

Epistemological Disjunctivism: Perception, Expression, and Self-Knowledge

Dorit Bar-On

Drew Johnson

University of Connecticut

Abstract

So-called basic self-knowledge (ordinary knowledge of one's present states of mind) can be seen as both 'baseless' and privileged. The spontaneous self-beliefs we have when we avow our states of mind do not appear to be formed on any particular epistemic basis (whether intro- or extro- spective). Nonetheless, on some views, these self-beliefs constitute instances of (privileged) knowledge. We are here interested in views on which true mental self-beliefs have internalist epistemic warrant that false ones lack. Such views are committed to a form of *disjunctivism* about basic self-knowledge. We begin by presenting an influential disjunctivist view about perceptual knowledge (Pritchard 2008, 2012, and elsewhere) and articulate a problem for it. We then consider two versions of disjunctivism about basic self-knowledge – one 'constitutivist', the other 'neo-expressivist' – and argue that both can avoid an analogue of this problem for self-knowledge. However,⁵ we give reasons for preferring the disjunctivism yielded by neo-expressivism. We conclude by considering briefly whether an acceptable disjunctivism about mental self-beliefs can point the way toward a sensible disjunctivism about perceptual beliefs.

1. Introduction

We seem to have privileged basic knowledge of our own present states of mind. As I say (out loud or to myself): "What a terrible headache!", or "I'd love a cup of tea!", I know as no one else does that I have a headache or would like a cup of tea. As a subject of mental states, I seem to be in a better position than anyone else to know how I feel or want or what's on my mind at a given moment. And I seem to enjoy this privileged epistemic position despite not having any justification (in the traditional, internalist sense, including inference, evidence, or observation) for the relevant self-beliefs. We are here interested in views of basic self-knowledge according to which, although basic self-beliefs are indeed not arrived at on any particular epistemic basis, they are nevertheless – in a sense that will be made clear – epistemically warranted; what provides the warrant for an individual's basic mental self-belief

is the very mental state the belief is about.¹ (So, for example, my self-belief that I'm feeling anxious right now is not only *made true* by my presently being in a state of feeling anxious; my belief is also epistemically warranted by my being in that state.) Indeed, it is this that enables my self-belief to represent an instance of knowledge.²

On this view of self-knowledge, *false* self-beliefs are not only false; they are also epistemically unwarranted, because there are no first-order mental states to make them either true *or* warranted. Views that follow this approach are committed to a form of *epistemological disjunctivism* about self-knowledge, because they entail that (internalist) epistemic warrant can vary between a true self-belief and its corresponding false self-belief.

Epistemological disjunctivism is usually discussed in the context of perceptual knowledge. According to one such account (Pritchard 2012), in paradigmatic cases of perceptual knowledge, S's perceptual belief that p is warranted by S's *seeing that* p (in the case of visual perception), where 'seeing that' is factive (that is, 'seeing that p' entails p), and where S has reflective access to the warrant for her belief. When S's perceptual belief is *false*, S does not see that p (because that would entail p), but only *seems* to see that p. Thus, both epistemological disjunctivism about perceptual knowledge, and the approach to self-knowledge mentioned above, hold that S's veridical belief that p has a different epistemic warrant from the warrant it would have had if p had been false.

¹ For simplicity, we will here be assuming a standard analysis of knowledge, according to which knowledge requires true belief that is warranted. For reasons abundantly discussed in the literature, knowledge may require something more. This should not matter for our concerns here. We set aside 'knowledge first' views such as Williamson 2000.

² In section 3.2 we articulate more thoroughly in what sense self-beliefs can count as warranted on these accounts, by elucidating a notion of epistemic *grounding*, distinct from epistemic *justification*.

In Section 2, we briefly contrast *metaphysical* disjunctivism about perceptual states (MDP) with *epistemological* disjunctivism about perceptual knowledge (EDP). We go on to present what we take to be a serious challenge to metaphysical disjunctivism about perception, due to Tyler Burge (2005; 2011). In Section 3, we offer some general comparisons between basic self-knowledge and perceptual knowledge, and discuss the sort of warrant appropriate for basic mental self-belief. In section 4, we consider a version of epistemological disjunctivism about basic self-knowledge (EDSK) that is associated with *constitutivist* views of self-knowledge and argue that it does not face the same challenge faced by EDP. However, there are difficulties with constitutivism, which speaks in favor of considering an alternative. In Section 5, we consider an alternative, *neo-expressivist* account that – we argue – has several advantages over constitutivism when it comes to articulating an acceptable version of EDSK. We conclude (in Section 6) by briefly considering whether an acceptable disjunctivism about mental self-beliefs can point the way toward a sensible disjunctivism about perceptual beliefs.³

2. Metaphysical and Epistemological Disjunctivism about Perception

It is standard, in discussions of disjunctivist accounts of perception and perceptual knowledge, to distinguish the following two cases. (For simplicity, we will focus primarily on *visual* perception.)

The Good Case: S seems to see an orange before her. S's sense organs and perceptual system are functioning properly, and the lighting and other environmental conditions conducive to accurate vision. There is an orange before S, and S does in fact see the

³ Throughout, unless otherwise noted, when we refer to self-knowledge/beliefs we are concerned with *basic* knowledge/beliefs concerning one's present states of mind.

orange. And so, if S has a perceptual belief that there's an orange in front of her, that belief will be true.

The Bad Case: S seems to see an orange before her. S's sense organs and perceptual system are functioning properly. However, S is undergoing a perceptual referential illusion where there is no orange (suppose there is a hologram of an orange instead, for instance), or is perceiving some other object that S cannot perceptually distinguish from an orange (e.g. a red grapefruit from a distance). There is no orange before S; S does not in fact see an orange.⁴ And so, if S has a perceptual belief that there's an orange in front of her, that belief will be false.

According to *metaphysical disjunctivism* about perception (MDP), as between a veridical perception of an orange, and a perceptual experience as of an orange caused by, say, a hologram of an orange, *there is no metaphysically relevant kind of perceptual experience in common*. Whatever commonalities there are between the experiences, they are not to be taken as underwritten by a single type of perceptual experience that the subject undergoes in the good and bad case.⁵

2.1 Burge's Objection to Disjunctivism

⁴ Typical formulations of the bad case appeal to hallucinations or illusions, without distinguishing the cause of illusion, sometimes alluding to evil demons, or brain-in-a-vat scenarios, or else not describing in much detail what sort of hallucination or illusion is under consideration. Our construal of the bad case is designed to accommodate Burge's reservation about not counting certain kinds of hallucination (such as hallucinations caused by drugs or brain-ticklings) as cases of perception (see Burge, 2005, 42).

⁵ Here is a representative statement of MDP:

"[T]he basic claim of disjunctivism can be put as follows: the experiences in the good case and the hallucinatory bad cases share no mental core, that is, there is no (experiential) mental kind that characterizes both cases." (Byrne and Logue, 2009, ix).

Tyler Burge (2005, 2011) has argued against disjunctivism about perception, which he takes to be the view that “that there is never an explanatorily relevant mental state type in common between (and specific to) a veridical perception and a referential perceptual illusion . . . [a]nd ... there is never a mental state type in common between (and specific to) perception of an object and perception of a would-be duplicate substitute for the object that would be, in the context, perceptually indiscernible to the perceiver” (Burge 2005, 25).⁶ Burge’s target seems to be *metaphysical* disjunctivism about perception (MDP). And he goes on to argue that MDP is incompatible with what the science of perception tells us about perceptual experience. The relevant point that Burge takes the science to support is what he calls the Proximity Principle:

The Proximity Principle: *“Holding constant the antecedent psychological set of the perceiver, a given type of proximal stimulation (over the whole body), together with associated internal afferent and efferent input into the perceptual system, will produce a given type of perceptual state, assuming that there is no malfunctioning in the system and no interference with the system. On any given occasion, given the total antecedent psychological state of the individual and system, the total proximal input together with internal input into the system suffices to produce a given type of perceptual state, assuming no malfunctioning or interference.”* (Burge 2005, 22, emphasis in original)

Consider again the good and the bad case. Assuming the conditions of the Proximity Principle are met, the Proximity Principle tells us that the perceptual state kind in the good

⁶ As Burge (2005, 42) notes, it may be mistaken for the metaphysical disjunctivist to regard all cases of illusion and hallucination as genuine perception. Their focus should be on cases that are genuinely perceptual. (See footnote 4.) But when it comes to epistemological disjunctivism, what matters is that experiences in good and bad cases are subjectively indistinguishable, whether or not those experiences are genuine perceptions.

case is the same as in the bad case. So disjunctivism is committed to accepting, while perceptual psychology denies, “that types of perceptual state can vary even when there is no change in proximal stimulation, internal input, and antecedent psychological states remain the same” (Burge, 2005, p. 22).⁷

If, as Burge claims, MDP is indeed incompatible with well-established empirical research in perceptual psychology, then MDP should be rejected.⁸ Note, however, that, even if one acknowledges a metaphysical commonality in perceptual experience between the good and the bad case, it remains to be seen what – if any – *epistemological* significance should be assigned to this commonality. Indeed, some self-proclaimed disjunctivists set aside metaphysical disjunctivism and instead focus directly on epistemological disjunctivism, which is specifically concerned with how the epistemic status of perceptual belief might vary between good and bad cases. For example, in articulating his preferred (McDowell-inspired) form of disjunctivism, Duncan Pritchard makes clear from the outset that his is a form of EDP, which he takes to be independent of MDP. The core thesis he puts forward is this:

EDP: “In paradigmatic cases of perceptual knowledge an agent, S, has perceptual knowledge that ϕ in virtue of being in possession of rational support, R, for her belief that ϕ which is both *factive* (i.e., *R*’s

⁷ Ram Neta has objected that, whether the Proximity Principle makes trouble for MDP depends on how the notion of proximity is understood. One would need to give sense to “proximal” that renders the Principle both plausible and inconsistent with MDP – something Burge does not seem to provide, and which we are not able to provide here, either.

⁸ Thus, proponents of the view would have to show that perceptual psychology is somehow not relevant to the evaluation of MDP. Perhaps they could argue, for instance, that perceptual psychology concerns states attributable to only to perceptual subsystems, whereas MDP concerns perceptual states attributable to (whole) perceivers. (See, e.g., McDowell 2010.) Burge disagrees. He writes: The theory of vision he describes “attributes states that are recognizably perceptual and recognizably states of individuals, not merely of subsystems” (2005, 22).

obtaining entails ϕ) and *reflectively accessible* to S" (Pritchard 2012, 13).

In the good case, according to (this version of) EDP, S can *see that* p. 'Seeing that' is factive, so S's *seeing that* p entails p. Moreover, disjunctivism adds, in the good case, S has reflective access to this factive reason, and it is in virtue of this reflective access that S's belief is justified in a way conducive to knowledge. When I believe there is an orange before me in the good case, I can know by reflection alone that what justifies my belief is that *I see* that there is an orange before me.

Pritchard maintains that EDP represents the commonsense view of perceptual knowledge, remarking that when one reports a perceptual belief, and someone challenges this belief ("How do you know there is an orange before you?"), a natural response to this challenge is to cite one's factive reason ("because I can see that there is an orange before me") (Pritchard 2012, 17-18). Now, it's true that one would offer the *same* reason in the bad case as well, though in that case one's reason won't be factive (since there's no orange to be seen). But the point is that, even in bad cases, we advert to what *would be* factive reasons in support of our knowledge claim. We do not offer as reasons what might in fact be common to both the good and the bad cases, namely, its *seeming* to us that we see that there is an orange.

2.2 EDP's Relation to MDP

If Burge is right that MDP is incompatible with empirical research in perceptual psychology, then, if it turned out that EDP entails MDP, EDP would have to be rejected too. Yet it seems that one could accept both EDP and the denial of MDP.⁹ At the same time, as Pritchard

⁹ Pritchard concurs: "It is reasonably clear that epistemological disjunctivism does not in itself entail metaphysical disjunctivism. For that the rational standing available to the agent in the normal veridical perceptual experiences and corresponding (introspectively indistinguishable) cases of illusion and

himself observes, it is difficult to see how EDP could be plausible *without* MDP (2012, 24). The difficulty in accepting EDP while denying MDP stems from the following lingering intuition. Suppose we hold fixed S's psychological states, the proximal stimulations of S's sense organs, etc., and *only* vary the distal causes of the proximal stimulations (thereby affecting whether S's resulting perceptual state is veridical or not). Then there seems to be no principled way for the epistemological disjunctivist to explain *in virtue of what* the warrant for the perceptual belief could vary between the good and the bad case. Accepting MDP makes EDP much more plausible, because then one can say, quite naturally, that the difference in the warrant for belief in the two sorts of case supervenes on differences in the nature of the perceptual state in the two sorts of cases.¹⁰

In the case of perceptual knowledge, it seems clear that one's perceptual experience has an important role to play in supplying the warrant for perceptual belief. As Pritchard recognizes, it is natural to offer the reason that one *sees* that p in support of one's belief that p, as opposed to simply offering p itself as the reason. Indeed, it seems that the involvement of one's perceptual experience is crucial for the relevant belief being *perceptual* in the first place. This suggests to us the following principle:

hallucination are radically different does not in itself entail that there is no common metaphysical essences to the experiences of the agent in both cases" (2012, p. 24).

¹⁰ One might think that the difference in the warrant S has for her perceptual belief in the good and bad cases is explained in terms of the reasons available to S in those cases, where what reason S has in turn depends on the relation between S's mental states and the world. So, in the good case, S's reason is a good one *because* S's mental states (including her perceptual experience) is related to the world in the right way, whereas in the bad case, S's mental states (which are of the same *type* as in the good case) are not appropriately related to the world. (Thanks to Matthew McGrath for suggesting this possible response.) A worry for this response is that it seems to present an externalist picture of warrant (as the nature of the relation of one's experience to the world is apparently not available for reflective access), whereas EDP (at least of Pritchard's variety) is an internalist view. How could the *internalistically* available reasons one has for belief vary between good and bad cases?

The Perceptual Warrant Principle (PWP): For a perceptual belief B formed on the basis of perceptual experience e, the epistemic warrant for B is wholly constituted by e.

PWP seems like an intuitively plausible principle, not to be rejected without some argument. Now, one might suggest, as an alternative, that the warrant for a perceptual belief B is secured through some *inference* of B from e. That is, one's perceptual beliefs, it might be thought, are based on one's perceptual experiences in the sense of being inferred from them. But this over-intellectualizes perceptual belief. It seems reasonable to credit nonhuman animals and young children with warranted perceptual beliefs, yet it may well not be plausible to credit them with the capacities necessary for making the relevant inferences. Instead, it is more plausible to think (along the lines of PWP) of perceptual experience itself as providing the warrant for perceptual belief. And this is something that epistemological disjunctivism can (and we think should) accept.¹¹ However, if EDP accepts this principle, it must add that perceptual experience in good cases is factive and reflectively accessible, while perceptual experience in bad cases is not factive. And this will likely involve commitment to MDP.¹²

The anti-disjunctivist point here is that it is mysterious how the contribution one's perceptual experience makes to one's warrant for belief could vary when the type of perceptual experience one is in is held fixed. If one accepts the Proximity Principle (see 2.1), and thus rejects MDP, then, if one also accepts PWP (i.e., that the perceptual experience is

¹¹ One might reject the idea that our perceptual experiences exhaust the internalist warrant for perceptual belief; perhaps there is something else, in addition to one's perceptual experience, which can play a warranting role for perceptual belief - for instance, the relation between that experience and states of affairs in the world. But (as mentioned in fn. 9) it is unclear to us how such an additional source of warrant could be something both reflectively accessible to an agent and something that can vary between good and bad cases.

¹²For an alternative argument that Pritchard's EDP is committed to MDP, see Cunningham (2016).

what constitutes the epistemic warrant for the perceptual belief), there is simply no room for pulling the good and the bad cases apart in terms of epistemic warrant.

Our tentative conclusion is that EDP is a problematic view.¹³ In what follows, we aim to articulate a plausible version of epistemological disjunctivism about basic self-knowledge (EDSK). If we are to be successful, we shall have to show that EDSK at least is not susceptible to the challenged we presented for EDP.

3. Epistemological Disjunctivism and Basic Self-Knowledge: Some Preliminaries

Epistemological disjunctivism about basic self-knowledge (EDSK) is the view that the warrant one has for a basic self-belief about a (current) state of mind – e.g. that one has a headache, or would like a cup of tea, or is thinking about the third premise of an argument – varies between veridical and nonveridical cases.¹⁴ In a veridical case, when one’s self-belief that one is in M is true, because one *is* in M, one has knowledge, since the belief is not only true but is also warranted by the state M itself. Being in M in some way provides one with the appropriate epistemic warrant for one’s self-belief. In a non-veridical case, one has no knowledge, not only because one’s self-belief is not true, but also because it is not warranted (or, at least, it lacks the kind of warrant had in the veridical case).

In Section 4 and 5, we consider two specific versions of EDSK: one *constitutivist* the other *neo-expressivist*. We will be arguing that:

¹³ But see Section 6, where we revisit this issue.

¹⁴ By ‘basic’ self-belief, we mean the ordinary, non-reflective beliefs one has about one’s own present mental states – paradigmatically, these are spontaneous beliefs one has when *avowing* (out loud or to oneself) e.g. “I’m so tired”, “I’m fed up with this”, “I’d like to leave now”, etc. Excluded are self-beliefs formed as a result of deliberate investigation or testimony, as when, for instance, one forms a self-belief as a result of discussion with a therapist or a friend’s input.

- (i) These versions of EDSK are not susceptible to the challenge earlier presented for EDP (2.1), so one cannot object to EDSK based on this challenge; and
- (ii) As between the two versions of EDSK – constitutivism and neo-expressivism – neo-expressivism is the more plausible view, insofar as it avoids other problems facing constitutivism.

However, before turning to these versions, some clarifications are in order.

3.1 Perceptual Beliefs vs. Basic Mental Self-Beliefs

We begin by offering some general comparisons between perceptual beliefs and basic mental self-beliefs.

- First, if basic self-knowledge were thought to be simply a species of perceptual knowledge – as per certain views of introspection¹⁵ – then EDSK would be just a special case of EDP.
- However, second, the literature is rife with arguments against the applicability of the perceptual model to basic self-knowledge. For example, in a series of influential articles, Sydney Shoemaker offered a detailed critique of this model, which we will not rehearse here.¹⁶ In any event, there is a clear disanalogy between the perceptual case and the case of basic self-knowledge. It is intuitively not clear what environmental conditions would be relevant to one’s possession of mental self-beliefs in the first place. And, moreover, it is not clear what would be the analogue in the self-knowledge case of the condition that “S’s sense organs and perceptual system are functioning properly”. Since the two versions of

¹⁵ See Fricker (1998), Macdonald (2007), Shoemaker (1986).

¹⁶ Shoemaker (1994).

EDSK we go on to consider both reject the perceptual model, we will simply set it aside here. (For the same reason, we will further assume that *inferential* models of basic self-knowledge should also be rejected.¹⁷)

- Third, views that reject the perceptual model of basic self-knowledge must deny that there is a perceptual experience in common between the good and the bad case of basic self-knowledge, because they hold that there is *no* perceptual experience involved in this type of knowledge at all. The views we go on to consider deny, in addition, that basic mental self-beliefs have any other source of epistemic justification (such as inference). A natural remaining candidate for warrant for basic self-beliefs on these views would seem to be the very mental states the beliefs concern – and this is obviously not something in common between veridical and non-veridical belief.¹⁸ And this would commit them to EDSK.¹⁹
- Yet, fourth, if the perceptual model is rejected, at least one commonsense motivation for EDP will not carry over to the case of self-knowledge. Recall Pritchard’s observation that the reason we would normally offer for a perceptual belief is a factive one: “How do you know that *p*?” – “I *see that p*”. Pritchard takes this to provide some positive support EDP. But if I were to ask you “How do you know you’d like a cup of tea?”, the natural answer would *not*

¹⁷ But see Byrne (2018), Cassam (2014), and Carruthers (2011) for recent defenses.

¹⁸ One might think that the self-beliefs could have the same warrant in virtue of the fact that they have the same content. But this would amount to holding that mental self-beliefs are self-warranting – a proposal we will here set aside.

¹⁹ A non-disjunctivist view that would be fashioned after Burge’s view of perceptual entitlement (see footnote 22) would not be available to them.

be “I *see/feel that* I’d like a cup of tea...”. To the extent that the question merits an answer, the sensible answer would be: “Well, because I *would* like a cup of tea”.²⁰

Still, this last answer itself can point toward a (related) commonsense support for EDSK. For it suggests that the reasons we would ordinarily cite for basic mental self-beliefs are the very facts that would make our self-beliefs true; we do not invoke some self-experience (e.g. “I *seem to be/feel as though I am* in pain”), in contrast with the perceptual case, where we *do* invoke our perceptual experiences. And this, if taken at face value, suggests that one’s belief that one is in M is warranted simply in virtue of one’s *being in* M. But, as this is a factive reason – one cannot have that same reason if one is not in M – EDSK follows.

- Finally, fifth, although we are concerned to explain the special epistemic security of basic self-knowledge, we by no means take mental self-beliefs to be infallible. Indeed, we take it that, at a minimum, a disjunctivist view of basic self-knowledge would need to make room for the possibility of veridical and non-veridical cases of basic self-beliefs that are indistinguishable from the self-believer’s perspective.

3.2 Varieties of Epistemic Warrant

Given the focus of epistemological disjunctivism on the question of warrant for belief, it will be helpful to mark some distinctions among species of epistemic warrant. These distinctions will allow us to clarify how the views we consider in the next two sections explain the distinctive nature of self-knowledge.

²⁰ The analog of this in the perceptual case – viz. “well, p!” – would seem tantamount to saying “I just know!”.

First, it should be noted that EDP (at least on Pritchard and McDowell's versions) adopts an *internalist* notion of warrant for veridical perceptual beliefs. A central feature of Pritchard's EDP is the claim that in paradigm cases of perceptual knowledge, one has *reflective access* to the rational support for one's perceptual belief.²¹ This commitment to reflective access is what qualifies EDP as an internalist view.²² The internalist component of EDP is part of what makes it an interesting view; what is interesting and distinctive about EDP is the combination of these ideas: (i) reasons for perceptual belief are reflectively accessible, *and* (ii) the reason for veridical perceptual beliefs are factive. This combination would be denied by traditional internalism, since that view would contend that reflectively accessible reasons cannot be factive in the case of perception, even if the reasons are sufficient for knowledge. (Pritchard 2012, 38).²³

The relevant point here is that if the views about basic self-knowledge that we go on to consider are to count as forms of epistemological disjunctivism (on the model of Pritchard's EDP), we should expect them to invoke an internalist conception of warrant as well. So in

²¹ Pritchard considers two ways of filling out the reflective access condition (2012, p. 36). First, there is *accessibilism*: "S's internalist epistemic support for believing that ϕ is constituted solely by facts S can know by reflection alone." The second version is *mentalism*: "S's internalist epistemic support for believing that ϕ is constituted solely by S's mental states." We shall restrict our understanding of reflective access to *accessibilism*, though it is worth considering what a disjunctivist account of self-knowledge might look like with a mentalist understanding of the internalist reflective access condition. If, as we go on to consider, self-beliefs about one's mental states are epistemically warranted by those mental states themselves, then a *mentalist* understanding of reflective access is clearly, and perhaps trivially, satisfied. Pritchard also notes, interestingly, that if EDP accepts mentalism, then it may also be committed to MDP (2012, pp. 36-37). Given our earlier argument, this is reason then for EDP *not* to accept the mentalist version of the reflective access condition.

²² As a starting point, we can consider a weak form of epistemic internalism, which Pappas (2014) calls 'Weak AJI': "One has a justified belief that p only if one can become aware by reflection of some essential justifier one then has for p." An essential justifier, here, is a justification for one's belief without which the belief would fail to be justified. Note, moreover, that this form of internalism only requires that one become *aware of* an essential justifier; one need not *recognize* this justifier as an essential justifier.

²³ Pritchard, 2012, advertises his disjunctivism as combining certain features of traditional internalism and traditional externalism. Pritchard and Neta, 2007, also make the case that McDowell's position combines internalist and externalist elements. The internalist conception of warrant that EDP employs is also recognized in Madison's 2010 overview of epistemic internalism.

articulating below the view of basic self-knowledge we prefer – the neo-expressivist view – we will assume a form of internalism that (we hope) should be acceptable to internalist epistemological disjunctivists. According to this form of internalism, to anticipate, for S to be internalistically warranted in believing p, S must be capable of citing, upon reflection, an essential reason R for her belief that p. (The notion of an essential reason here – adapting Pappas’ notion of ‘essential justifier’ (2014) – is that of a reason without which there would be insufficient warrant for belief.)

Second, we would like to reject the epistemological assumption that when a belief is warranted, its warrant must consist in some particular epistemic basis, or method, that an agent employs in order to arrive at the belief. This assumption is what Bar-On calls the Distinct Epistemic Basis requirement, which (as applied to self-knowledge) states: “If I can be said to have privileged knowledge that I am in a certain state of mind, then this knowledge must have some *distinct epistemic basis*; there must be a special epistemic method or route (a special ‘way of knowing’) that I use to obtain this knowledge” (2004, 344). Applied to knowledge generally, the Distinct Epistemic Basis requirement implies that all knowledge involves a special ‘way of knowing’; it would be incorrect to say that one knows something unless there is a way that one knows – there would have to be some distinct method one employs, some ‘evidential work’ one engages in, in arriving at the relevant belief.²⁴ Distinct epistemic bases for belief include: inference, testimony, observation, memory, etc. - in short, methods of forming a belief that

²⁴ Zimmerman also identifies and rejects a similar assumption in addressing a purported puzzle about self-knowledge: “Our first-order beliefs themselves provide grounds for our second-order introspective beliefs. One’s reason for thinking that one believes that p is the very fact that one believes that p. Why haven’t philosophers embraced this simple answer? *Many epistemologists have assumed that we can justify our beliefs only through observation or inference*” (2006, p. 338, emphasis added).

confer justification on that belief. In rejecting the Distinct Epistemic Basis requirement, we open the possibility that some beliefs may count as warranted, even if the believer has not employed any distinct method or done any evidential work. One can be warranted in holding beliefs that one just finds oneself with, without having engaged (now or at any point in the past) in some justifying belief-forming method. So, although employing a distinct epistemic method in arriving at a belief is sufficient for warranting the belief to some extent, it is not necessary for warrant.

We thus follow several authors in distinguishing two species of epistemic warrant: *justification* and *entitlement*.²⁵ We understand *justification* internalistically, as requiring (i) that one employ a distinct epistemic basis in forming a belief, and (ii) that the believer must have reflective access to the basis for her belief.²⁶ We understand *entitlement* to represent a way for belief to be warranted in the absence of (i): one can be entitled even if one has done no ‘evidential work’ (or has employed no special method) to earn that entitlement (Wright 2004, 174).²⁷ This way of thinking about entitlement captures the rejection of the Distinct Epistemic Basis requirement articulated above.

On Dretske’s understanding of entitlement, we are entitled to beliefs that we cannot avoid having (Dretske, 2000). Wright (2004) holds that although one may *accept* the propositions to which one is entitled, entitlement does not provide any justificatory evidence for the truth of those propositions. A worry for these views of entitlement is that they do not

²⁵ See Altschul, 2011, Wright, 2004, Burge, 2003, and Dretske, 2000.

²⁶ This way of understanding justification is meant to align with a Cartesian, rationalist view of justification.

²⁷ We take inspiration here from Wright’s 2004, which portrays entitlement as an epistemic good that does not require one to do any evidential work.

explain what makes entitlement an *epistemically* good-making feature - they do not assign a clear *positive* epistemic status to entitlements – one that, for example, systematically connects the entitlement with truth.

However, we think there is a category of entitled belief that *is* appropriately connected with truth: we call this category *grounded belief*. Here ‘being grounded’ is to be understood in the sense of being *anchored*, rather than in the sense of being made on some specific *grounds*.²⁸ The grounding for one’s belief provides reason for belief, albeit not the sort of reason envisaged by the Distinct Epistemic Basis requirement. The views of self-knowledge we consider (constitutivism and neo-expressivism) will take basic self-beliefs to provide paradigm examples of *grounded* (in the present sense) *yet baseless* belief.

One might worry that views that take basic self-belief to be ones we are entitled to (in virtue of being grounded) but have no justification for (because not satisfying the Distinct Epistemic Basis requirement) cannot count as forms of epistemological disjunctivism. As discussed earlier, EDP employs an *internalist* notion of epistemic warrant, but on some views of entitlement (e.g. Burge, 2003), entitlement is an *externalist* type of warrant. However, the notion of entitlement we employ in this paper is neutral with respect to the (epistemological) internalism/externalism debate. It is at least a theoretical possibility for grounded beliefs to be internalistically warranted. While (in our terminology) such beliefs would not count as epistemically *justified*, they would still be internalistically warranted. In the case of basic self-knowledge, all that would be required for self-belief to be internalistically warranted is for one

²⁸ There is a similar use of the notion of ‘grounding’ in contemporary discussions of theories of truth, as well as metaphysics.

to have reflective access to that which provides the grounding for one's self-belief, and all that requires is for one to be able to provide her reason for the belief in terms of that grounding. Granting that one's self-beliefs are grounded in the mental states that those self-beliefs concern, and granting (as is plausible) that one can typically cite her mental states when providing reasons for her mental self-beliefs, it will turn out that we have reflective access to the grounding for our basic self-beliefs.

We are now in a position to evaluate two different views that embrace disjunctivism about basic self-knowledge.

4. Constitutivism about Basic Self-Knowledge

4.1 Constitutivism

Proponents of *constitutivism* maintain that spontaneous mental self-beliefs count knowledge courtesy of *constitutive* connections they bear to the states of mind they are about.²⁹ Believing and other states of mind – both 'evaluative' attitudes, such as intending, desiring, hoping, and 'receptive' experiential states – are 'intrinsically known'. We have privileged knowledge of such states "not because we are in a specially good position to form second-order beliefs about that reliably track their existence, but because their existence is normally constituted by our knowing assessment" them (Boyle, 2011, p. 237). Thus, "in the normal and basic case, [e.g.] believing P and knowing oneself to believe P are not two cognitive states; they are two aspects of one cognitive state" (2011, p. 228).³⁰ This means that the constitutivist is committed to a strong thesis of *self-intimation*: in the specific case of mental

²⁹ See Peacocke, 1999, Zimmerman, 2006, Coliva, 2012. For critical discussion, see Bar-On, 2004, pp. 388ff., 2009, and 2017.

³⁰ For a more functionalist version, see Zimmerman 2006, p. 343f. (expounding Shoemaker's view).

states, it is metaphysically impossible for you to be in a state of mind M and not know it, since your M, and your self-belief that you are in M, are in effect not 'distinct existences'. (It is for this reason that it is a mistake to think of basic self-knowledge, as do introspectionist views, in terms of a kind of perception-like tracking of "some independent condition." (2011, p. 239)).

Constitutivism, then, takes the epistemic status of (true) self-beliefs as knowledge to be underwritten by the *metaphysics* of mental states: these states have beliefs 'built into' them. Possessing a basic self-belief that you are in a mental state is simply guaranteed by your *being* in that state. In particular, mental self-beliefs are not understood as ones you *arrive at* using observation, inference, or any other epistemic method; for having these (second-order) beliefs is simply a constituent or aspect of being in the relevant states. But, despite having no epistemic *basis*, these beliefs – according to the constitutivist – are still epistemically warranted. For they can be said to be epistemically grounded (in our sense – see 3.2) in the first-order mental states they are about.³¹ Thus, on the constitutivist view, the first-order states make the second-order beliefs both true and epistemically warranted.³²

This means that constitutivism is committed to a form of *epistemological* disjunctivism: when one has a second-order belief, *either* that belief is epistemically grounded in the first-order mental state that makes it true, or there is no first-order mental state that makes the belief true, and so no first-order mental state to provide epistemic grounding for the belief. So a true and a false self-belief that one is in M have different epistemic grounding; the true belief

³¹ Recall the use of terminology laid out in section 3.2.

³² See Zimmerman, 2006, p. 343ff. Zimmerman frames his discussion primarily in terms of second-order beliefs about first-order beliefs. But constitutivists (including Zimmerman and Boyle) typically extend the constitutivist claim to other first-order mental states.

will be grounded in the state it is about, while the false belief cannot be. If a self-belief's epistemic warrant depends only on its epistemic grounding, a true and a false self-belief cannot share their epistemic warrant; hence epistemological disjunctivism.

As noted earlier, constitutivists reject perceptual (as well as inferential) models of basic self-knowledge. There is, on this view, a contrast between basic self-knowledge and perceptual knowledge, in that mental self-beliefs are not metaphysically independent of the states they are about the way perceptual beliefs are independent of the observable facts they are about.³³ Whereas in the perceptual case, given the metaphysical independence of the relevant belief and fact, there is an 'epistemic distance' one must traverse to attain knowledge (one's perceptual beliefs must hook up to the worldly facts in the right way), there is no such distance to traverse in the case of self-beliefs.

4.2 Constitutivism and EDSK

We are now in a position to see how the constitutivist avoids the challenge presented earlier for EDP. On our diagnosis, the challenge arises because, if we follow Burge in insisting that we must acknowledge the presence of an explanatorily relevant perceptual state in common between the good and the bad cases (to avoid contradicting empirical psychology), it becomes very difficult to deny the epistemic relevance of such a state. Even if perceptual beliefs are not arrived at via *inferences* from perceptual experiences, it seems difficult not to regard the perceptual beliefs as epistemically warranted by these experiences. The difficulty for EDP is then to explain how the good case and the bad case can differ in terms of their epistemic warrant, given that they are based on perceptual experiences of the same type. But, according

³³ See e.g. Zimmerman 2006, p. 342 and Boyle, 2010.

to the constitutivist, mental self-beliefs are not *based on* anything; in particular, they are not based on any (possibly shared type of) experience concerning how one's mental state *appears* to one. So there is no pressure to acknowledge an epistemically relevant psychological commonality between the true and false self-believer in terms of some perceptual self-experience. But then the road is clear for the constitutivist to propose that the good and bad case differ in terms of their epistemic warrant. In the good case, *S*'s basic self-belief is epistemically warranted, because *grounded* in the relevant first-order mental state (which is also what makes the self-belief true). Not so in the bad case. Whereas EDP faces an apparently insurmountable challenge, the constitutivist's EDSK does not.

However, the constitutivist idea that being in a mental state is metaphysically inseparable from believing that one is in it is problematic, and there are several extant objections to constitutivism both as a view about the metaphysics of mental states and as a view of basic self-knowledge. For reasons of space, we only mention one problem that is the most relevant to our concerns here.³⁴

Consider a *false* belief that one is in *M*. A false mental self-belief cannot be properly 'built into' – or be inseparable from – the state of mind it is about. Being false, there is no state of mind for it to be 'built into'! So the constitutivist needs to tell us what to make of the straightforward possibility of false basic mental self-beliefs. However the constitutivist accommodates such self-beliefs, we think she faces the following *prima facie* difficulty. (One main advantage we see for the alternative account we consider in the next section is its

³⁴ For a critical discussion of constitutivism, see Bar-On, 2004, 2009, and 2017. (And see also Bar-On, "Ordinary Self-Knowledge: Base-less Yet Privileged?", in progress.)

potential for avoiding this dilemma.) A key component in the commonsense explanandum of basic self-knowledge is the privileged epistemic status of basic mental self-beliefs – the fact that, unlike other beliefs, they are strongly presumed to be true and to enjoy a special entitlement (even though mental self-believers are not expected to possess, or required to offer, any evidence for the beliefs).³⁵ The constitutivist purports to explain this special epistemic status by appealing to the *metaphysics* of mental states: the fact that mental states are partially constituted by self-beliefs. But that explanation can now be seen as insufficient, in light of the possibility of false self-beliefs. A false basic mental self-belief, on the constitutivist account, is not only false; it also lacks *all* warrant, since the only source of epistemic warrant for a basic belief that one is in M (given EDSK) is the state that grounds it (M). But this means that false and true mental self-beliefs cannot be said to enjoy the same secure epistemic status – the status we ordinarily accord them. That status, it seems, is indifferent to whether the belief is true or false. A false (nonevidential) mental self-belief would still seem to enjoy the same presumptive privileged status that would be accorded to a corresponding true belief – simply in virtue of being a (nonevidential) *mental self*-belief. (This is a point of disanalogy between ordinary mental self-beliefs, on the one hand, and perceptual, as well as third-person mental beliefs, and bodily self-beliefs, on the other. See below, Section 5.)

This means that an acceptable version of EDSK has to make room for *some* epistemic commonality between true and false mental self-beliefs. Thus, whereas EDP (as we saw earlier)

³⁵ Bar-On 2004 (esp. Ch.s 1 & 9), argues that this special epistemic status (aka ‘first-person authority’) constitutes a more salient and stable component of our commonsense view than the self-intimation that constitutivism tries to capture and that it provides a more feasible explanandum for a philosophical theory of basic self-knowledge.

struggles to explain in virtue of what perceptual beliefs in the good and the bad cases *vary* in terms of epistemic warrant, constitutivism faces a difficulty in explaining what mental self-beliefs in good and bad cases have *in common* in terms of their epistemic warrant. (As we shall argue below, neo-expressivism is – by contrast – well placed to capture both the epistemic similarities and the epistemic differences between true and false mental self-beliefs.)

5. Neo-Expressivism and Basic Self-Knowledge

We now turn to an alternative view of basic self-knowledge – *neo-expressivism* – which, like constitutivism, endorses a form of EDSK, but which avoids the problem facing constitutivism mentioned in the previous section. Like constitutivism, neo-expressivism rejects the perceptual model of basic self-knowledge and denies that basic self-beliefs must be arrived at via some distinct epistemic route or formed on some distinct epistemic basis in order to qualify as instances of genuine (and privileged) knowledge. And this allows neo-expressivism also to avoid the main challenge facing EDP. However, unlike constitutivism, neo-expressivism is not committed to the idea that there is a metaphysical dependence between mental self-beliefs and the states they are about, and so it can straightforwardly acknowledge the possibility of false basic self-beliefs and can account for the epistemic commonality between true and false mental self-beliefs.

5.1 Neo-Expressivism

Bar-On (2004, 2011, 2009, 2012, and elsewhere) develops an account of the *distinctive epistemic security of avowals* (where avowals are understood to be nonreflective, or ‘unstudied’, self-attributions of occurrent mental states that are made on no distinct epistemic basis). On Bar-On’s view, avowals are distinctive *acts* that serve to express *in the act sense* (‘a-

express') the very mental state that the avower self-attributes, in addition to expressing the (higher-order) belief that she is in the state. So, for example, in avowing (out loud or silently) "I am so glad to see you" you a-express your delight at seeing your addressee, and not just your *belief* that you are happy to see her. This renders the avowal – as an act – similar to expressive acts such as saying (or thinking) "It's so good to see you", or just giving a hug, though understood as a *product* – or in terms of the *expressive vehicles* used – it is different *semantically* from these other expressions. ("I'm so glad to see you" expresses in the *semantic* sense – 's-expresses' – a self-ascription, true iff you are glad to see your addressee, whereas "It's so good to see you" s-expresses a non- self-ascriptive proposition, and the hug does not s-express any proposition.) Bar-On argues that her account enables us to capture the intuitive epistemic contrasts between avowals and all other attributions that merely *report* contingent matters (including third-person mental attributions, perceptual judgments, and all bodily self-attributions, as well as all evidential mental self-attributions). Notably, the account helps explain why avowals are not open to ordinary doubts, epistemic challenges or requests for reasons; they are not subject to simple correction or defeat. (This is a point the constitutivist apparently struggles to explain in the case of false avowals). At the same time, it also accommodates the semantic continuities between avowals and truth-evaluable pronouncements, as well as semantic differences between avowals and other expressions.

Now the question often at the center of discussions of basic self-knowledge is how our mental self-*beliefs* can be taken to manifest things we *know* about ourselves, and know in a unique and privileged way. This question is especially pressing for any view that – like neo-expressivism – rejects the Distinct Epistemic Basis requirement, and maintains that the special

epistemic status of our avowals is not due to the fact that the self-beliefs they express are arrived at them through the use of a special epistemic method (such as perception, evidential inference, self-interpretation, etc.).³⁶ As regards accounts of basic self-knowledge that, specifically, explain the special status by appeal to the reliability of an internal perception-like mechanism, Bar-On argues that the best such accounts can do is treat the contrasts between avowals and other ascriptions as on a par with the epistemic contrasts between first-person and third-person attributions of certain bodily states. However, assimilating the epistemic status of avowals to that of bodily self-reports fails to do justice to the distinctive security of basic mental self-beliefs as ordinarily understood.³⁷

5.2 Immunity to Error³⁸

Consider self-attributions issued through proprioception or kinesthesia. Such self-attributions share a certain epistemic feature with avowals: they are ‘identification-free’. In normal circumstances, if I say or think: “My legs are crossed” or “I’m spinning around”, then, epistemically speaking, my self-attribution does not rely on my *recognizing* some individual as myself and taking that person’s legs to be crossed, etc. (Indeed, I have no more reason for thinking that *someone’s* legs are crossed than whatever reason I have for thinking that *my* legs are crossed.) To use a phrase due to Gareth Evans and Sydney Shoemaker, bodily self-

³⁶ For Bar-On’s rejection of the perceptual view, and of the Epistemic Approach in general, see her 2004, Ch. 3, 4, 2009, and Bar-On and Long 2001.

³⁷ See Bar-On, 2004, esp. Ch. 4 and *passim*.

³⁸ For fuller discussion, see Bar-On 2004, 2009, and 2012.

attributions of this kind are “immune to error through misidentification” (IEM, for short).^{39,40} When a self-attribution of the form “I am F” is IEM, then, although I may fail to be F – so my self-attribution may be false – there is no room for me to think: Someone is F, but is it *me*? I cannot sensibly doubt that it is *me* who has the relevant property F without doubting that *someone* has it. (Contrast this with a case in which I, e.g., tell that I have \$500 in my bank account by consulting the bank teller's screen. Here I *can* wonder whether someone – though not me – has \$500 in her account.)

Importantly, immunity to error through misidentification does not reflect a special recognitional success, that is, success in identifying – by recognizing – the “right” individual of whom to predicate F. Quite the opposite: if anything, this kind of immunity reflects the *absence* of recognitional identification.⁴¹ Still, on Evans' analysis, self-attributions that are IEM can represent *knowledge* that we gain about ourselves in a distinctive way. Evans notes that we possess two general capacities for gaining information about some of *our own* states and properties – the capacity to “perceive our own bodies” (which includes “our proprioceptive sense, our sense of balance, of heat and cold, and pressure”), and the capacity to determine our own “position, orientation, and relation to other objects in the world ... upon the basis of

³⁹ See below, where we distinguish IEM₁ from IEM₂.

⁴⁰ For discussion of the phenomenon of immunity to error through misidentification, see Wittgenstein 1958, pp. 66–67, Evans, 1982 (esp. ch. 7, sec. 2), Shoemaker, 1968, and Wright, 1998, pp. 18–20. For a very useful analysis, see Pryor, 1999.

⁴¹ There is a “thin” sense in which I do identify myself as the subject of the ascription, for I do manage to *refer* to myself, and referring is picking out, and picking out is a form of identifying. It may be useful to distinguish between the *referential* notion of identifying (a semantic notion) and the *recognitional* notion of identifying (which is an epistemic notion). Compare Evans, 1982, p. 218. For a related discussion of the differences between semantic and epistemic aspects of perception, see Burge 2005, pp. 6–9.

our perceptions of the world.”⁴² When a subject gains information of the form “I am F” (for the relevant range of F's) using one of these capacities, Evans remarks, “[t]here just does not appear to be a gap between the subject's having information (or appearing to have information), in the appropriate way, that the property of being F is instantiated, and his having information (or appearing to have information that that *he* is F; for him to have, or to appear to have, the information that the property is instantiated just is for it to appear to him that *he* is F” (1982, 221). So self-attributions of the form “I am F” that are IEM can represent secure knowledge I have that *I* myself am F, *even though* they do not rest on recognitional identification of the one who is F.

Now, Bar-On's neo-expressivist account of basic self-knowledge begins with the suggestion that paradigmatic present-tense nonreflective mental self-attributions of the form “I am in M” or “I'm M(ing) (*that*) c” (where M is a mental state and c is its putative intentional content⁴³) are issued ‘in the avowing mode’: they are not made on the basis of (inner) observation, tacit deliberation, evidence, or interpretation. And they are *immune to error through misascription* – IEM₂ – in addition to being immune to error through misidentification – (which we'll now label) IEM₁.⁴⁴ Suppose I produce a self-attribution such as “I am nervous about this dog”, or “I am hoping that you'll join us tonight”, or “I am feeling tired”. Epistemically

⁴² Evans, 1982, pp. 220, 222.

⁴³ The phrase “intentional content” is here used to cover both intentional object (e.g. “I'm afraid of *the dog*”) and propositional content (e.g. “I'm hoping that *it won't rain today*”).

⁴⁴ Bar-On first introduces the notion of ascriptive immunity to error to characterize avowals' distinctive security in 2000 and develops it in 2004, 2004a, 2009, and 2012. In his 2012, Wright offers a discussion of immunity to error that has several notable points of contact with Bar-On's, though he does not make the connection – crucial to Bar-On's neo-expressivism – between avowals' distinctive immunity to error and their expressive character. Moreover, assuming that avowals are taken to be made on no epistemic basis, Bar-On takes it to be necessary to supplement the negative characterization in terms of immunity to error with a positive epistemological account of what qualifies them as knowledge.

speaking, my self-attribution does not rely on my recognition of a state in me as having a certain character and content, based on how it *appears* to me (say, as a result of introspection). I do not recognize a state of mine as being M (as opposed to M') and having content *c* (rather than *c'*). (Indeed, I have no reason – specific to the occasion on which I issue the avowal – for thinking that I am in *some* state or other, and that it has some content, other than whatever reason I have for thinking that I am in mental state M with content *c*.⁴⁵) Consequently, in making the self-attribution, I am not subject to the possibility of a *brute* recognitional error (where a brute error is one that is simply due to the world failing to cooperate, rather than being due to some kind of failure of the subject's faculties).⁴⁶ And this is in good part what makes the perceptual model inappropriate for mental self-beliefs. (Note, however, that to say that (basic) mental self-attributions are not open to brute recognitional errors of identification

⁴⁵ In the case of mental ascriptions to others, I *do* typically have such independent reasons. And mental *self*-attributions can also take a more evidential character – when they rely on e.g. testimony, observation or interpretation for determining, e.g., whether one is scared of something or what one is scared of. But this is not the case when one is simply *avowing* being in a mental state.

⁴⁶ For discussion of the notion of brute error and its connection to immunity to error, see Bar-On 2004, pp. 9f., 183, 200f., 332f. Bar-On's neo-expressivist view (unlike standard expressivist views) makes room for the possibility of false avowals. However, on her view, false avowals constitute species of *expressive failures*; they are not examples of brute errors. Expressive failures, Bar-On argues, are to be understood in *psychological*, rather than epistemological terms. A false avowal, thus understood, involves a self-attribution that is *wrong* (because false) but not because the avower has *gone* wrong, in moving from an internal 'judgment of appearance' to the relevant self-belief. (See Bar-On 2004, Ch. 8, where an explanation of e.g. self-deception and wishful thinking is provided along these lines.)

Doyle (this volume) argues that accommodating the possibility of false basic self-beliefs (at least ones that concern one's first-order *beliefs*) requires allowing that there are 'ringers' for the states the beliefs are about – viz., states of the individual that are "subjectively indistinguishable between good and bad cases" and that "mislead us when we err about our own minds" (p.XX). However, this ignores the above-mentioned option, of explaining false self-beliefs on the model of expressive failures. That model rejects the idea that when a subject errs about her own mind this must be because she has been *misled*, or fooled by a 'ringer'. (For a directly relevant discussion of the status of internal 'impostors', see Bar-On 2004a.) Doyle argues that appeal to an epistemically relevant element that is subjectively indistinguishable between the good and the bad case (what he refers to as a 'conscious judgment', an 'inner assertion') is necessary in order to explain "how the possession of self-knowledge is intelligible from the subject's point of view" (p. X). We think that the dual immunity to error discussed here should help in providing such an explanation, insofar as this immunity is something to which self-believers have reflective access. A fuller explanation will have to await another occasion.

or ascription is not to say that they are absolutely infallible or incorrigible. It is just to say that they are protected from a much wider array of epistemic errors, doubts, and corrections than other attributions, including, specifically, proprioceptive and kinesthetic self-reports.⁴⁷

Interestingly, perceptual beliefs *are* open to brute error; a perceptual belief can be false even when the believer's perceptual faculties are working perfectly. Accordingly, perceptual beliefs are *not* immune to brute recognitional error in the way that basic mental self-beliefs are. An adequate disjunctivist account of basic self-knowledge should explain the immunities to error of basic mental self-beliefs, but there is no pressure for EDP to provide a parallel explanation for perceptual beliefs.

The characterization of avowals' distinctive security as a matter of their being IEM₂ (in addition to IEM₁) provides a suitably tempered interpretation of the familiar claim that our own basic mental self-beliefs are absolutely protected from epistemic challenge (doubt, correction, falsification). The dual immunity to error can in part explain why, if we consider your present thought that you are, say, feeling annoyed by your friend, the self-belief it manifests appears at once to lack any distinct epistemic basis and yet to be especially likely to be something you know (as no one else does). And it explains why basic mental self-attributions are not subject to perception-like *brute* recognitional errors.⁴⁸

The above characterization does not commit one either to Cartesian privileged access, or, indeed, to any distinctively secure epistemic basis on which avowals supposedly are made.

⁴⁷ For discussion of the possibility of false avowals, see Bar-On, 2004, esp. pp. 320-335 and 394ff.

⁴⁸ On occasion, one might 'second-guess' an earlier avowal - "I thought I was annoyed at you, but really I was frustrated by this puzzle." Note, however, that when this occurs, one is (on the occasion of second-guessing) not expressing frustration in the *avowing* mode, but offering a self-report from a third-personal stance. The later self-attribution does not represent a *basic* self-belief; and its occurrence does not reveal the earlier avowal to involve a brute error.

Moreover, it does not require supposing that the secure epistemic status of these self-attributions is vouchsafed via some conceptual or metaphysical guarantee, as per constitutivism. However, by itself, the characterization does not yet give us a full account of avowals' status. For one thing, we need to understand the *source* of the additional immunity to error enjoyed by basic mental self-attributions.⁴⁹ *Why* is it that they are not only IEM₁ but also IEM₂? In addition, we have seen that immunity to error, in general, is no guarantee of *truth*. So we need to understand why basic mental self-attributions contrast with all other ascriptions in being so strongly presumed to be true; and, given that they are made on no epistemic basis, we need to know what provides *positive* epistemic warrant for them.

5.3 Immunity to Error, Expression, and Knowledge⁵⁰

By way of appreciating the expressive character of spontaneously produced mental self-attributions, consider a case in which you falsely (though sincerely) avow "My tooth hurts!" (say, at the dentist's chair, as the drill approaches your mouth). Under the circumstances, you might have equally said: "Ow!", or emitted a yelp, or winced. Given that, it does not seem plausible to regard the avowal – but not the "Ow!" or the wince – as an upshot of *mistaking* (say) fear of the approaching drill for pain, resulting in a brute error about your state. You falsely think or judge that your tooth hurts, but you didn't come by this judgment in consequence of being fooled by the appearance of your internal state. It seems much more plausible to regard the avowal as, just like the yelp, something forced out of you, though in this

⁴⁹ In the case of proprioceptive reports, the source of their immunity to error through misidentification has to do with our possessing special mechanisms for obtaining information concerning our *own* bodies. See Evans (1982), Ch. 7.

⁵⁰ For a more extended presentation of the ideas summarized here, see Bar-On 2011 (see also 2009 and 2012). Full discussion appears in Bar-On 2004, Ch.s 6-8.

unusual case not by an actual toothache, but rather by the priming effect of fear (fueled, perhaps, by painful dental history). The avowal, as an act, is on a par with the “Ow!” or the yelp/wince, in being an expressive act. It is no more plausible to regard the avowal as manifesting an *evidence- or observation-based* self-belief than it is to regard the spontaneous yelp/wince as so based. Although we say that a true basic self-belief is *grounded* in the mental state the belief concerns, this is not the same as saying that the self-belief has the mental state as a distinct *epistemic basis*. As we have understood the notion, having a distinct epistemic basis involves employing an epistemic method in *forming* the relevant belief that one can cite in support of that belief. This is absent in the case of basic self-belief. The self-belief one has when avowing her mental state is a belief she simply finds herself with.

On the neo-expressivist account, one who avows her state of mind is *speaking from* that state (using a self-ascriptive vehicle), rather than merely *reporting* (or otherwise informing of) a self-belief she has arrived at on the basis of the way her state appears to her. Insofar as the self-attribution made when avowing is not made on any recognitional basis, it is not subject to brute errors of misidentification or misascription. This is also why self-attributions produced in the avowing mode are so strongly presumed to be true, since to take someone to be avowing M is just to take it that she has given voice to her M – which means taking the self-attribution to be true. The relevant presumption of truth does not amount to a conceptual guarantee of truth; it is defeasible. We *can* make sense of a spontaneous self-attribution of an occurrent state of mind being false, as in the dentist case, or certain cases of self-deception, implicit bias, and so on. Still, as noted earlier, in such cases we do not suppose the avower to be subject to a purely brute ‘recognitional’ error, of the sort that afflicts one who is subject to a perceptual

illusion. Falsity in these cases is not plausibly taken to be due to a subject's simply mis-taking of one of her mental states for another; we expect it to have its source in some psychological irregularity, impairment, failure, or interference.

Note that on the present account, there is no special difficulty in accommodating the straightforward possibility of false basic self-beliefs and the doxastic commonality between the good and bad cases of self-knowledge. The person who is subject to self-deception, or to some priming effect, would be just as disposed to (sincerely) avow: "I am in M" as the person who truly believes she is in M. By ordinary standards, that person would be straightforwardly credited with having a self-belief as naturally as the person who has a true self-belief. The neo-expressivist, in contrast with the constitutivist, can make room for this, since she allows what the constitutivist denies: that there is (in general) an ontological independence between mental self-beliefs and the mental states that would make such beliefs true or false. At the same time, like the constitutivist, the neo-expressivist denies that (basic) mental self-beliefs must be epistemically *based on* internal experiences of appearances, which experiences, moreover, explain the beliefs' positive epistemic status.

5.4 Neo-Expressivism and EDSK

If basic mental self-beliefs indeed have no epistemic basis, must we suppose – as do constitutivists – that such beliefs can only be epistemically warranted insofar as they are epistemically grounded, and that they can only be epistemically grounded if they are true – that is, if the believer is in the state that her belief is about? If so, then epistemic disjunctivism about basic self-knowledge follows.

However, our discussion of neo-expressivism so far opens up another possibility.⁵¹ Insofar as basic mental self-beliefs are not based on any inference, evidence, or self-experience, it doesn't seem that one could have *justification* for the relevant self-attributions (see earlier discussion of justification in Section 3.2). However, this does not mean that one cannot have epistemic *warrant* for the self-belief, in the sense of being *entitled* to it. As a subject of bodily states with normal physiological capacities, if I think, in the ordinary way, that my legs are crossed, I can be said to be entitled to believe that it is *my* legs that are crossed, even if I have relied on no evidence, and have done nothing to ascertain the identity of the person whose legs are crossed, say, by ruling out imposters. I can perhaps be said to be *entitled by default* to this belief.⁵² Similarly, given the neo-expressivist analysis of avowals' security (in terms of a distinctive immunity to error), perhaps, as a subject of mental states with normal psychological capacities, I am in a unique position to give voice to a state of mind by self-attributing it, and this renders me entitled by default to the relevant self-belief, even if I have not done any evidential work or have not ruled out alternatives to the attribution I make.⁵³

If it is correct to think that basic mental self-beliefs enjoy entitlement by default, however, one might wonder whether neo-expressivism represents a form of disjunctivism at all. After all, both true and false self-beliefs enjoy this sort of entitlement, and so in this respect are alike in epistemic warrant.⁵⁴ Yet the defining feature of epistemological disjunctivism is the

⁵¹ For discussion, see Bar-On, 2004, 381ff.

⁵² See Bar-On, 2004, 374ff. for discussion of entitlement by default.

⁵³ To say that one is entitled by default is, of course, not to say that one's self-belief is *true* by default. Truth, on the present proposal, is simply a matter of whether one is in M or not, which may in general be independent of whether one *thinks* (or judges, or believes) that one is in M.

⁵⁴ One may also worry, in addition, that entitlement by default is not sufficient to constitute a positive epistemic good. (See above, 3.2.)

commitment to there being a difference in epistemic warrant between belief in the good vs. the bad case.

We think that despite positing an epistemic commonality between true and false basic self-belief in terms of entitlement, neo-expressivism can still count as a form of epistemological disjunctivism. This is simply because the epistemic warrant for a true self-belief is not exhausted by entitlement by default; a true self-belief enjoys another sort of epistemic warrant that is *not* had by the relevant false self-beliefs.

Like the constitutivist, the neo-expressivist can insist that a true basic mental self-belief will have epistemic *grounding* – in *the very state* the belief is about (in addition to the believer being entitled to it by default). Not so when it comes to the corresponding false self-belief. When one is not in M, one can still believe that one in M, and even be entitled by default to that belief; but one is not (fully) warranted, since one's belief fails to be epistemically grounded.⁵⁵

Note that this constitutes what one might call a *weak epistemological disjunctivism*. A *strong* epistemological disjunctivism would deny that there is *any* epistemic warrant shared by beliefs in good and bad cases. A weak epistemological disjunctivism would allow that belief in good and bad cases can share some sources of warrant, yet still have different epistemological statuses: beliefs in bad cases will not be *fully* warranted (in a way that allows for knowledge), while beliefs in good cases are fully warranted (in a way that renders them instances of knowledge). One advantage of the weaker version is that it straightforwardly accommodates

⁵⁵ Bar-On 2004, 388ff. spells out a view along these lines and offers some neo-expressivist motivations for it.

the possibility of false basic mental self-beliefs, and allows that such self-beliefs can be as epistemically innocent – and even enjoy some of the same entitlement – as their true counterparts, without invoking any epistemically relevant self-experiences. At the same time, it acknowledges that *being* in the relevant mental state has bearing on the epistemic support one has for one’s self-belief (without bringing the self-belief in its train). This can be seen as an advantage, inasmuch as it captures something about subjects’ epistemic position in ordinary circumstances: both the true and the false believer, if pressed, would offer *M* itself as the reason for their belief that they are in *M* (as well as for avowing *M*).

6. Some Concluding Remarks

One of our central aims in this paper has been to show that EDSK is better placed to avoid a serious challenge that faces EDP (presented in 2.1). But one might now wonder whether our discussion of EDSK points the way to a version of *EDP* that can *also* avoid the challenge. Unlike in the case of basic self-knowledge, in the case of perceptual knowledge, adopting a non-perceptual model is obviously not an option. Still, the proponent of EDP may not be left without response. By way of conclusion, we offer some reflections on behalf of this proponent, leaving further discussion for future work.

First, when posing the problem for EDP (in 2.1), we noted that for EDP to accept the Perceptual Warrant Principle (PWP) (according to which the warrant for perceptual belief is constituted by perceptual experience) *likely* involves commitment to MDP. We say that it *likely* involves MDP, rather than *requiring* such commitment, because there is another theoretical option open to epistemological disjunctivism. One can accept both EDP and PWP *without* accepting that perceptual experience in the good case is *radically* different from perceptual

experience in the bad case. One need only accept, instead, that while perceptual experience is of the same fundamental metaphysical *type* in the good and the bad case, there are nevertheless significant *differences* between the experiences (this is what Byrne and Logue call the Moderate View about perceptual experience - note that this view is *not* itself a version of MDP, but is instead a competitor; 2009, pp. x-xi). These differences, one might argue, could then underwrite the difference in warrant between belief in good and bad cases that *epistemological* disjunctivism claims. We are not familiar with anyone who explicitly endorses the Moderate View of perceptual experience as a supplement to epistemological disjunctivism, but this combination of views may be an interesting way forward for epistemological disjunctivists concerned to avoid the criticism presented in section 2.1.⁵⁶

Second, recall from our discussion of neo-expressivism that the dual immunity to error in the case of avowals provides a default entitlement for basic mental self-beliefs, where this default entitlement is shared between true and false basic self-beliefs. A *true* basic self-belief gains additional epistemic warrant because it is epistemically grounded in the mental state it is about. This feature of the view qualifies it as a form of epistemological disjunctivism, since, given this feature, a true self-belief would still have a different warrant from a corresponding false one.

Perhaps a similar move is available to EDP. It may be that one is epistemically entitled to one's perceptual beliefs because of the nature of one's perceptual experience. Granting that

⁵⁶ It should be noted that alternatively, on McDowell's brand of disjunctivism, it seems that in the case of veridical perceptual belief, the perceptual state itself contains the fact that the belief concerns. On this view, there is no 'epistemic distance' one needs to traverse for one's (true) perceptual belief to amount to knowledge. And this parallels the neo-expressivist view proposed here. (Thanks to Ram Neta for pointing out this parallel.) However, it is precisely this feature of the McDowellian view that makes it controversial as an account of perceptual knowledge.

one's perceptual beliefs in the good and the bad case are epistemically based on the same type of perceptual experience, one is equally entitled to one's belief in the good case as one is in the bad case. Nevertheless, the epistemological disjunctivist might add that there is a further source of warrant available in the good case that is not available in the bad case; namely, the factive reason that one sees that such-and-such. However, while in the case of self-belief it is clear that the epistemic grounding for one's believing comes from the very mental state that the belief concerns, the point that is at issue when it comes to assessing EDP is whether or not perceptual beliefs enjoy *any* kind of epistemic warrant other than what is provided by the nature of one's perceptual experience – this cannot simply be assumed. But our observations at least point the way toward a weaker epistemological disjunctivism about perception, which – like the weaker disjunctivism about basic self-knowledge – accepts that beliefs in good and bad cases may share *some* kind of epistemic warrant.⁵⁷

⁵⁷ We thank Casey Doyle, Joe Milburn, Matt McGrath, and Ram Neta for helpful comments on earlier drafts of this paper. We thank the Language and Mind group at Arche for questions and discussion on a presentation of this paper.

References

- Altschul, J. (2011). Epistemic entitlement. In *The Internet Encyclopedia of Philosophy*. Retrieved from <http://www.iep.utm.edu/ep-en/>
- Bar-On, D. (2000). Speaking my mind. *Philosophical Topics*, 28(2), 1-34. DOI: [10.5840/philtopics200028216](https://doi.org/10.5840/philtopics200028216)
- Bar-On, D. (2004a). Externalism and self-knowledge: Content, use, and expression. *Nous*, 38(3), 430-55. DOI: [10.1111/j.0029-4624.2004.00477.x](https://doi.org/10.1111/j.0029-4624.2004.00477.x)
- Bar-On, D. (2004b). *Speaking my mind: Expression and self-knowledge*. Oxford: Oxford University Press.
- Bar-On, D. (2009). First-person authority: Dualism, constitutivism, and neo-expressivism. *Erkenntnis*, 71(1), 53-71. DOI: [10.1007/s10670-009-9173-y](https://doi.org/10.1007/s10670-009-9173-y)
- Bar-On, D. (2011). Externalism and skepticism: Recognition, expression, and self-knowledge. In A. Coliva (ed.), *Self-knowledge and the self*. Oxford: Oxford University Press.
- Bar-On, D. (2012). Expression, truth, and reality: Some variations on themes from Wright. In C. Wright & A. Coliva (eds.), *Mind, meaning, and knowledge: Themes from the philosophy of Crispin Wright* (162-192). Oxford: Oxford University Press.
- Bar-On, D. (2015). Transparency, expression, and self-knowledge. *Philosophical Explorations*, 18(2), 134-152. DOI: [10.1080/13869795.2015.1032334](https://doi.org/10.1080/13869795.2015.1032334)
- Bar-On, D., and Long, D. (2001). Avowals and first-person privilege. *Philosophy and Phenomenological Research*, 52(2), 311-335. DOI: [10.1111/j.1933-1592.2001.tb00058.x](https://doi.org/10.1111/j.1933-1592.2001.tb00058.x)
- Boyle, M. (2010). Bar-On on self-knowledge and expression. *Acta Analytica*, 25(1), 9-20. DOI: [10.1007/s12136-009-0075-z](https://doi.org/10.1007/s12136-009-0075-z)
- Boyle, M. (2011). Transparent self-knowledge. *Proceedings of the Aristotelian Society Supplementary Volume lxxxv*, 223-241. DOI: [10.1111/j.1467-8349.2011.00204.x](https://doi.org/10.1111/j.1467-8349.2011.00204.x)
- Burge, T. (1993). Content preservation. *The Philosophical Review*, 102(4), 457-488. DOI: [10.2307/1523046](https://doi.org/10.2307/1523046)
- Burge T. (1996). Our entitlement to self-knowledge. *Proceedings of the Aristotelian Society*, 96(1), 91-116. DOI: [10.1093/aristotelian/96.1.117](https://doi.org/10.1093/aristotelian/96.1.117)
- Burge, T. (2003). Perceptual entitlement. *Philosophy and Phenomenological Research*, 67(3), 503-548. DOI: [10.1111/j.1933-1592.2003.tb00307.x](https://doi.org/10.1111/j.1933-1592.2003.tb00307.x)
- Burge, T. (2005). Disjunctivism and perceptual psychology. *Philosophical Topics*, 33(1), 1-78. DOI: [10.5840/philtopics20053311](https://doi.org/10.5840/philtopics20053311)
- Burge, T. (2011) Disjunctivism again. *Philosophical Explorations*, 14(1), 43-80. DOI: [10.1080/13869795.2011.544400](https://doi.org/10.1080/13869795.2011.544400)
- Burge, T. (2013a). Epistemic warrant. In T. Burge (ed.), *Cognition Through Understanding: Philosophical Essays, Volume 3* (521-33). Oxford: Oxford University Press.
- Burge, T. (2013b). Introduction. In T. Burge (ed.), *Cognition Through*

- Understanding: Philosophical Essays, Volume 3* (1-52). Oxford: Oxford University Press.
- Byrne, A. (2005). Introspection. *Philosophical Topics*, 33(1), 79-104.
- Byrne, A. (2018). *Transparency and self-knowledge*. Oxford: Oxford University Press.
- Byrne, A., and Logue, H. (2008). Either/or. In A. Haddock and F. Macpherson (eds.), *Disjunctivism: Perception, action, knowledge*, 57-94. Oxford: Oxford University Press.
- Byrne, A., and Logue, H. (2009). Introduction. In A. Byrne and H. Logue (eds.), *Disjunctivism: Contemporary readings* (vii-xxix). Cambridge: MIT Press.
- Carruthers, P. (2011). *The opacity of mind: An integrative theory of self-knowledge*. Oxford: Oxford University Press.
- Cassam, Q. (2014). *Self-knowledge for humans*. Oxford: Oxford University Press.
- Coliva, A. (2012). One variety of self-knowledge: Constitutivism as constructivism. In A. Coliva (ed.), *The Self and Self-Knowledge* (212-242). Oxford: Oxford University Press.
- Cunningham, J.J. (2016). Reflective Epistemological Disjunctivism. *Episteme*, 13(1): 111-132. DOI:10.1017/epi.2015.62
- Dretske, F. (2000). Entitlement: Epistemic rights without epistemic duties? *Philosophy and Phenomenological Research*, 60(3), 591-606. DOI: [10.2307/2653817](https://doi.org/10.2307/2653817)
- Evans, G. (1982). *Varieties of reference*. Oxford: Oxford University Press.
- Fricker, E. (1998). Self-Knowledge: Special access versus artefact of grammar - a dichotomy rejected. In C. Wright, B.C. Smith, and C. Macdonald (eds.), *Knowing Our Own Minds* (155). Oxford: Oxford University Press.
- Haddock, A., and Macpherson, F. (2008). Introduction. In A. Haddock and F. Macpherson (eds.), *Disjunctivism: Perception, action, knowledge* (1-25). Oxford: Oxford University Press.
- Macdonald, C. (2007). Introspection and authoritative self-knowledge. *Erkenntnis*, 67(2), 355-372. DOI: [10.1007/s10670-007-9072-z](https://doi.org/10.1007/s10670-007-9072-z)
- Madison, B.J.C. (2010). Epistemic Internalism. *Philosophy Compass*, 5(10): 840-853. DOI: 10.1111/j.1747-9991.2010.00333.x
- Martin, M.G.F. (2009). The reality of appearances. In A. Byrne and H. Logue (eds.), *Disjunctivism: Contemporary readings*, 91-116. Cambridge: MIT Press.
- McDowell, J. (2009). Selections from criteria, defeasibility, and knowledge. In A. Byrne and H. Logue (eds.), *Disjunctivism: Contemporary readings* (75-90). Cambridge: MIT Press.
- McDowell, J. (2010). Tyler Burge on disjunctivism. *Philosophical Explorations*, 13(3), 243-255. DOI: [10.1080/13869795.2010.501905](https://doi.org/10.1080/13869795.2010.501905)
- McDowell, J. (2013). Tyler Burge on disjunctivism. *Philosophical Explorations*, 16(3), 259-279. DOI: [10.1080/13869795.2013.808693](https://doi.org/10.1080/13869795.2013.808693)
- Neta, R., and Pritchard, D. (2007). McDowell and the New Evil Genius.

- Philosophy and Phenomenological Research*, 74(2): 381-396. DOI: [10.1111/j.1933-1592.2007.00023.x](https://doi.org/10.1111/j.1933-1592.2007.00023.x)
- Pappas, G. (2014). Internalist vs. Externalist Conceptions of Epistemic Justification. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2017 Edition). URL: [<https://plato.stanford.edu/archives/fall2017/entries/justep-intext/>](https://plato.stanford.edu/archives/fall2017/entries/justep-intext/)
- Peacocke, C. (1996). Our entitlement to self-knowledge: Entitlement, self-knowledge, and conceptual redeployment. *Proceedings of the Aristotelian Society*, 96(1), 117-158. DOI: [10.1093/aristotelian/96.1.117](https://doi.org/10.1093/aristotelian/96.1.117)
- Peacocke, C. (1999). *Being known*. Oxford: Oxford University Press.
- Pritchard, D. (2008). McDowellian neo-Mooreanism. In A. Haddock and F. Macpherson (eds.), *Disjunctivism: Perception, action, and knowledge* (283-310). Oxford: Oxford University Press.
- Pritchard, D. (2012). *Epistemological disjunctivism*. Oxford: Oxford University Press.
- Pryor, J. (1999). Immunity to error through misidentification. *Philosophical Topics*, 26(1-2), 271-304.
- Shoemaker, S. (1968). Self-reference and self-awareness. *Journal of Philosophy*, 65(October), 555-567. DOI: [10.2307/2024121](https://doi.org/10.2307/2024121)
- Shoemaker, S. (1986). Introspection and the self. *Midwest Studies in Philosophy*, 10(1), 101-120. DOI: [10.1111/j.1475-4975.1987.tb00536.x](https://doi.org/10.1111/j.1475-4975.1987.tb00536.x)
- Shoemaker, S. (1994). Self-knowledge and 'inner sense'. *Philosophy and Phenomenological Research*, 54, 249-314.
- Snowdon, P. (2009). The objects of perceptual experience. In A. Byrne and H. Logue (eds.), *Disjunctivism: Contemporary readings*, 49-74. Cambridge: MIT Press.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.
- Wittgenstein, L. (1958). *The blue and brown books*. Harper & Row.
- Wright, C. (1998). Self-knowledge: the Wittgensteinian legacy. *Royal Institute of Philosophy Supplement*, 43, 101-122. DOI: [10.1017/S135824610000432X](https://doi.org/10.1017/S135824610000432X)
- Wright, C. (2004). Warrant for nothing (and foundations for free)? *Aristotelian Society Supplementary Volume*, 78(1), 167-212. DOI: [10.1111/j.0309-7013.2004.00121.x](https://doi.org/10.1111/j.0309-7013.2004.00121.x)
- Wright, C. (2012). Reflections on Francois Recanati's 'immunity to error through misidentification: what it is and where it comes from'. In S. Prosser and F. Recanati (eds.), *Immunity to error through misidentification: New essays*, (247-280). Cambridge University Press.
- Zimmerman, A. (2006). Basic self-knowledge: Answering Peacocke's criticisms of constitutivism. *Philosophical Studies*, 128, 337-379. DOI: [10.1007/s11098-004-7797-y](https://doi.org/10.1007/s11098-004-7797-y)