, it is not clear at all why we should adopt a more complex explanation that treats inquiry termination as a type of *action* in need of motivation.¹¹

¹¹. Kappel and Moeller try to address this kind of objection to their proposal, but see Ridge (2018: 143–46) and Boulton and Köhler (2020: 744–45) for convincing rebuttals.

Could epistemic expressivists appeal to the *loose* sense of motivation to argue that epistemic judgments have practical motivational force linked to terminating inquiry? Yes, in theory. But it is utterly unclear why they should be tempted to do so—even bracketing Boulton and Köhler’s important challenge against accounts of motivational force based on loose motivation. The problem raised in the previous paragraph can be summarized as follows: It is theoretically unmotivated to assume that making a judgment of the form *S knows that P* would strictly motivate agent A to terminate inquiry with respect to P given that we have a simpler alternative explanation for why people stop inquiring in this context. This is just as true when we replace the word ‘strictly’ by ‘loosely’, and it seems to have the same implications for a version of Kappel and Moeller’s proposal that appeals to loose motivation. Altogether, the foregoing suggests that no variation on their account could help epistemic expressivists explain the motivational force of epistemic judgments. Moreover, absent other explicit proposals about how actions can serve as motivational targets, the analysis of this section seems to rule out the possibility of claiming that epistemic judgments have *practical* motivational force.

3.3. *Epistemic Judgments and Habit-Related Motivational Force*

The next option that we want to discuss is that epistemic judgments have *habit-related motivational force (HMF)*, that is, that they motivate *habits*. This is a view that Kyriacou (2012) seems to flirt with. He defends the idea that people hold an attitude of endorsement toward specific habits as part of some epistemic judgments. Moreover, the type of habits he has in mind are habits of *belief-fixation*. Here is one such habit: the habit of trusting what reliable sources of information say (Kyriacou 2012: 222). This is the type of habit that would lead someone (e.g.) to adopt the belief that dolphins are mammals just after forming the belief that Mary is a reliable source of information as well as the belief that Mary says that dolphins are mammals.

Though Kyriacou himself doesn’t explicitly propose an account of the motivational force of epistemic judgments, it would be natural to use his work as a basis to make the following proposal:

HMF-Belief-fixation: When an agent A makes an epistemic judgment, the judgment motivates A to form or maintain those *habits of belief-fixation* that are the objects of the endorsement part of the judgment.

Does this proposal fare better than the previous ones? Unfortunately, no. There are two possible readings of HMF-Belief-fixation, depending on whether

the notion of motivation employed is understood as *strict* or *loose*. And each interpretation faces an objection similar to the one afflicting the corresponding DMF principle discussed in Section 3.1.

Let us start with the loose-motivation interpretation. The central point of the objection that we presented in Section 3.1 to DMF-Loose is that it relies on a notion (*loose motivation*) that is left largely unexplained. This point is just as damaging when applied to the version of HMF-Belief-fixation involving loose motivation. Changing one of the relata of the motivation relation from *beliefs* to *habits of belief-fixation* doesn't change the dialectical situation in any relevant way.

Turn to the version that invokes strict motivation. The problem with that interpretation is that it seems to have a commitment that is about as objectionable as DMF-Strict's commitment to direct doxastic voluntarism. That it has such a commitment becomes obvious when we realize that forming a new habit of belief-fixation is very similar to just adopting a new belief. Take, for instance, the habit of trusting what reliable sources of information say. And now consider the following belief: the belief *that if S is a reliable source of information and S says that P, then P is (likely) true*. This belief can do the same belief-generating work as the habit just cited. In fact, for any habit of belief-fixation that can be stated in natural language, we can identify a content-related belief which does the same belief-generating work. Now, the version of HMF-Belief-fixation that invokes strict motivation entails that desire-like states can cause the adoption of new habits of belief-fixation in a way that involves A being properly described as adopting those habits as a direct result of *intending* to acquire them. In other words, it entails that we can have *direct control* over our habits of belief-fixation. But consider the difference between saying that someone can have direct control over specific habits of belief-fixation *versus* saying that someone can have direct control over the relevant content-related beliefs. Since we have just established that habits of belief-fixation and the corresponding beliefs play a similar role in one's mental economy, if the latter statement is objectionable, the former should be as well.

It follows that neither reading of HMF-Belief-fixation can do the trick. This in turn indicates that epistemic expressivists cannot maintain that habits of belief-fixation are plausible motivational targets for epistemic judgments. Considerations similar to the one just reviewed also undermine the strategy of claiming related entities—entities such as *procedures* or *policies* about how to revise one's own beliefs, and *dispositions* to form beliefs in certain ways—as motivational targets.

That being said, these considerations do not by themselves spell the end of the idea that epistemic judgments have habit-related motivational force. A variation on the principle just assessed is worth discussing too:

HMF-Inquiry-management: When an agent A makes an epistemic judgment, the judgment motivates A to form or maintain those *habits of inquiry management* that are the objects of the endorsement part of the judgment.

This principle assumes people endorse habits of *inquiry management* as part of their epistemic judgments rather than habits of *belief-fixation*. As understood here, an *inquiry* is a set of actions that someone performs to increase the chances of receiving reliable information about a topic of interest. Given this definition, the following habits will count as habits of inquiry management: the habit of writing down the structure of an argument using formal tools (e.g., propositional calculus, predicate calculus) to assess its validity when its validity is otherwise hard to evaluate, and the habit of verbally or digitally verifying with trusted sources whether a given person, group of people, institution, newspaper, or website has a good epistemic reputation before physically putting yourself in a situation to receive information from it. The main attractive feature of HMF-Inquiry-management is that it allows its proponents to escape any problematic commitments in the vicinity of direct doxastic voluntarism due to its focus on inquiry-related *actions*.

However, this principle faces a major issue of its own. To see what the issue is, it helps to focus on specific types of judgments. Suppose that A makes a justification judgment of the form *S justifiably believes that P*. Epistemic judgments of this form are what Kyriacou's expressivist theory aims to deal with in the first place. Adapted to HMF-Inquiry-management, one of his main proposals is that justification judgments include, among other things, a conative attitude of endorsement of the habits of inquiry management *in virtue of which S has acquired the belief that P* (Kyriacou 2012: 224). The issue with HMF-Inquiry-management is that in general A does not have enough information about S's habits of inquiry management to identify those specific habits in virtue of which S has acquired the belief that P. For instance, let's take US Chief Medical Advisor Anthony Fauci's stated belief that "the hyper or aberrant inflammatory or immunological response [to COVID-19] gives as much to the morbidity and mortality as the actual virus replication itself".¹² We assume that, when prompted, many people would be willing to say that Fauci *justifiably* believes that proposition. Yet, the habits of inquiry management which have led Fauci to believe this presumably stem from complex strategies learned over many years about how to read academic papers in immunology, how to analyze complex medical data and figures, when and how to reach out to other experts when he has doubts about

12. Fauci uttered this sentence in a White House press briefing. See <https://www.whitehouse.gov/briefing-room/press-briefings/2021/02/22/press-briefing-by-white-house-covid-19-response-team-and-public-health-officials-6/>

certain findings, how to lead and supervise the employees at National Institute of Allergy and Infectious Diseases so that the information they provide him is reliable, etc. Therefore, except practicing immunologists, very few will have ideas about the habits of management inquiry that have led Fauci to believe the proposition cited above.

Importantly, we think that this kind of situation—where A makes a justification judgment without possessing enough information to identify the specific habits of inquiry management that have led S to believe that P—doesn't just come up in cases where S has high levels of expertise and P pertains to that expertise. It also takes place in many mundane contexts. Most of the time, when an agent A attributes justification or knowledge that P to some subject S, what guides A's reasoning on the matter is A's rough idea about S's general epistemic character (whether and to which extent S displays specific epistemic virtues or vices) as well as A's assessment of how S's institutional, geographical, and/or social position allows S to receive reliable information about topics related to P. If we are right about this, it follows that in general A does not have enough information about S's habits of inquiry management to identify the habits in virtue of which S has come to believe that P. And this thereby causes trouble for HMF-Inquiry-management because A cannot be motivated to form or maintain habits whose identity they cannot ascertain.

We thus conclude that HMF-Inquiry-management is implausible. Moreover, beyond the two principles reviewed above, it is hard to come up with any other proposal in line with the thesis that epistemic judgments have habit-related motivational force. Therefore, the discussion in this section undermines this thesis.

3.4. *Epistemic Judgments and Affective Motivational Force*

We now want to consider the view that epistemic judgments motivate *affective states*, or dispositions to form such states. In that case, we will talk about the purported *affective motivational force* (AMF) of epistemic judgments. An important challenge here is to determine a plausible target affective state for epistemic judgments to motivate. To our knowledge, Gibbard (2012) is the only author to have made a concrete proposal about the affective motivational role of epistemic judgments. He talks about a disposition to experience “[a]n ‘Oops’ response” following certain types of first-personal epistemic judgments about what one ought to believe (Gibbard 2012: 173).

A natural way of interpreting his suggestion is as follows:

AMF-Feeling-of-error: When an agent A makes an epistemic judgment of the form *For any proposition P, I ought to believe P if conditions C are met*

(or some other equivalent judgment), the judgment motivates A to form a disposition to experience a feeling that they have made a reasoning error when the following conditions obtain: (i) the issue of whether P is true has come up to A's consciousness; (ii) A believes that C are met; (iii) A occurrently believes that they do not believe that P.¹³

Consider the following scenario. Michael forms the judgment *I ought to believe what (the majority of) life scientists say about COVID-19 vaccines*. Then, he receives evidence through a reliable source that most life scientists believe that some vaccine V targeting the virus is safe. However, when asked by a friend whether he thinks that V is safe, he says 'No' and he's sincere. AMF-Feeling-of-error predicts that, in these conditions, Michael would immediately experience a feeling of having made a reasoning error.

Now, we are happy to grant that many people experience such a feeling of error when conditions like (i)–(iii) are satisfied. However, our main goal in this section is to make the case that epistemic expressivists cannot appeal to AMF-Feeling-of-error to characterize the motivational force of epistemic judgments. To do so, we will argue that epistemic expressivists who adopt this proposal cannot provide a satisfactory account of the *dispositional profile* of that feeling. Yet, they need to be able to offer such an account because the denial to do so would amount to the costly rejection of a highly entrenched view in philosophy of mind: the view that token mental states belonging to a certain state-type (e.g., belief, desire, fear, anger, joy, feeling of surprise) have a characteristic dispositional profile.¹⁴

More specifically, we will argue that epistemic expressivists can't explain the *downstream effects* of the feeling of error in someone's mental economy. Suppose that a certain agent A makes an epistemic judgment of the form *For any proposition P, I ought to believe P if conditions C are met*, and there is a proposition P such that conditions (i)–(iii) are satisfied. According to AMF-Feeling-of-error, the agent will then experience a feeling of error. Explaining the downstream effects of the feeling of error amounts to answering the following question: What kind(s) of actions or mental states does that feeling usually get A to do or to form? In what follows, we claim that epistemic expressivists cannot ade-

13. Nothing in the analysis that follows hangs on whether Gibbard's suggestion involves motivation in the strict sense or in the loose sense. So, we've left AMF-Feeling-of-error unspecified in that regard.

14. This view is entailed by many major positions about the metaphysics of mental states—reductive functionalism, Dennettian instrumentalism, and liberal dispositionalism (e.g., Schwitzgebel 2002)—because they individuate state-types at least partly in terms of tokens' dispositional role in the agent's mental economy. Moreover, even among people who reject such positions, many accept the view cited in the main text because they take the dispositional role of token mental states to be *reliably correlated* to their state-type and content. This includes non-reductive functionalists as well as some phenomenal intentionality theorists (e.g., Mendelovici 2018).

quately address this question. To be clear: We are not assuming that epistemic expressivists need to say that the feeling of error itself has a specific *motivational* role. They don't have to say this, as far as we can tell. However, they still have to be able to tell us about the downstream component of the feeling's disposition profile—what the feeling *does* once it's produced in a particular situation.

We can start with an intuitive answer to the question raised in the previous paragraph: the feeling of error gets the agent to *directly* form the belief that P (i.e., without the intervention of any additional intermediary mental state). Unfortunately, as intuitive as it sounds, epistemic expressivists can't avail themselves of this answer. This becomes obvious when we ask whether the feeling of error has intentional content and, if so, what it is. On the one hand, epistemic expressivists cannot say that the feeling of error has intentional content unrelated to the targeted belief (something like *Reasoning error now!*), or perhaps even no intentional content at all. They cannot say this because, if they did, they would block any plausible explanation as to how this feeling could directly lead to the adoption of the right belief—the belief that P—among an infinity of alternative beliefs. For instance, they would leave no space to explain why Michael adopts the belief *that V is safe* following the production of the feeling of error in the scenario described above as opposed to, say, the belief *that the twelfth president of the United States once sported a mustache*.

On the other hand, suppose that epistemic expressivists maintain that the feeling has intentional content that *includes* information specific to the targeted belief (something like *An error was made in reasoning about the safety of vaccine V* in the previous scenario). Then, they will commit themselves to the view that what we've been calling *the feeling of error* is in fact a belief or includes a belief as a subpart. After all, the feeling of error is now hypothesized to have conceptual content (e.g., *An error was made in reasoning about the safety of vaccine V*) as well as the typical dispositional profile of a belief. It has the typical dispositional profile of a belief in that (a) it is formed directly following the reception of evidence that its content is true (e.g., Michael's feeling of error is directly formed following the reception of evidence that he has made an error in reasoning about the safety of vaccine V); (b) it directly guides the formation of other beliefs whose content is systematically related to its own content (e.g., it directly guides Michael to form the belief *that V is safe*). Moreover, by committing themselves to the view that what we have been calling *the feeling of error* is itself a belief or includes a belief as a subpart, epistemic expressivists would thereby endorse a variation of the claim that epistemic judgments motivate *dispositions to form beliefs in certain ways*, a claim which we have argued in Section 3.3 that epistemic expressivists ought to avoid.¹⁵

15. There is another reason for epistemic expressivists to be wary of committing themselves to the view that the feeling of error is itself a belief or includes a belief as a subpart. It is that it

An upshot of the previous paragraph is that epistemic expressivists should *deny* that it is part of the dispositional profile of the feeling of error to get the agent to directly adopt the belief that P. It is worth noting that this is something which Gibbard (2012) himself seems to deny. For him, it is the first-personal epistemic judgment itself that plays the role of getting the agent to directly adopt the belief that P following the production of the feeling of error. After all, in line with his previous work, he holds that an epistemic judgment is a specific type of contingency plan about how to revise one's own beliefs (Gibbard 2012: 169ff.). Here is the kind of model that he has in mind about the role of the feeling of error in someone's mental economy. At time t_1 , A makes the judgment J of the form *For any proposition P, I ought to believe P if conditions C are met*. This immediately motivates the formation of the *disposition* to produce the feeling of error in conditions (i)–(iii). Then, at a later time t_2 , the sudden satisfaction of conditions (i)–(iii) triggers the production of the feeling of error. The feeling of error then alerts the agent that something has gone wrong in one of their spontaneous inferences and that they need to engage in a different kind of reasoning procedure. A few moments after t_2 , the original judgment J is re-activated and thereby motivates A to directly adopt the belief that P.

This is an elegant account. But, if we are right that he subscribes to this model, it commits Gibbard to the claim that first-personal epistemic judgments motivate both the formation of a disposition to produce the feeling of error in conditions similar to (i)–(iii) *and* belief revision. And, in terms of the current discussion, the issue with this claim is simple: it assumes that epistemic judgments have a doxastic motivational role, so it faces the same objections we raised against that option in Section 3.1.

What other type of account of the downstream effects of the feeling of error could epistemic expressivists adopt? Here is another tempting idea: the feeling simply brings attention to the fact that a reasoning mistake has been made *in the last few seconds* so that the agent can start recalling the last few steps in their theoretical reasoning and make the relevant correction. However, to stay clear of the issues just raised, it is stipulated that neither does the feeling of error or the first-personal epistemic judgment directly influence belief revision. It is left to A's other mental states to initiate the relevant belief change.

As appealing as it may seem, epistemic expressivists can't resort to this proposal either. Here is why. Even if a feeling of error brings A's attention to the fact that they've just made mistake, this proposal doesn't leave any space to explain how the error will be *corrected*. Presumably, if A made a mistake in the first place, it is because they had beliefs, perceptions or intuitive reactions that led them to

entails the controversial thesis in philosophy of emotions that some affective states are or include beliefs. See, e.g., Scarantino (2010) for a critical discussion of this thesis.

forgo adopting the belief that P when the issue of whether P is true came up to their consciousness. How is A's knowledge they've just made a mistake going to help them adopt the belief that P if none of A's mental states is poised to trigger the adoption of the belief that P and A possesses some mental states which dispo-
se them to remain neutral toward P (or, worse, to adopt the belief that not-P)?¹⁶

This ends the list of accounts of the downstream effects of the feeling of error that we wanted to examine. These accounts are arguably the most promising, and we have found them all wanting for the purpose of showing that AMF-Feeling-of-error constitutes the basis of a reasonable explanation of the motivational force of epistemic judgments. This strongly suggests that AMF-Feeling-of-error isn't a plausible contender in that regard.

What does that mean for the more general proposal that epistemic judgments have *affective* motivational force? There remains one important possibility that we haven't mentioned yet—namely, that epistemic judgments motivate *reactive attitudes* (i.e., emotions such as resentment, indignation, and gratitude) toward other people due to their stated views, arguments and beliefs. However, we will bracket this idea here due to its close affinity to the response to the epistemic motivation problem that we develop in Section 4. According to that response, epistemic judgments have *practical* motivational force in virtue of directly motivating *verbal actions of criticism and praise*. Moreover, we think that what makes reactive attitudes seem plausible as motivational targets for epistemic judgments is precisely the emotions' connection to verbal actions of criticism and praise. The rest of the reactive attitudes' characteristics, such as their specific phenomenal character or their role in producing other types of actions (e.g., the particular body movements associated with moral indignation), seem less relevant in that context. If we are right about this, then whatever considerations militate in favor of claiming reactive attitudes as motivational targets may provide even stronger support for the view that we advance in Section 4.¹⁷

Besides reactive attitudes and the hypothesized feeling of error described above, it is unclear to us what other affective states could be targeted by epis-

16. Some may be tempted to answer this question by saying that, for as long as A upholds a first-personal judgment of the form *For any proposition P, I ought to believe P if conditions C are met*, they will possess a corresponding belief with the content *For any proposition P such that conditions C are met, it is (probably) the case that P*. But this move is a non-starter. It raises a number of deeply problematic issues about the relationship between the first-personal epistemic judgment and the corresponding belief. Just to give one example: If the latter state is acquired and abandoned exactly at the same time as the former state, it means that they have an extremely similar dispositional profile. How can it be that one is a straightforward belief and the other is conative state?

17. Could epistemic expressivists repurpose the view that we put forward in Section 4 by conceiving of reactive attitudes as variables that mediate the connection between the attitude of epistemic endorsement and the relevant types of verbal actions that we describe there? Absolutely. And we don't have any problems with that strategy—that is, other than the fact that it faces a similar challenge to the one that we raise in Section 5.

temic judgments and what role these affective states could play in our mental economy. So, barring responses along the lines of the one we propose in Section 4, the foregoing suggests that epistemic expressivists can't hold that epistemic judgments have *affective* motivational force.

3.5. *Epistemic Judgments and Conative States*

Turn to the view that epistemic judgments motivate *conative* states. We want to highlight here that there is a straightforward reason why epistemic expressivists can't appeal to it in order to address the challenge of characterizing the motivational force of epistemic judgments. The reason is that adopting this view just pushes the challenge one level up without providing a new angle on the underlying issues. As we noted at the beginning of Section 3, motivational force is an essential characteristic of conative states. So, the suggestion that epistemic judgments motivate a subsequent conative state raises the question of what the motivational force of *that* state is. Here, we have similar options as those described above: The subsequent conative state could motivate doxastic states, actions, habits and related entities, affective states, conative states, or a combination of the previous options. And similar issues to the ones discussed in the previous sections, as well as in Section 3.6, will arise for each of these options.

3.6. *A Combination of the Above*

The last possibility that we will discuss is that epistemic judgments motivate a *combination* of the elements mentioned above. There are two different ways of developing this option. On the one hand, epistemic expressivists could claim that different types of epistemic judgments have distinct motivational targets. They could say, for instance, that *first-personal* epistemic judgments motivate doxastic states, whereas *third-personal* epistemic judgments motivate actions. On the other hand, epistemic expressivists could maintain that epistemic judgments of the very same type (e.g., first-personal epistemic judgments) motivate at once a number of the elements surveyed in the previous sections. In this section, we will argue that neither version of this option provides additional purchase to adequately explain the motivational force of epistemic judgments.

Consider the first approach. Take two types of epistemic judgments that have different motivational targets. Call the two types of judgments T_1 and T_2 respectively. Now focus on T_1 . Proponents of this approach must say what is T_1 's motivational target as well as what type of motivational relation T_1 has with the target (i.e., strict motivation or loose motivation). But then the issues we raised for

the five types of targets discussed in previous sections (doxastic states, actions, habits and related entities, affective states, and conative states) will also arise in the case of T_1 . After all, assuming that T_2 has a different motivational target doesn't resolve any of these issues for T_1 . Hence, appealing to the first approach doesn't bring any new tools to explain the motivational force of T_1 . Ditto for T_2 . And so, it follows that this approach doesn't provide a new angle on the challenge that occupies us.

Let us turn to the second approach. It is committed to the idea that epistemic judgments of at least one type — call it type T — motivate two motivational targets at once. Call these targets MT_1 and MT_2 respectively. To satisfactorily describe the motivational force of T , proponents of this approach must provide an account of the motivational force of T with regard to MT_1 as well as an account of the motivational force of T with regard to MT_2 . This means that they must explain, among other things, what type of motivational target MT_1 is and what type of motivational relation judgments of type T have with MT_1 . But, again, the issues we raised for the five previous options will arise in this context too. The assumption that T also motivates another type of target, namely MT_2 , won't mitigate them in any way. Therefore, it follows that this second approach also fails to offer the required purchase.

It is worth considering an example. Ridge (2018) develops a theory of the motivational force of a specific subset of epistemic judgements that happens to embody this second approach. The theory presupposes the following claim associated with the idea of *pragmatic encroachment* in epistemology:

[W]hen the costs of having a false belief are very high and the costs of suspending judgment until more evidence is gathered are very low, the level of epistemic justification one must have to count as 'sufficient' [to be warranted in believing that P] is plausibly higher than when the costs of having a false belief are lower or the costs of gathering more evidence are very high. (Ridge 2018: 145)

From this, Ridge argues that at least one kind of epistemic judgment has motivational force, namely judgments about what someone has *sufficient epistemic reason to believe*. This is because the above claim strongly indicates that "these epistemic judgments have as part of their content that there is sufficient practical reason for someone to form a belief *rather than* gather more evidence or whatever" (Ridge 2018: 147, his emphasis). To judge that there is sufficient epistemic reason to believe that P , in other words, is in part to judge that there is sufficient practical reason to *form a view about whether P rather than continuing to inquire*. Since, by hypothesis, practical normative judgments like the latter are motivating, judgments about sufficient epistemic reasons are also motivating.

We are happy to grant that pragmatic factors may make a difference as to whether a subject is aptly (or normally) described as having sufficient epistemic reason to believe *P*. However, we don't think that Ridge's proposal can be of use in arguing that epistemic judgments have motivational force in a way that avoids the issues raised above. And that's because it implements the second approach just discussed. On Ridge's theory, judgments about what there is sufficient reason to believe motivate us to *form a belief one way or the other* as well as *to stop inquiring*. In other words, Ridge's proposal amounts to a form of the doxastic motivation view assessed in Section 3.1 in combination with a new version of PMF-Inquiry-termination applied to sufficient-epistemic-reason judgments. Now, Ridge pleads, on behalf of his theory, that this version of PMF-Inquiry-termination avoids the objection that we outlined in Section 3.2 against the original version of PMF-Inquiry-termination applied to knowledge attributions, and we are tempted to agree. But the problem is that his proposal still faces the same issues that we outlined in Section 3.1 about the doxastic motivation view.

This concludes our presentation of the epistemic motivation problem. Recall that it is based on two claims: that epistemic expressivists face an explanatory challenge to provide an account of the motivational force of epistemic judgments, and that they cannot adequately tackle that challenge. Though the case for the first claim was straightforward, the argument for the second claim required a longer discussion. The argument involved a detailed analysis in favor of rejecting the six most promising options for the motivational targets of epistemic judgments. Taken together, the considerations reviewed in Section 3 thus threaten to show that epistemic expressivism is untenable.

4. A Verbal-Action Account

We will now provide what we think is the best response to the epistemic motivation problem. It seeks to undermine the second claim on which the problem is based by putting forward a novel account of the motivational force of the state of epistemic endorsement. More specifically, the response involves arguing that epistemic endorsement motivates certain types of *actions*. We've already seen one proposal along those lines in Section 3.2: Kappel and Moeller (2014) argue that judging that *S* knows that *P* motivates the attributor to *terminate inquiry with respect to P*. We also saw there that there are good reasons to reject their proposal. The problems faced by Kappel and Moeller's proposal have even led Boulton and Köhler (2020) to conclude that it is unlikely that epistemic judgment could motivate *any type* of action.

But we suspect that this is too hasty. We believe that there is a rather natural suggestion for the practical motivational force of epistemic endorse-

ment if we look at seminal work on moral expressivism. Many expressivist theories of epistemic discourse (e.g., Chrisman 2007; Field 2009) take the state of epistemic endorsement to have as its object a set of *epistemic norms*, that is, norms that guide belief revision. With this in mind, perhaps epistemic endorsement strictly motivates specific kinds of *verbal actions*, such as verbally naming the epistemic norms that the agent endorses as well as publicly criticizing (or praising) other people's arguments based on whether these arguments clash (or conform) with the agent's norms. This proposal mirrors some of the main suggestions of Gibbard (1990) about the role of verbal actions in accounting for the motivational force of moral judgment. Here is a representative passage:

The state of accepting a norm, in short, is identified by its place in a syndrome of tendencies toward action and *avowal*—a syndrome produced by the *language-infused* system of coordination peculiar to human beings. The system works through discussion of absent situations, and it allows for the delicate adjustments of coordination that human social life requires. The syndrome that manifests accepting a norm takes in normative discussion and normative governance. In this normative discussion, in unrestrained contexts, *one tends to avow the norm*. One tends to be influenced by the avowals of others, and to be responsive to their demands for consistency. Normative governance by the norm is a tendency to conform to it. Accepting a norm is whatever psychic state, if any, gives rise to *this syndrome of avowal of the norm* and governance by it. (Gibbard 1990: 75, our emphasis)

Our main interest here concerns the references to *avowal* and *normative discussion*. By 'discussion of absent situations', Gibbard means discussions that pertain to past or hypothetical scenarios and that aim to elicit people's judgment about the actions performed in these scenarios. He clarifies his notion of avowal as follows:

By "avowal" here, I mean to include a wide range of expressions we might count as taking a position in normative discussion—in the discussion of absent states of affairs. The simplest kind might be evincing an emotion toward the absent situation. Other kinds of avowal are more explicit: we may express a hypothetical decision in words, or label an action in words that are emotively charged. To understand acceptance of norms, we need to look to such avowal—the kind of avowal from which consensus can emerge, and which may persist even when there is no approach to consensus. (Gibbard 1990: 73)

Both quotes highlight that, on Gibbard's (1990) account, accepting a moral norm has motivational force related to a variety of verbal actions—actions such as stating the norm itself as well as publicly criticizing (or praising) people's actions in various past or hypothetical scenarios based on whether those actions clash (or conform) with the norm.

At first sight, this seems to offer a promising model for understanding the practical motivational force of epistemic endorsement. Suppose that an agent possesses a state of epistemic endorsement whose object is a certain set *N* of epistemic norms. Suppose furthermore that the agent witnesses some speaker *S* making a public argument. The idea under consideration is that this state of epistemic endorsement strictly motivates, at least to some extent, the agent to act as follows: (i) to verbally state specific elements of *N* if the agent thinks it is relevant to the assessment of *S*'s argument; (ii) to verbally criticize *S*'s utterance if the argument violates *N*; (iii) to verbally praise *S*'s utterance if the argument conforms to *N*. Verbal criticism or praise could come in a variety of formulations, but some obvious possibilities include mere rejections (e.g., "What you said is false"), mere acceptance (e.g., "I agree with everything she said"), critical analyses of various length (e.g., "So, your whole case rests on Jack's testimony? But we all know that Jack is a pathological liar!"), claims to the effect that a given argument violates an element of *N* (e.g., "That can't be right, you're flouting the principle of noncontradiction!").

There are two strong considerations in favor of adopting this account of the motivational force of epistemic endorsement. First, humans across a wide variety of cultures (if not all cultures) are disposed to engage in these sorts of verbal actions (Mercier 2011). So, we already need to posit motivational states to specifically explain where these actions come from. The account thereby also avoids a major drawback of Kappel and Moeller's proposal: that of basing their entire explanation of the motivational force of epistemic judgments on actions of a kind whose existence is highly debatable.

Second, assuming that the type of expressivist account of moral discourse developed by Gibbard (1990) is on the right track, the account developed here suggests a substantial parallel between the conative attitudes expressed by moral claims and those expressed by epistemic claims. It suggests that both types of conative attitudes have strict motivational force tied to *verbal actions*. Epistemic expressivists could appeal to this parallel to maintain that the psychological mechanisms driving epistemic discourse are of the same kind as the mechanisms driving moral discourse or that one general set of psychological mechanisms for normative discussion drives both types of discourses. It would thus allow epistemic expressivists who are sympathetic to moral expressivism (which is all of them, to our knowledge) to provide a unified picture of their commitments in philosophy of mind.

Overall, these two considerations should make our response look like a promising solution to the epistemic motivation problem. However, it faces an important challenge, to which we turn now.

5. An Externalist Challenge to Our Verbal-Action Account

We developed in the previous section what we think is the most promising response to the epistemic motivation problem. In this section, we want to highlight that the response raises an important challenge of its own, a challenge that pertains to the very psychology of epistemic criticism and praise.

The challenge that we have in mind comes into view when we start asking why we should posit a mental state whose function is to motivate verbal actions of the precise types identified in Section 4. We have already noted above that there is strong evidence (Mercier 2011) that we possess dispositions to perform the verbal actions of the relevant types. However, it is not clear that the existence of these dispositions is best explained *by assuming the existence of a state whose role is to motivate such specific types of actions*. The dispositions in question could very well arise due to other types of mental states, such as general desires to communicate what we think about other people's utterances to them and others. For instance, suppose that we have a general desire to criticize (or praise) people depending on whether what they say clashes (or coheres) with what we are disposed to say about a certain event or scenario. Such a desire, when combined with a belief that someone has said something that clashes (or coheres) with what we would have said about a certain event or scenario, would lead to the formation of an intention to verbally criticize (or praise) the person's utterances.

Why would anyone posit desires like that? Because they make sense of the general argumentative nature of human communication. Any human agent is disposed to verbally criticize or praise other people's utterances based on their compatibility with the agent's own beliefs, desires, tastes, values, etc.—and not just in virtue of violating high-level precepts discussed in philosophy under the rubric of 'epistemic norms' (high-level precepts like "Form beliefs in a way that satisfy *modus ponens*"). For instance, Julie might criticize one of Michael's arguments if that argument is incompatible with her belief that it rained this morning or that Mary ate an apple two hours ago. It doesn't matter that these beliefs are utterly peripheral to Julie's worldview. If she thinks Michael's argument is worth criticizing, she may well bring up the fact that it is incompatible with these otherwise inconsequential beliefs. More generally, various theoretical and empirical considerations suggest that humans have cognitive mechanisms that specifically monitor for accuracy in communication and that motivate people to

argue with someone with whom they disagree in a large variety of contexts (see, e.g., Mercier & Sperber 2011 for a thorough review).

It follows that anyone who chooses our response to the epistemic motivation problem faces an important challenge: to explain why the postulation of a state whose function is to motivate verbal actions of the types identified in Section 4 provides a theoretical payoff over and above positing the type of general desires for accuracy in communication just described. We are not in a position to show that this project *can't* be done, but it strikes us as difficult.

Part of what makes the challenge pressing is that, of all the potential expressivist and cognitivist theories of epistemic discourse, only expressivist theories committed to our response (or something much like it, see fn. 17) are liable to the challenge. That is so for two reasons. First, the class of claims whose utterances plausibly count as *verbal actions of epistemic criticism or praise* is much larger than the class of claims that are the focus of the debate between epistemic expressivists and epistemic cognitivists, the claims that we called 'epistemic claims' in Section 2. The former class includes claims like "What you said is false", which do not count as *epistemic claims* for the purposes of that debate. So, our response notwithstanding, expressivists and cognitivists don't need to take any stance on the type of mental state that motivates the utterances of claims of the former class. Of course, statements belonging to the latter class are sometimes used to criticize or praise others' epistemic behavior (e.g., "You don't *know* this"). But this is where the second reason becomes relevant.

Second, expressivists and cognitivists don't have to say that the mental state(s) *expressed* by an epistemic claim are exactly the same as those that *motivate* the claim. To take a relevant example: Epistemic cognitivists are committed to the view that epistemic claims express straightforward beliefs. Yet, just possessing a belief isn't enough to lead someone to produce an utterance stating the content of that belief. Moreover, it is not the case that people have cognitive processes that lead them to immediately state the content of every belief that becomes occurrent or conscious (Carruthers 2018). So, there has to be a specific conative state that drives some agent A to say something about how bad (or good) they believe someone else's epistemic behavior to be in a specific context, even if A already has beliefs about that behavior. Epistemic cognitivists should therefore be more than happy to fill out the motivational gap here by appealing to general desires for accuracy in communication.

It should be clear at this point that the challenge presented in this section mirrors some of the work of externalists about moral motivation (e.g., Railton 1986; Brink 1989; Svavarsdottir 1999). *Externalism about moral motivation* is the view that the reliable connection between moral judgments and actions that conform to these judgments is due to contingent, conative mental states (like a desire to do the morally right action or a desire to be a moral person) which are

distinct from moral judgments themselves. Similarly, the challenge we develop here could be conceived as based on a form of *externalism about motivation for epistemic criticism and praise*. On this view, the reliable connection between epistemic judgments and verbal utterances of epistemic criticism and praise is due to contingent, conative mental states (like a general desire to criticize people depending on whether what they say clashes with what we are disposed to say about a certain event or scenario) which are distinct from the epistemic judgments themselves.

Where does this leave us? In Section 3, we argued against the possibility for epistemic expressivists of endorsing any of the extant accounts of the motivational force of epistemic judgments from the metaepistemological literature. Therefore, insofar as the externalist challenge that we raise here to our verbal-action account is unanswerable, this may suggest that there is no adequate solution to the problem. As we have said already, we are not sure whether the challenge can be met or not. But if it cannot be met, then our work would support rejecting epistemic expressivism altogether.

6. Conclusion

We started this paper by formulating a new problem for epistemic expressivism based on two claims. The first is that it is a constraint on any expressivist theory of epistemic discourse—be it pure or ecumenical—that it be able to provide an account of the motivational force of epistemic judgments. The second claim is that no such theory can satisfy the constraint. The combination of both claims gave rise to what we called *the epistemic motivation problem*. We then developed what we think is the best response to our problem, which involves claiming that the state of epistemic endorsement motivates certain kinds of verbal actions, namely acts of verbally criticizing and praising (through public uses of natural language) the arguments presented by other people. We finally argued that this response may not be a genuine solution unless a further externalist challenge about the nature of the motivation involved in producing these verbal actions is addressed.

Due to the inadequacies of previous proposals for the motivational force of epistemic judgment (see Section 3), a failure of our proposed response may suggest that epistemic expressivism cannot escape the epistemic motivation problem. If so, it will turn out that one cannot give an expressivist account of epistemic discourse at all. If not, then our analysis will provide a promising way to develop and defend epistemic expressivism's commitments in philosophy of mind and philosophy of action.

Acknowledgments

We thank the Editor and the two anonymous reviewers for comments that have led to substantial additions and improvements throughout the paper. Alexandre Duval gratefully acknowledges support for his initial research on the paper from the Social Sciences and Humanities Research Council of Canada (grant 756-2020-0385) and the philosophy department of the University of British Columbia. He also received financial support from the Templeton World Charity Foundation (grant TWCF0539) while working on the final version of the manuscript. Charles Côté-Bouchard thanks the Social Sciences and Humanities Research Council of Canada (grant 756-2019-0366) and the philosophy department of the Université de Montréal.

References

- Ahlstrom-Vij, Kristoffer (2013). Moderate Epistemic Expressivism. *Philosophical Studies*, 163(2), 337–57.
- Alston, William P. (1988). The Deontological Conception of Epistemic Justification. *Philosophical Perspectives*, 2, 257–99.
- Beddor, Robert (2016). *Reduction in Epistemology*. PhD dissertation, Rutgers University.
- Blackburn, Simon (1996). Securing the Nots: Moral Epistemology for the Quasi-Realist. In Walter Sinnott-Armstrong and Mark Timmons (Eds.), *Moral Knowledge? New Readings in Moral Epistemology* (82–100). Oxford University Press.
- Blackburn, Simon (2001). *Ruling Passions: A Theory of Practical Reasoning*. Clarendon Press.
- Boult, Cameron and Sebastian Köhler (2020). Epistemic Judgement and Motivation. *The Philosophical Quarterly*, 70(281), 738–58.
- Brink, David O. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge University Press.
- Carruthers, Peter (2018). Implicit Versus Explicit Attitudes: Differing Manifestations of the Same Representational Structures? *Review of Philosophy and Psychology*, 9, 51–72.
- Chrisman, Matthew (2007). From Epistemic Contextualism to Epistemic Expressivism. *Philosophical Studies*, 135(2), 225–54.
- Chrisman, Matthew (2012). Epistemic Expressivism. *Philosophy Compass*, 7(2), 118–26.
- Chuard, Phillipe and Nicholas Southwood (2009). Epistemic Norms Without Voluntary Control. *Noûs*, 43(4), 599–632.
- Cuneo, Terence (2007). *The Normative Web: An Argument for Moral Realism*. Oxford University Press.
- Dreier, James (2004). Meta-Ethics and the Problem of Creeping Minimalism. *Philosophical Perspectives*, 18(1), 23–44.
- Dreier, James (2015). Another World: The Metaethics and Metametaethics of Reasons Fundamentalism. In Robert N. Johnson and Michael Smith (Eds.), *Passions and Projections: Themes from the Philosophy of Simon Blackburn* (155–71). Oxford University Press.

- Field, Hartry (2009). Epistemology Without Metaphysics. *Philosophical Studies*, 143(2), 249–90.
- Field, Hartry (2018). Epistemology from an Evaluativist Perspective. *Philosophers' Imprint*, 18(12), 1–23.
- Gibbard, Allan (1990). *Wise Choices, Apt Feelings*. Harvard University Press.
- Gibbard, Allan (2003). *Thinking How to Live*. Harvard University Press.
- Gibbard, Allan (2012). *Meaning and Normativity*. Oxford University Press.
- Kappel, Klemens (2010). Expressivism about Knowledge and the Value of Knowledge. *Acta Analytica*, 25(2), 175–94.
- Kappel, Klemens (2011). Is Epistemic Expressivism Dialectically Incoherent? *Dialectica*, 65(1), 49–69.
- Kappel, Klemens and Emil F. L. Moeller (2014). Epistemic Expressivism and the Argument from Motivation. *Synthese*, 191(7), 1529–47.
- Kvanvig, Jonathan L. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge University Press.
- Kyriacou, Christos (2010). *The Epistemic Justification Puzzle*. PhD dissertation, University of Edinburgh.
- Kyriacou, Christos (2012). Habits-Expressivism about Epistemic Justification. *Philosophical Papers*, 41(2), 209–37.
- Lynch, Michael P. (2009). Truth, Value and Epistemic Expressivism. *Philosophy and Phenomenological Research*, 79(1), 76–97.
- Mendelovici, Angela (2018). *The Phenomenal Basis of Intentionality*. Oxford University Press.
- Mercier, Hugo (2011). On the Universality of Argumentative Reasoning. *Journal of Cognition and Culture*, 11(1–2), 85–113.
- Mercier, Hugo and Dan Sperber (2011). Why Do Humans Reason? Arguments for an Argumentative Theory. *Behavioral and Brain Sciences*, 34(2), 57–74.
- Mitova, Veli (2011). Epistemic Motivation: Towards a Metaethics of Belief. In Andrew Reisner and Asbjørn Steglich-Petersen (Eds.), *Reasons for Belief* (54–74). Cambridge University Press.
- Peacocke, Christopher (1992). *A Study of Concepts*. MIT Press.
- Railton, Peter (1986). Moral Realism. *The Philosophical Review*, 95(2), 163–207.
- Ridge, Michael (2007). Epistemology for Ecumenical Expressivists. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 81, 83–108.
- Ridge, Michael (2018). How to Be an Epistemic Expressivist. In Conor McHugh, Jonathan Way, and Daniel Whiting (Eds.), *Metaepistemology* (141–58). Oxford University Press.
- Scarantino, Andrea (2010). Insights and Blindspots of the Cognitivist Theory of Emotions. *British Journal for the Philosophy of Science*, 61(4), 729–68.
- Schafer, Karl (2014). Doxastic Planning and Epistemic Internalism. *Synthese*, 191(12), 2571–91.
- Schroeder, Mark (2010). *Noncognitivism in Ethics*. Routledge.
- Schwitzgebel, Eric (2002). A Phenomenal, Dispositional Account of Belief. *Noûs*, 36(2), 249–75.
- Sinclair, Neil (2007). Expressivism and the Practicality of Moral Convictions. *The Journal of Value Inquiry*, 41(2–4), 201–20.
- Svavarsdóttir, Sigrún (1999). Moral Cognitivism and Motivation. *The Philosophical Review*, 108(2), 161–219.