

BRUNO MÖLDER *

BELIEF: MINIMAL NON-REALISM AND THE CONSTITUTIVE ROLE OF INTERPRETATION

SUMMARY: Pośłajko’s minimal non-realism about belief combines a deflationary account of belief’s existence, a denial that beliefs are natural properties, and a revisionary stance toward the current folk concept of belief. This paper argues that rejecting or revising the folk concept is not necessary for anti-realism and develops a framework that allows for a more nuanced articulation of realist and anti-realist options about belief. The second half of the paper defends interpretivism, another deflationary position alongside minimal non-realism. In particular, it clarifies the constitutive role of interpretation in what it takes to have mental states, and the relationship between canonical ascribability and the possession of mental properties.

KEYWORDS: belief, minimal non-realism, anti-realism, interpretivism, causation, constitution, folk psychology.

1. Introduction

Krzysztof Pośłajko’s *Unreal Beliefs: An Anti-Realist Approach in the Philosophy of Mind* is a meticulously argued defence of anti-realism about belief. It is clearly written and largely congenial to my own philosophical outlook. In my own work, however, I have generally tried to avoid using broad labels such as “realism” and “anti-realism” in order to minimise confusion (Mölder, 2010, p. 14). Since these terms can be understood in multiple ways, I have been wary

* University of Tartu, Department of Philosophy. E-mail: bruno.moelder@ut.ee. ORCID: 0000-0002-0136-9729.

of the risk that once one adopts, say, the ominous label of “anti-realism”, people assume that one incurs certain commitments that may come with the label even if the actual position is more intricate and nuanced. For this reason, I have found it more productive to work with related but more specific notions, such as “objectivity” or “recognition-dependence”. Of course, it is possible to define the term “realism” more narrowly from the outset, and this is what Pośłajko does. However, this approach carries the risk that certain intuitive aspects commonly associated with realist or anti-realist positions might be left out, and some formulations can then sound paradoxical, such as the claim that beliefs exist but are not real.

Let us begin with a brief summary of how Pośłajko construes the debate over beliefs. As already hinted, he distinguishes the question of the existence of beliefs from the question of the reality of beliefs. On the standard Quinean understanding, the reality and existence of something are the same: things that exist are real, and things that do not exist are not real. In contrast, Pośłajko adopts the deflationary view that existence questions are not substantial. Following Thomasson (2015), he regards answering existence questions as an easy matter—one only has to assess whether the application criteria for the relevant term are met. In this way, philosophical questions about the existence of beliefs can be set aside as trivial. The existence of beliefs no longer poses a deep philosophical problem. In fact, it is difficult to find anti-realist philosophers who would deny the existence of beliefs in this deflationary sense. How, then, should we distinguish between belief realism and belief anti-realism?

For Pośłajko, the main divide between realists and anti-realists concerning belief lies not in the question of whether beliefs exist but in whether they are real. He takes real properties to be natural in David Lewis’s (1983) sense of “sparse” properties. For Pośłajko (2024, p. 51), there are three hallmarks of natural properties: they form natural kinds, since they are objectively similar; they are referred to in causal explanations; and they stand in relations to fundamental or “elite” natural properties.¹ Pośłajko argues that beliefs fail to be natural in all three respects. The position he proposes with regard to beliefs is “minimal non-realism”. He articulates this position through the following five points (pp. 64–65):

1. Belief attributions can be true and often are; there are beliefs.
2. Attributed belief states share no objectively similar properties.
3. Beliefs are not natural kinds.
4. Beliefs do not play a causal role, and belief attributions do not have “genuine explanatory power in the context of causal explanations”.
5. Belief contents cannot be naturalised.

¹ Henceforth, I refer to Pośłajko’s (2024) only by page numbers.

I find myself in agreement with almost every item on this list. In Mölder (2010), I also adopt a deflationary approach to truth and regard the existence of beliefs *qua* pleonastic entities as ontologically unproblematic (as Posłajko also acknowledges on p. 36). I do not think that there are natural belief properties and have argued that mental states (including beliefs) do not form natural kinds (Mölder, 2019; 2024). My views on mental causation are a bit more nuanced, as I distinguish between causal efficacy and causal relevance. I do think that there is no efficient mental causation, but explanations in terms of mental states could be causally relevant in the sense of Jackson and Pettit (1990): if successful, causally relevant explanations indicate that there are some causally efficacious non-mental processes taking place in the brain. I do not find that the objections to mental causation that Posłajko develops in sect. 6.2 apply to this very light notion of causal relevance. Finally, I have also never truly believed in the prospects of a naturalistic account of mental content.

Given these broad agreements on non-realism about belief, are there any differences between us worth discussing? There are two such issues that I will focus on in this paper. First, I argue that Posłajko's more specific way of setting up his position, and particularly the revisionary character of his account of belief, creates problems for how his account should be labelled. The second point of difference concerns interpretivism, a view that Posłajko has criticised on various counts. He does not see the need to regard beliefs as being constitutively dependent on interpretation and prefers minimal non-realism, which is arguably simpler than interpretivism. In what follows, I address both points.

2. The Varieties of Realism and Anti-Realism

In more general terms, Posłajko (p. 1) presents his minimal non-realism about belief as a conjunction of three statements: (1) beliefs exist in a deflationary sense; (2) beliefs are not real; and (3) the folk-psychological concept of belief is mistaken and requires amelioration. While the first two claims are also covered in the five points above, the third clause is somewhat surprising and may be seen as in tension with the other claims. It is not entirely clear why denying that beliefs are real is not already sufficient for anti-realism about belief. Why bring the issue of the correctness of folk psychology to bear on this question? Let us explore this further.

My contention is that Posłajko's notion of realism (and, respectively, anti-realism) collapses two notions that should be kept separate for the sake of clarity. One is realism as a metaphysical view about the nature of an entity or a class of entities; the other is realism as a view about a particular discourse (cf. Demeter, 2009). Thus, to be a realist about beliefs is to take beliefs as real entities—if one is a physicalist, then beliefs would be physical entities, but one could also be

a non-natural realist about beliefs and regard them as, say, non-physical states.² But surely one way to be an anti-realist about beliefs is to deny that beliefs are natural properties in Lewis's sense. Note that nothing has been said about folk psychology yet. This is because realism/anti-realism about folk psychology can be separated from the question of realism/anti-realism about belief. Realism about the folk-psychological discourse is foremost a semantic thesis, which states that (i) mentalistic ascriptions in folk-psychological terms are truth-apt statements capable of representing mental facts (in some sense or other), and (ii) when we make these ascriptions, our aim is to represent these facts successfully.³ One way of being an anti-realist about folk-psychological discourse is to accept (i), but deny (ii). Mental fictionalism is such a position, as it takes folk ascriptions to be true or false, but denies that our aim in making these ascriptions is representational. Another kind of anti-realism about folk-psychological discourse would deny (i), namely, that folk-psychological sentences are truth-apt. This would amount to a kind of non-cognitivism about folk psychology.⁴

Metaphysical views on entities can be combined with various views on discourse. While entity-realism about beliefs is most naturally combined with realism about folk-psychological discourse (a Fodorian position), an entity-realist about beliefs can also be an anti-realist about folk psychology. This would be the position of someone who takes beliefs to be a natural kind and thinks that folk psychology is not a good guide to the nature of belief. She could hold that folk psychology mainly fulfils regulative or evaluative functions instead of a descriptive one. Realism about folk psychology combined with anti-realism about beliefs would amount to an error theory about folk psychology, which in the philosophy of mind, is exemplified by eliminativism. On that view, folk-psychological discourse is representational, but fails, as folk is wrong about the nature of belief. Another way to combine realism about folk-psychological discourse with belief anti-realism would be to read (i)–(ii) in a deflationary manner and deny the natural kindhood of beliefs. Anti-realism on both counts would

² In principle, entities could be “sparse” or “natural” in Lewis's sense without being naturalistic or physical.

³ These correspond to the first two elements of realism in Wright (1995, p. 198), who also adds a metaphysical clause (see also Drayson, 2022, p. 148; Mölder, 2022, p. 321). However, the metaphysical view on entities and the semantic view on discourses are separable. This is made clear by Demeter (2009), who has served as an obvious inspiration here. Yet, while Demeter distinguishes between entity realism and discourse realism, he understands entity-realism in terms of ontological commitment, thus linking it to the question of existence. In the context of Pośłajko's book, entity realism can be given more substantial content in terms of natural properties.

⁴ For a recent work on mental fictionalism, see Demeter, Parent, and Toon's (2022). Pośłajko discusses non-cognitivism about belief on pp. 11–14. A recent and more sophisticated non-cognitivist approach, which allows belief ascriptions to be truth-conditional, is presented by Fernández Castro (2023a; 2023b). He defends an expressivist view of folk psychology, according to which ascribing mental states is not an attempt to describe the subject's mind but rather to express an evaluative attitude.

deny that beliefs are natural entities and would be either fictionalist or non-cognitivist about the folk-psychological discourse, depending on whether one rejects (ii) or (i).

As this discussion hopefully shows, one gets a more nuanced elaboration of the realism issue by distinguishing between realism about entities and realism about discourse. Given that these stances are independent, a realist about beliefs *qua* entities need not take a stance on the status of the folk-psychological concept of belief. A realist who takes beliefs to be natural kinds usually assumes a distinction between belief as an entity and our ordinary concept of belief, and does not care much about the latter. It is enough that the folk concept of belief sets the process of empirical study in motion. Demeter (2009, p. 66) describes this perspective on the folk notion in this way: “Folk psychology may provide initial contingent, and thus possibly false, descriptions of entities to start with, but it has no special importance beyond that”. Analogously, an anti-realist about belief who does not think that beliefs are natural properties is not bound to assume that the folk concept of belief is flawed (from a certain angle, my own position could be characterised in such terms).

Posłajko’s treatment of folk-psychological concepts actually includes both semantic (or epistemic) and metaphysical aspects. He distinguishes between the application criteria of a notion and its metaphysical commitments (pp. 123–128). The notion of belief has superficial application criteria: beliefs are normally attributed on the basis of information about how people behave and what they say. Information about processes in their internal cognitive system or brain is not needed. On the inferentialist conception that Posłajko adopts, the metaphysical commitments of a term are fixed by its inferential relations, and according to him, these commitments are substantial. Namely, when we attribute a belief to someone, we are thereby licensed to infer that the person has an internal state that causes her behaviour.

Posłajko finds these commitments to internal causal states flawed and proposes revising our ordinary concept of belief. This revision should replace the current substantial concept of belief with a superficial one (p. 131). However, Posłajko does not elaborate on how exactly this revision should be effected. He notes, incorporating Haslanger’s (2012) distinction between manifest and operative concepts, that only our manifest concept of belief—that is, the concept we can explicitly articulate—involves metaphysically substantial commitments, whereas the operative concept of belief, the one we use implicitly, is shallow (p. 132). Yet even if we assume that operatively people are already on the right track, it is still unclear how to get from a philosophical critique to actually getting people to change the way they think about beliefs (and other mental states with similar substantial commitments).

Given that Posłajko incorporates the revision of the ordinary concept of belief into his approach, his theory acquires a dual and shifting character—his stance on the overall concept of belief is very different before and after the revision. Before the revision, the notion of belief carries substantial commitments,

and thus current belief attributions must be false. There are no such things that would satisfy the substantial commitments of the folk concept. Since the revision has not yet taken place, Pośłajko's stance currently amounts to a version of eliminativism.⁵ After the revision, however, when only the superficial notion of belief is retained, belief ascriptions can be true in this minimal sense. He writes:

Thus, it might indeed be the case that belief attributions might be seen as substantially false if we treat them as employing the substantial notion of belief; however, the same belief attributions might be treated as being minimally true if we saw them as employing the new thinner notion of belief. (p. 141)

The claim that the folk concept of belief carries substantial but flawed metaphysical commitments provides Pośłajko with a way to justify why revising the folk concept is part of his construal of minimal non-realism. He can be read as arguing that non-realism about belief requires not only rejecting the metaphysical reality of beliefs but also recognising that the concept itself misrepresents the world (cf. p. 136).

While acknowledging this, I still find that separating entity realism from discourse realism allows for a more nuanced representation of positions, including those that do not burden folk psychology with strong commitments. From this point of view, one could reject belief realism without having to say that the folk concept is mistaken or needs fixing, since its commitments are already minimal and it might still serve its role even if the world does not contain beliefs as natural entities.

Pośłajko seems to assume that a match with the folk concept is required for realism, and a mismatch for anti-realism. For example, he notes that, in addition to showing that beliefs are not natural properties, one also needs "some sort of commitment to the idea that the folk notion of belief is in the wrong with respect to mental reality" (p. 117). In contrast, from the opposing point of view, if it turns out that "the folk concept of belief matches mental reality" (p. 152), then the position should be regarded as the realist one.

However, it seems to me that a match between the folk concept and reality is not always required for belief realism, and that anti-realism does not always imply a mismatch. Let us presume that belief could have either substantial metaphysics, that is, be a natural property, or shallow metaphysics, that is, not be a natural property. Folk psychology, or more particularly the folk concept of belief, could carry either substantial or minimal commitments. Then we get a two-dimensional matrix with four positions.⁶ (1) Belief is a natural property

⁵ Pośłajko would no doubt object to the "eliminativist" label, since he wants to preserve the term "belief" and does not claim that beliefs do not exist. However, what is preserved after the revision is only the shallow shell of a concept, while its inferential core has been eliminated (see also pp. 136, 172).

⁶ A structurally similar two-dimensional matrix is used by Mun (2016) to distinguish positions in theories of emotion according to their metaphysical commitments and views

and folk psychology of belief is substantial. There is a match between folk-psychological notion and the metaphysics of belief. This is a position of robust realism about belief. (2) Belief is not a natural property, but folk psychology carries substantial commitments. In that case, folk psychology is mistaken about belief. This corresponds to Pośłajko's view of belief before the conceptual revision and amounts to a kind of error-theory or anti-realism about belief of an eliminativist variety. (3) Belief is a natural property, but folk psychology has minimal commitments. As folk psychology is superficialist, the folk concept of belief does not capture the rich metaphysics of belief—there is a mismatch between the concept and mental reality. (4) Belief is not a natural property, and folk psychology is austere. This would be Pośłajko's view of belief after the conceptual revision.⁷ In this case, the minimal folk concept matches the minimal reality of belief.

As we can see from this list of positions, there are two ways in which folk psychology could be in error given a mismatch between the folk concept of belief and reality, but not both cases are versions of anti-realism. While (2) is anti-realist about belief, (3) could still count as a version of realism about belief, since it takes beliefs as natural properties not fully captured by the folk concept. Also, of the two cases in which the folk concept matches reality, only (1) would count as realism in the straightforward sense, while (4) could be labelled either as anti-realism (if one puts weight on the idea that belief is not a natural property) or as a version of minimal realism (if one follows Pośłajko and proceeds from the “match” idea).

Compared with the earlier distinction between entity realism about beliefs and realism about folk psychology as discourse, this two-dimensional matrix focuses on the match between the metaphysics of belief and folk commitments. It contests Pośłajko's assumption that realism requires a match between the folk concept and reality, while anti-realism implies a mismatch. The matrix shows that this is not necessarily the case. This provides further support for the point that Pośłajko's inclusion of the third clause in his minimal non-realism is not strictly required by the rejection of belief realism.

Here, then, lies the tension in Pośłajko's position. The third clause (that the folk concept of belief embodies false assumptions) is the reason why Pośłajko's account is revisionary about belief. But in being so, he holds different views on the concept of belief at different times (pre- and post-revision). Basically, he moves from a version of eliminativism (position 2) to a version of realism (posi-

on the reference of the ordinary term “emotion”. The present matrix applies a comparable strategy to belief and folk psychology in order to distinguish between types of realism and anti-realism.

⁷ Pośłajko discusses this option on p. 152, and concedes that it would be “straightforwardly realistic”, but dismisses it, since he regards the minimal commitments (as assumed by Bogdan, 1993; Horgan, 1993) as still too strong. However, if we assume that the minimal folk notion of belief is exactly as superficial as Pośłajko takes it to be, then his post-revision view falls into category (4).

tion 4—“realism” in his terms). As things stand, he is still a kind of eliminativist about the metaphysically loaded concept of belief, and will be a realist in the future, provided the revision succeeds.⁸ In fact, his overall view is not easy to name, as its parts are in tension. Poślajko calls his position “minimal non-realism”, and the label “non-realism” is warranted by his stance that the current conception of belief is faulty (p. 152). However, if we proceed from his assumption that a match with reality is required for realism, then post-revision he could no longer be called a “non-realist”, as the future revised concept of belief would match reality and would no longer be faulty. Conversely, if the rejection of beliefs as natural properties is enough for anti-realism, as I suggested, then he could keep the “anti-realist” or “non-realist” label throughout. But in that case, the third clause about the folk concept being faulty is not a necessary part of anti-realism.

It is not entirely clear what commitments are left for the concept of belief after the revision, according to Poślajko. Recall that he assumes a version of the inferentialist theory of meaning, on which the meaning of a term includes epistemic application criteria and a set of licensed inferences that embody its metaphysical commitments (pp. 126–127). Having set up the distinction between these two components of meaning, Poślajko describes an alternative linguistic community that employs *only* the superficial concept of belief and remarks that “perhaps it is the one we should be using” (p. 128). If this is his suggestion, then the revised concept of belief would incur not just fewer metaphysical commitments, but no commitments at all. In that case, the notion of belief would lack an inferential component, that is, applying the term “belief” would not license any inferences. One half of the meaning of the term “belief” goes missing after the revision. What remains is just the superficial shell of the application criterion, which “specifies that we are justified to ascribe beliefs to a given subject if the subject in question satisfies fairly weak behavioural criteria” (p. 127).

However, this leaves it rather mysterious why we ascribe beliefs at all. What purpose does belief ascription serve, even if the behavioural criteria are satisfied? When discussing why the wholesale elimination of folk-psychological vocabulary would not work, Poślajko writes (p. 169): “We need a way of describing humans and making sense of their actions that is easily acquired and not too demanding of our cognitive capacities”. Indeed—but making sense of people’s actions by ascribing beliefs to them presupposes that there is some kind of connection between those actions and beliefs. If this relationship were merely justificatory and non-causal, a deeper explanation of what made the action happen would still be required. If making sense of an action involves explaining why it occurred by reference to the agent’s beliefs, then, given the widespread view of

⁸ Poślajko could respond that since the operative concept of belief is already shallow, it need not be changed, and so his view on that concept remains stable. However, this does not resolve the tension. First, as we saw before, his stance towards the *overall* concept of belief clearly changes, as he moves from regarding belief ascriptions as false to regarding them as true. Second, a more convincing case needs to be made for showing that, in the case of the notion of belief, the manifest and operative concepts do really come apart.

folk-psychological explanation, the relationship would have a causal character. In fact, Pośłajko (pp. 118–123) stresses that, pace Curry (2018), on the current folk conception, beliefs really are causes. It is precisely this causal character that he proposes to remove from the concept of belief. He is then faced with the challenging task of explaining how the revised, superficial concept of belief can be compatible with the explanatory aim of belief ascription, which is to make sense of our actions.

Ascertaining the commitments that people actually incur when they ascribe beliefs might well be a task that cannot be settled from the armchair, so the following remarks are just my opinion. I think that the commitments that Pośłajko saddles folk psychology with are too strong. As already noted, he assumes that the folk are committed to beliefs being internal representational states that cause behaviour. In particular, he writes that the inferential norms comprising the concept of belief license us “to infer from the claim ‘*x* believes that *p*’ to the claim that ‘*x* is in some definite internal state which is causally responsible for *x*’s behaviour” (p. 127). It is hard to believe that a quotidian concept involves so specific inferential licences to internal mechanisms. If there are inferential links in folk psychology, those usually connect the belief to other mental states and to the state of affairs that the belief is about. Pośłajko concedes that people cannot articulate this rule but can follow it implicitly. However, this latter claim requires independent evidence—evidence that would support such strong commitments rather than merely minimal ones.

It is more plausible to assume that folk commitments about belief are rather minimal—that people have beliefs, which are about something, and that they play a causal role in relation to actions.⁹ There is no need to assume that the folk are committed to the existence of mental representations or concrete internal mental states. It is the task of philosophy to provide an account that respects folk commitments while explaining them in an ontologically unproblematic way, rather than dismissing them or treating them as misguided. On the interpretivist view that I favour, the possession of beliefs, understood in this minimal sense, can be accounted for in terms of people being interpretable as having these contentful beliefs, and these beliefs could be causally *relevant* even if they are not causally *efficacious* internal states (Mölder, 2010, Chapter 7). That would also explain why we attribute beliefs to people—it is because we find them causally relevant to what they do.

⁹ Given the availability of an interpretivist gloss, this could be understood even more minimally than Horgan’s (1993, p. 289) austere folk psychology, which holds that “the states it posits are realized in humans, somehow or other, by states that causally interact in ways that render the generalizations of FP true”.

3. The Bearable Lightness of Interpretation

This brings us to interpretivism, the view that having mental states is a matter of interpretation.¹⁰ In its classic forms, the view was developed by Dennett (1987) and Davidson (1973; 1974). In the remaining part of the paper, I focus on Poślajko's critique of my own version of interpretivism and offer a long-overdue response to it.

While interpretivism is briefly discussed in Poślajko (2024), his stance in the book rests on a more thorough analysis of the view and of my own specific position in his earlier work (Poślajko, 2020). In that paper, he construes interpretivism as the position that interpretation is constitutive of propositional attitudes. Following Byrne (1998) and Kriegel (2010), he unpacks this as the claim that facts about mental states are judgement-dependent, that is, dependent on what mental states the interpreter would interpret one as having.¹¹ Poślajko also notes that adopting a deflationary take on the existence of mental states is the best way for an interpretivist to avoid succumbing to mental fictionalism. Interpretivism is thus seen as consisting of two main claims that are independent: "deflationary ontological realism about the existence of mental states, and an additional metaphysical assumption that mental states are judgement-dependent" (Poślajko, 2020, p. 717). He agrees with the deflationary approach but finds interpretivism "an unnecessarily convoluted version" of deflationism (Poślajko, 2020, p. 715).¹²

Poślajko (2020, pp. 718–719) argues that subscribing to judgement-dependence leads to troubles with the epistemology of mental states. Specifically, the issue concerns the capacities of the interpreter whose judgements are constitutive of the possession of mental states. Specifying these capacities gives rise to a dilemma. If the interpreter has ordinary capacities, then the account cannot accommodate mistaken attributions—for if ordinary interpretations are constitutive, there is no room for error. If, however, the interpreter is idealised, then she knows too much, and her judgements will be remote from those that people would routinely give. Both horns of this dilemma are rather unattractive. Poślajko (2020, pp. 719–720) acknowledges that my version of interpretivism does not encounter the dilemma; however, he argues that my account rejects

¹⁰ In this section, I write broadly about mental states and set aside the differences between propositional attitudes (such as belief) and other types of mental states.

¹¹ I should note that Byrne (1998, p. 204; see also p. 206) does not take the judgement-dependent analysis to express the idea of one set of facts being constituted by another. He formulates the analysis as a biconditional and points out that both sides stand on equal footing, with neither being more basic.

¹² In another paper, Poślajko (2023, p. 928) repeats this point, claiming that interpretivism incorporates "a problematic component in the form of the claim about attitudes being constituted by the process of interpretation". He regards this component as "an unnecessary complication which makes the interpretivist version of the general deflationary/superficial approach far less palatable than the original generic dish".

judgement-dependence and thus becomes indistinguishable from a general deflationist approach, losing its interpretivist flavour.

Let us examine this claim in more detail. To begin with, I will briefly summarise my own position, the ascription theory (Mölder, 2010; 2017). It is an interpretivist account of the possession conditions of mental states, that is, of what it takes to have them. My claim is that people have those mental states that are canonically ascribable to them. Canonical ascription sets a standard for correctness, so not every ordinary ascription would count as true. A canonical ascription is one that integrates information from various ascription sources in a coherent way and would not require revision if it were actually made. The second clause relaxes the condition to a range of acceptable ascriptions that fall short of the ascription an ideal interpreter would make. In this way, the canonical ascription is, in principle, within the reach of ordinary people. The mental properties whose possession is fixed by canonical ascription are conceived in a deflationary manner. These properties are pleonastic: metaphysically, to have them amounts to nothing more than being describable by mental predicates.

Avoiding the aforementioned epistemological issues does not hang on the rejection of judgement-dependence. The notion of canonical ascription does the work of evading the dilemma. Nevertheless, Pośłajko argues that my account fails to make mental properties dependent on interpretation, since I reject judgement-dependence, which he defines as the view that “it is both necessary and sufficient that one is interpreted as possessing mental states in order to possess them” (Pośłajko, 2020, pp. 719–720). However, I do not think that I reject judgement-dependence in the relevant sense, once it is suitably clarified.

First, I never rejected the view that interpretability is necessary for the possession of mental states, as I regard interpretability as ineliminable from the possession condition. Second, the answer to the question of whether interpretation is sufficient for the possession of mental states depends on how much is packed into the notion of interpretation or ascription (Mölder, 2010, pp. 77–78). The view I reject is the one according to which the act of ascription alone suffices to constitute mental state possession, as such an act could be whimsical, showing no regard for the evidence. I hold that the contribution of the world and of the subject of ascription is also important. This contribution enters the picture through the ascription sources, which include information about the subject’s behaviour, the stimuli affecting her, her personal background, and other mental states ascribable to her. But when the notion of interpretation is understood more broadly, as including the ascription sources and the ways of making them coherent in accordance with folk principles of interpretation, then such interpretation would be both necessary and sufficient for the possession of mental states. On this broader reading, the conditions that make interpretation appropriate, such as coherence with the evidence, must also be included in order for interpretation to be sufficient. The way I phrased the connection between possession and interpretation (for ascription is a kind of interpretation)—“to have a belief that *p* is to have it canonically ascribable”—makes it clear that canonical ascribability is

both necessary and sufficient for belief possession (Mölder, 2010, p. 171; also quoted by Pośłajko, 2020, p. 720).

I have also argued that mental states are recognition-dependent, and this can be taken as a reason for offering an interpretivist account of their possession (Mölder, 2010, pp. 158, 202–204; Mölder, 2024). The idea is that a state qualifies as mental only when it is specifiable in mental terms. However, this specification is not intrinsic to the state itself. Rather, the applicability of mental terms depends on various external factors (as captured by the ascription sources), and interpretation is required to determine which mental terms are appropriate. Thus, possessing a mental state depends on the interpreter's ability to recognise someone as being in a state with a specific mental description. Recognition-dependence can be seen as a form of judgement-dependence.

Still, I have not yet answered the concerns raised by Pośłajko (2020). There are two worries. One is the “no-distinction challenge”, as interpretivism seems not to differ from other deflationary accounts of the mind. The other is the worry that my account does not explain how exactly mental states such as beliefs depend on interpretation. I will address these concerns in turn.

The no-distinction challenge asks how the ascription theory differs from generic deflationism, which states that an object *O* exists if and only if the application conditions of the term “*O*” are met. What is the distinctive role for interpretation to play in the case of the mind? Pośłajko (2020, p. 721) also points out that one could be a deflationist about boulders. What is peculiar about boulders is that they do not form a natural kind, but are just a category of large rocks, whose diameter exceeds 25,6 cm. One could thus claim that boulders exist, since the application conditions of “boulder” are fulfilled, and that the property of being a boulder is pleonastic. This does not commit the boulder-deflationist to the claim that boulders are constituted by our acts of classification. Could an interpretivist maintain that beliefs (and other mental states) differ from boulders? If not, then it cannot be said that either is dependent on or constituted by interpretation.

Indeed, to avoid collapsing into general deflationism, interpretivism needs to maintain its distinctive interpretive element. It is also not a good idea to be a global interpretivist. Instead, its range of application should be kept more restricted. For if everything depends on interpretation, then interpretation-dependence would not be distinctive anymore. I do not object to applying interpretivism to certain other social, non-mental entities that may require some form of canonical determination, provided we care enough about clearly delineating their boundaries. This does not mean that interpretivism generalises to boulders. A distinction can be drawn between fixing the application conditions of a term and fulfilling those conditions. In some cases, fixing the application conditions involves conventional elements, but once established, their application is automatic and depends solely on empirical facts. In other cases, however, even after the application conditions are fixed, fulfilling them requires interpretation. Thus, even if being a belief and being a boulder are both pleonastic properties, the relevant distinction lies between cases where meeting the application conditions

of a term is straightforwardly empirical (as in the case of boulders) and cases where meeting them is sensitive to interpretation, as with mental states.

Hence, I would stick to the claim that the possession of mental states is constituted by interpretation, and this is what makes interpretivism distinct from other deflationary views. The crucial point is that interpretation is required to fulfil the application conditions. A canonical interpretation seeks the most coherent ascription (within a reasonable range) given the ascription sources and their relevance. This is based on the ascriber's competence, which is uncodifiable. It cannot be made fully explicit, and this is why the interpreter is ineliminable. In other words, a proper application of a mental term *is* already an interpretation.

It is crucial to stress that mental state ascription is not simply a matter of applying fixed criteria in a straightforward manner. Making an ascription requires constructing a coherent arrangement of information from various ascription sources. This arrangement is unique in each case of applying a mental term, such as "belief". This contrasts with cases that do not involve interpretation—cases in which a term applies automatically once its application conditions are fixed and empirically satisfied.

Posłajko's second challenge takes the form of an invitation to explain the sense of dependence that holds between interpretation and mental states in my view: "what is the sense of the notion of metaphysical dependency that is adopted by the proponent of interpretivism once the idea of judgement-dependence, in the sense which requires an actual interpretation to take place, is rejected?" (Posłajko, 2020, p. 722). This challenge is motivated by the fact that my account does not require actual ascription but instead settles for interpretability or ascribability. Posłajko's worry is that a merely possible interpretation is metaphysically too light to support the constitution claim. I have already explained that I do not reject the idea that interpretability (considered broadly) is necessary and sufficient for the possession of mental states. The question now is whether it is strong enough for constitution.

I believe it is. Remember that mental properties are pleonastic. Having pleonastic properties is a lightweight affair, and ascribability, although light, is strong enough to secure what it takes to have pleonastic properties. Talk in terms of having pleonastic properties can be replaced by talk about the applicability of predicates. Thus, the issue boils down to the applicability of mental predicates. In my reply to the first challenge, I already explained the constitutive role of interpretation in this. To ensure that mental predicates are applicable to someone, (mere) canonical ascribability is exactly at the right level of strength.

Pleonastic facts, such as having a pleonastic property, do not require constitution in any more substantial metaphysical sense. What is needed is simply that it is possible for an individual to fall under a specific mental description, given the right conditions. Not just any mental description will do; the conditions of canonicity must be met. In every case where someone is describable in mental terms, it is in virtue of the pleonastic fact that a relevant mental state is canon-

ically ascribable to them. In short: lightweight facts call for a lightweight form of constitution.

If my responses to Pośłajko's challenges are successful, this will have implications for his overall argumentative strategy. He rejects the deflationary version of interpretivism in favour of his minimal non-realism, on the grounds that the former incorporates the claim of constitutive interpretation-dependence. But if I am right that any ascription of mental states or application of mental terms already involves interpretation, then interpretability—although lightweight—must also be an inescapable component of minimal non-realism.

4. Conclusion

This paper has examined Pośłajko's minimal non-realism about belief, which combines three claims: that beliefs exist in a deflationary sense, that beliefs are not natural properties, and that the folk concept of belief is mistaken and requires revision. I argued that the third claim introduces unnecessary complications into his account, to the extent that it is unclear whether the position still qualifies as anti-realist. The more nuanced framework provided in the paper shows that one can reject belief realism without committing to the view that the folk concept is mistaken. This alternative view does not regard beliefs as natural properties but preserves the explanatory usefulness of folk psychology.

I disputed Pośłajko's assumption that the folk concept of belief involves strong metaphysical commitments to internal causal states. When he rejects these commitments, it becomes mysterious why people attribute beliefs at all if they bear no causal relevance. Instead, I propose that the shallow application criteria for the ordinary concept of belief are largely consistent with its minimal metaphysical commitments. This speaks in favour of a deflationary account of belief that retains its explanatory role in everyday discourse without invoking substantial metaphysics.

In the second part of the paper, I responded to Pośłajko's challenges to interpretivism. I clarified my stance on the judgement-dependence of mental states on interpretation. I argued that interpretivism does not lose its distinctive flavour even within the deflationary framework. Interpretability is constitutive of mental state possession, because only through interpretation can one fix which pleonastic mental properties the subject has. The condition for possessing a pleonastic property is precisely the condition under which it can be ascribed. Finally, I pointed out that canonical ascribability is not too lightweight to ground the constitution claim; on the contrary, it is exactly right for securing that someone is describable in such-and-such mental terms. So I think that interpretivism as a position is doing quite well—any reports of its death (as in Pośłajko, 2023, p. 929) are exaggerated, and the use of obituary discourse is premature.

REFERENCES

- Bogdan, R. J. (1993). The Architectural Nonchalance of Commonsense Psychology. *Mind and Language*, 8(2), 189–205.
- Byrne, A. (1998). Interpretivism. *European Review of Philosophy*, 3, 199–223.
- Curry, D. S. (2018). Beliefs as Inner Causes: The (Lack of) Evidence. *Philosophical Psychology*, 31(6), 850–877.
- Davidson, D. (1973). Radical Interpretation. *Dialectica*, 27, 313–328.
- Davidson, D. (1974). Belief and the Basis of Meaning. *Synthese*, 27, 309–323.
- Demeter, T. (2009). Two Kinds of Mental Realism. *Journal for General Philosophy of Science*, 40(1), 59–71.
- Demeter, T., Parent, T., Toon, A. (Eds.). (2022). *Mental Fictionalism: Philosophical Explorations*. London; New York: Routledge.
- Dennett, D. C. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.
- Drayson, Z. (2022). What We Talk About When We Talk About Mental States. In T. Demeter, T. Parent, A. Toon (Eds.), *Mental Fictionalism: Philosophical Explorations* (pp. 147–159). London; New York: Routledge.
- Fernández Castro, V. (2023a). An Expressivist Approach to Folk Psychological Ascriptions. *Philosophical Explorations*, 27(1), 86–105.
- Fernández Castro, V. (2023b). Folk Psychology Without Metaphysics: An Expressivist Approach. *South African Journal of Philosophy*, 42(2), 128–143.
- Haslanger, S. (2012). *Resisting Reality: Social Construction and Social Critique*. New York: Oxford University Press.
- Horgan, T. (1993). The Austere Ideology of Folk Psychology. *Mind and Language*, 8(2), 282–297.
- Jackson, F., Pettit, P. (1990). Program Explanation: A General Perspective. *Analysis*, 50(2), 107–117.
- Kriegel, U. (2010). Interpretation: Its Scope and Limits. In A. Hazlett (Ed.), *New Waves in Metaphysics* (pp. 111–135). London: Palgrave Macmillan.
- Lewis, D. K. (1983). New Work for a Theory of Universals. *Australasian Journal of Philosophy*, 61(4), 343–377.
- Mölder, B. (2010). *Mind Ascribed: An Elaboration and Defence of Interpretivism*. Amsterdam: John Benjamins.
- Mölder, B. (2017). Mind Re-Ascribed. *Studia Philosophica Estonica*, 10(2), 55–104.
- Mölder, B. (2019). Vaimuseisundid pole loomulikud liigid. [Mental States Are Not Natural Kinds]. In R. Altnurme (Ed.), *Humanitaarteadused ja kunstid 100-aastases rahvuslikoolis* (pp. 173–185). Tartu: Tartu University Press.
- Mölder, B. (2022). Interpretivism and Mental Fictionalism. In T. Demeter, T. Parent, A. Toon (Eds.), *Mental Fictionalism: Philosophical Explorations* (pp. 320–339). London and New York: Routledge.
- Mölder, B. (2024). Mental Kinds and Practical Realism. In A. Mets, E. Lõhkivi, P. Mürsepp, J. Eigi-Watkin (Eds.), *Practical Realist Philosophy of Science: Reflecting on Rein Vihtalemm's Ideas* (pp. 143–159). Lanham, MD: Rowman & Littlefield (Lexington Books).

- Mun, C. (2016). Natural Kinds, Social Constructions, and Ordinary Language: Clarifying the Crisis in the Science of Emotion. *Journal of Social Ontology*, 2(2), 247–269.
- Posłajko, K. (2020). Can Deflationism Save Interpretivism? *Philosophia*, 48(2), 709–725.
- Posłajko, K. (2023). Folk Psychological Models and the Metaphysics of Belief. A Reply to Curry. *Philosophia*, 51(2), 919–931.
- Posłajko, K. (2024). *Unreal Beliefs. An Anti-Realist Approach in the Metaphysics of Mind*. London: Bloomsbury Academic.
- Thomasson, A. L. (2015). *Ontology Made Easy*. Oxford: OUP.
- Wright, C. (1995). Can There Be a Rationally Compelling Argument for Anti-Realism About Ordinary (‘Folk’) Psychology? *Philosophical Issues*, 6, 197–221.