But Is It Interpretivism?

Sören Häggqvist

Department of Philosophy, Stockholm University

In this critical notice I raise a couple of questions concerning Mölder's ambitious metaphysics, aimed at underpinning his Ascription Theory. I argue that some of the points he takes to depend on this metaphysics are in fact independent of it. I further question whether the relation between the mental and the physical is quite so unlike relations between special science entities and physics as Mölder suggests. Finally I relate Mölder's Ascription Theory in very compressed form and suggest that although its loosening of the strictures on what evidence an ascriber of mental attitudes may avail herself of, it is not clear that the theory can really do without rationality considerations of the sort emphasised by Davidson and Dennett—at least if it is to count as a species of interpretivism.

Keywords: interpretivism, metaphysics, mental states, coherence, holism, Davidson, Dennett

1. The metaphysics

This very rich book offers the most detailed, systematic, and sustained defense to date of interpretivism, as a general position in philosophy of mind. It also proposes a novel elaboration of the view: a theory Mölder calls *ascriptivism*. Mölder's theory is importantly different from the views of the position's founding fathers, Dennett and Davidson.

Unlike Dennett, for instance, Mölder takes his metaphysics seriously. A hefty second chapter lays out the ontological groundwork. The many points and distinctions made here are not idle but do real work in subsequent discussions. A central theme, to which Mölder returns over and over throughout the book, is his insistence on a distinction between pleonastic entities versus what he calls "the natural basis" (Mölder 2010, 21).[1] The latter is what really exists. The former are what we get through what (Schiffer

Corresponding author's address: Sören Häggqvist, Department of Philosophy, Stockholm University, 106 91 Stockholm, Sweden. Email: soren.haggqvist@philosophy.su.se.

[1] Unless indicated otherwise all references with just a page number will be to (Mölder 2010).

Studia Philosophica Estonica (2017) **10.2**, 8–17
Published online: October 2017

Online ISSN: 1736–5899
www.spe.ut.ee
http://dx.doi.org/10.12697/spe.2017.10.2.03

2003) calls "something-from-nothing transformations" like the following inference (22):

> Stephen is clever. Therefore, Stephen has the property of being clever.

Applying existential quantification to the conclusions of such arguments, one gets a quickly swelling ontology of properties,[2] whose place in the natural world and whose relations to other properties may become philosophical quandaries. Mölder does not object to the inference, but to the inflation of ontology that results from taking it at face value: he calls talk of properties as though they were real existents, or capturing description-independent natures of things, "inflationary" (20, 23). And he holds that very often, such talk is misguided, since the "properties" discussed are merely pleonastic. For the distinction between natural and pleonastic entities, Mölder prefers to talk of different "strata" rather than "levels" (31). This seems wise, since, as he notes, levels-talk often involves relations between entities which in his view belong to the same stratum.[3]

Mölder is not committed to any particular theory of natural existents. But he says that the qualifier 'natural' is used to signal allegiance to a broadly naturalist metaphysics (21); that "what are commonly taken to be natural kinds surely belong to the natural basis" (23); and that "the natural basis consists of ontologically basic *physical* entities and their configurations" (33, italics in original). He also hints that individuation of such ontologically basic entities may go by causal powers (27), making foundational metaphysics a project informed by and perhaps continuous with science (28).

Mental events, states, and properties are not natural existents on this view: they belong to the pleonastic stratum. Moreover, neither mind itself nor any psychological category are natural kinds with a hidden nature open to a posteriori discovery (143).[4] This is the basis of Mölder's claim that possession of mental states is inextricably tied to our everyday psychological concepts, as these are reflected in folk psychology.[5] Note that this claim is

[2] Provided one also takes the quantification to have ontological import. Mölder focusses on generation of pleonastic properties, but as he mentions, similar mechanisms may be used to generate e.g. facts, particulars and other apparent denizens of ontology.

[3] For instance, both usual macro-micro distinctions and personal-subpersonal distinctions "might be confined only to the pleonastic" (31, fn. 17); conversely, "natural properties can form hierarchies of levels" (31, fn. 17).

[4] Mölder here (143–144) invokes Ellis (2001), whose demanding criteria for natural kindhood—observer-independence, categorical distinctness and instrinsicality—are obviously not met by mental states. He also invokes Dennett (1994), but does not comment on the natural reading of Dennett as questioning the very notion of natural kinds invoked by e.g. Putnam (1975).

[5] It should be noted that Mölder does not take folk psychological platitudes to exhaust or implicitly define mental concepts, whether in the way suggested by Lewis (1972, 1994) or

intended to fall short of saying that mental states themselves *are* just what our everyday concepts take them to be. Ascription theory is offered only as a theory of what it is to have mental states, and not of their nature. It is one thing to specify possession conditions, another to give an account of what the "constitution and individuating features" of a mental state are (12). While not true of pleonastic entities across the board, in the case of mental states, there is nothing more to be told about this than what goes into the meanings of mental terms, i.e. into mental concepts. But ascription theory is not a theory of meaning and does not yield one (and this is a chief difference with Davidson). Rather it presupposes that mental vocabulary is already in place, equipped with public lexical meanings. The same goes, in fact, for the meaning of every term. While ascription theory does specify application conditions for mental predicates—in other words, for possession of pleonastic mental properties—these specifications presume stable meanings of words in general; hence, there is no special problem of mental content or intentionality: "in an ascription such as 'John believes that there is a dog,' the content of John's belief is specified by terms that already have a common meaning" (151). Mölder is optimistic about the prospects of such a prior, general theory of meaning, even granted that such a theory cannot, if it is to be non-circularly joined with ascription theory, invoke (a thinker's being in) mental states (152).

However, if mental entities are on the whole gotten on the cheap through something-from-nothing transformations, why is not their "nature" exhausted by the application conditions of the corresponding predicates? After all, ascription theory aims to tell what it takes to belong to the extension of these predicates. Why hold that the nature of mental states—given that they are merely pleonastic—lies in something as intensional as (non-mental, public) concepts? Indeed, what is the point of talking of a "nature" here?

To the first of these questions, perhaps an answer might be that the ascription theory actually does not provide criteria for belonging to the extensions of mental predicates, and that there is some slack between correct and warranted applications/attributions, even though reference to attribution is partly constitutive for actual inclusion in the extensions. However, Mölder does hold that *canonical* ascriptions do determine the mental facts, in the sense of determining which mental attributions are correct. Yet, possession conditions do not line up with individual mental states but provide only "a general scheme that explicates the relations between mental phenomena and their possession"; mental states "are not individuated by possession conditions" (171). Perhaps this provides a basis for a different answer to the first

Braddon-Mitchell and Jackson (1996). But he argues that this is a problem for functionalism, and not for him (133–138).

question. But I am not sure about this, nor what his response to the other two questions is. Despite its overall clarity and analytic sophistication, there are points where, as in this case, the book did not make it entirely clear (to me) what Mölder's position is.

Another instance of this is the question whether there is anything special about the relation between the mental and the physical, as compared to the relation between various special sciences and physics. This question crops up in most discussions about reduction, autonomy, and causation in recent decades' philosophy of mind. Mölder makes a number of important points pertinent to this issue in chapter 2. He rejects supervenience, rightly I believe, as needlessly introducing more problems that it solves (even if one gets the definition of it technically right and unambigious). He suggests, and I agree, that assuming supervenience amounts to postulating an asymmetric dependence-relation which itself seems in need of more explanation than it can provide (35). I also accept his rejection of even weak supervenience of the mental on neural properties. But this point, like his rejection of token identity theories, seems to depend straightforwardly on intepretivism, rather than on the metaphysical distinction between a natural basis and pleonastic entities (or between the deflationary and inflationary). Concerning token identity, Mölder writes: "On the deflationary view … the application of a mental description is nothing of the sort that needs to be exhausted by or based on a physical particular, nor is it intelligible how it could be." (54) But this too seems like something an interpretivist should hold irrespective of her metaphysics.

Where the metaphysical machinery is really made to do some work is in Mölder's rejection of functionalism. In an unusually ambitious discussion of realisation (57–70), Mölder rejects role functionalism on the grounds that role functional properties cannot be distinct from their realisers. And he seems here to suggest that there is a peculiar problem besetting any attempt to account for a realisation relation between natural base properties and mental properties, a problem distinct from other cases of seemingly multiply realisable properties (65–66). But here one may ask why the mental case is considered to be so special. Given the wide reach of the arguments Mölder uses against (role) functionalism, is there any room for functional properties at all? Perhaps Mölder would reply that since natural properties should be individuated by causal powers, the natural base itself consists entirely of functional properties (in a sense), but there seems little room for multiple realizability at the natural stratum.

Conversely, when Mölder says that asciption theory is physicalist "in the sense of Braddon-Mitchell and Jackson (1996)" and so reckons with an ontology only of "the entities recognised by the physical sciences and their

naturalist extensions" (33), it is natural to ask just why psychology, as Mölder clearly assumes, is not among these. An answer might presumably invoke a distinction between intensional and extensional contexts, but Mölder does not elaborate on this (this is one place where Davidson was actually clearer). And the question remains what to make of all the (presumably pleonastic) entities apparently reckoned with in those extensions of the physical sciences.

2. The theory

Mölder's theory holds that the possession condition for, say, "…believes that it is raining" is given by what is canonically ascribable to a subject X, according to the schema (175):

> Possession condition for X *f*'s that *p*: *f*'s that *p* is canonically ascribable to X.

Canonical ascription, in turn, is defined this way, where '*As*' abbreviates an ascription statement of the form 'X *f*'s that *p*' (175):

> *As* is a *canonical ascription statement* $=_{df}$ *As* is a statement in approximation to the maximum coherence with the ascription sources $c_1, \ldots, c_4$ and if *As* were in fact ascribed, it would not require revision.

There are, according to Mölder, four main sources of ascriptions (hence the numbering above): other contentful ascribable mental states; the subject's behaviour; the subject's environmental stimuli; and the subject's history or background (161–162). I will first make a few expository comments and then raise a couple of questions about these definitions and the theory they are used in.

First, the possession condition does not require that any actual ascription takes place. Ascribability is a modal notion, although Mölder stresses that he wishes to make an actual ascription, as it were, nearby possibility when it does not occur. Second, the fact that an ascription does not require revision does not imply that it is unrevisable (174). It just means that it is stable or settled. Third, although she is not mentioned explicitly in the definiens, the definition of a canonical ascription statement presupposes or makes implicit reference to an ascriber (Mölder sometimes talks about "canonical ascribers"). This is not a postulated ideal figure, but anyone of us much of the time in our folk-psychological dealings with each other. Fourth, although anyone can occupy the ascriber's role, we may not know when our ascriptions fulfil the requirements for canonicality. Fifth, insofar as there may be several approximations, given the evidence from the ascription sources, to maximum coherence, there may also be several valid ascriptions;

but if even a canonical ascriber is faced with a choice between incompatible ascriptions, this is a reason for her to not (yet) accept any of them as canonical (184). Whatever differences remain between canonical ascriptions will, Mölder holds, perhaps lead to plurality (of mental state possession) rather than indeterminacy (183). Sixth, and crucially, the theory requires coherence (or approximate maximum coherence) among the ascriber's pieces of evidence, but *not* among the attributee's beliefs or other attitudes.

It is a sympathetic streak in Mölder's theory that he allows an interpreter not just common sense but also the wisdom about others encoded in folk psychology. Mölder eschews imposing on interpreters knowledge of how to construct T-theories from a narrow evidence base of holding-true, as well as any canons of decision theory and probability calculus. The evidence allowed for intepretation is wide-ranging. But while Davidson and Dennett, arguably, have a story about what it is to be a believer, Mölder simply defers to folk psychology on this issue. Thus, in discussing the "liberalism objection" exemplified by Block's Blockhead example, he writes: "since mental predicates are ascribed from the folk-psychological standpoint, this perspective has already circumscribed the range of potential subjects of interpretation" (193). This is slightly disappointing, given that ascription theory is not intended as a psychologically realistic description of how psychological states are attributed but as a metaphysical account of what their possession consists in (even if Mölder holds that not just the states, but also the subjects possessing them and the ascibers attributing them, are merely pleonastic entities).

Connectedly, it might be asked how much or little evidence is required—from the four types of source Mölder reckons with—in order to warrant an ascription. Presumably, if the subject belongs to the circumscribed range (a normal adult human, say), very little evidence may warrant ascription. Mölder has more to say about coherence between pieces of evidence than about volume, and the interesting question, perhaps, is whether even in the limiting case of just seeing a person standing in the sun—but about whom one knows nothing else—is enough for ascribing to her the belief that the sun is shining (this is not zero evidence, of course, but draws on environmental stimuli).[6] After all, the less the amount of evidence is, the easier will it tend to be to achieve coherence among it. But perhaps there is a limit where ascriptions based on very meagre evidence—and thus bordering

[6] Mölder has a sophisticated technical discussion of coherence, drawing on Thagard (2000) and applying his algorithm to the problem of approximating maximal coherence for asciption statements (165–170). A problem with Thagard's approach is that it finesses the philosophically important questions of what makes for the weights of contraints it takes as input (cf. Glymour 1992).

on charity—cannot satisfy the condition of not requiring revision and thus reaching canonical status.

Curiously, despite Mölder's labeling ascription theory a version of interpretivism, interpretation in any more hermeneutical sense seems to play a much lesser role for him than in Davidson's and Dennett's theories. Mölder's account allows a very broad spectrum of sources or evidence for ascriptions. This is perhaps connected with his interesting rejection of one of the central ideas in Davidson's and Dennett's work: that we are required to ascribe minimal rationality to the subject we're attributing intentional attitudes to. Their rationality requirement, simply put, is that the subject has largely coherent attitudes, acts in ways rationalisable given those attitudes, and holds beliefs largely true by the interpreter's lights (Davidson) or has wishes and beliefs congruent with her biological endowment, niche, and needs (Dennett).

Mölder writes: "I am inclined to think that the quarrel over rationality is a red herring in discussion about interpretation." (119). And his theory does not require even coherence among the attributee's beliefs, for we may see how evidence from the subject's behaviour and environment warrants ascription, for example in the light of evidence from her idiosynchratic history, of inconsistent beliefs. At the same time, Mölder accepts the holism of the mental, and notes that "mental states are ascribable in holistic bundles" (123). He also notes that such holistic ascription is tied up with the need to make beliefs and other attitudes "intelligible" (123). Holism is also invoked as a reason to allow other ascribable mental states as a source of evidence (161). Presumably the ascribability of "other" mental states is constrained in similar ways, though, so there is a question whether it is not misleading to represent this as a source of evidence on a par with the three other ones. And in both Dennett and, especially, Davidson, the rationality assumptions made on the subject's part are tied to the fact that mental holism is chiefly content holism: inconsistencies among ascribed beliefs undermine not only the project of making sense of the interpretee, but of making sense of individual beliefs (desires, etc) as well. Here, Mölder's approach is very different. As we saw, he takes the meaning of "X believes that p" to be already settled whenever an attribution is (or could be made), for any substitution-instance. This is fine, but the question where the need for coherence on the subject's part becomes pressing, is how to make sense of—make intelligible—a case such as Dennett's (1987, 14):

> A biologist friend of mine was once called on the telephone by a man in a bar who wanted him to settle a bet. The man asked: "Are rabbits birds?" "No" said the biologist. "Damn!" said the man as he hung up. Now could he *really* have believed that rabbits were birds? Could anyone really and truly be attributed that belief? Perhaps, but it would take a bit of a story to bring us to accept it.

The story needed seems to be one precisely granting a measure of coherence on the interpretee's beliefs. Short of that, it is hard to see how to warrant ascription. Similar considerations motivate Davidson's insistence on coherence. I am not sure why Mölder appears less impressed by such considerations, but I suspect that this may be because of meaning-theoretical views not made explicit in the book. However, even if there is no question over the meaning of the statements used to ascribe beliefs, the problem of how incoherent chunks of beliefs—as opposed to particular false beliefs—might be ascribable does not go away, and I would have welcomed a discussion of why Mölder does not find Dennett's and Davidson's arguments against this possibility compelling.

This issue is connected with another one on which Mölder says very little: of what to make of evidence consisting of specifically verbal behaviour. In Davidson, of course, this issue has centre stage, and Mölder differs from Davidson not just in not aiming to provide a theory of meaning, but in accepting already interpreted speech of a subject into the evidence base for ascriptions of attitudes. But as the Dennett example makes vivid, issues of interpretation often enter even if such evidence is granted, and they bear on what an ascriber is to make of the evidence.

Moreover, there is the rationality requirement on relations between beliefs, desires, and actions. Mölder specifically insists on a distinction between intelligibility and rationality, where the latter "cannot be a constitutive requirement for interpretation" (120). As illustration, he uses an example in (Levin 1988): a man wishes to clean his carpet and belives that he can do so by using his vacuum cleaner. But instead of using it, he washes his car. This is intelligible, Levin and Mölder say, if the man regularly is reminded of car washing when thinking about vacuum cleaners, and even generally behaves according to the regularity "If [the subject] wants Q and believes that P brings about Q, and P reminds [the subject] of R, then [the subject] will do R" (Levin 1988, 212, quoted in Mölder, 120). And so it is. But in order to know that the man behaves in accordance with this psychological regularity, can an interpreter proceed without rationality assumptions connecting his actions in general with his beliefs and desires in a less idiosynchratic way? This seems dubious.

A final note on Dennett. Mölder makes a valiant attempt to package Dennett's views in his own preferred metaphysics (115–116). But in discussing Dennett, Mölder often appears to equate patterns, as Dennett uses the notion, with mental states (with content), and to take Dennett's theory as one where "intentional patterns" are ascribed. Although Dennett is elusive on the subject of patterns, I think that it is fairly clear that he takes the "real patterns" (Dennett 1991), whose reality grant a measure of realism to his

view, as not themselves the attitudes whose ascription they motivate. Nor do I think that McLaughlin and O'Leary-Hawthorne's 1994 reading of Dennett, on which attitude talk concerns internal dispositions to behaviour, is correct (Mölder suggests that Dennett 1987 supports their reading, but I beg to differ). Dennett's talk of patterns goes from the clear cases of compression theory, via the still fairly clear cases of Conway's Game of Life, to the case of patterns grounding (sort of) the intentional stance; when it reaches the latter, it is arguably just a suggestive but elusive metaphor. But in order to make sense of Dennett's rationales for the intentional stance, I believe that patterns here must be simply patterns in behaviour. These do not supervene on internal dispositions, however, for they depend on how the world surrounding the subject is. So they are not themselves "intentional"; and in fact, in the quotations Mölder provides in connection with his discussions of "intentional patterns" (e.g. on 116), Dennett actually does not speak of *intentional* patterns, but simply of (various kinds of) patterns.

In sum, although I am not convinced that ascription theory is the best version of intepretivism available—or even that it *is* a form of interpretivism —Mölder's book is enormously rich in ideas and detailed discussion, and it is the most penetrating study of the general approach I have seen. I do remain suspicious of the streak of metaphysical puritanism enveloping Mölder's ascriptionism, and I am unsure—despite Mölder's repeated insistence to the contrary—how essential his Schifferian metaphysics is to his theory.

Bibliography

Braddon-Mitchell, D. and Jackson, F. (1996). *Philosophy of Mind and Cognition*, Blackwell, Oxford.

Dennett, D. (1987). *The Intentional Stance*, MIT Press, Cambridge, MA.

Dennett, D. (1991). Real patterns, *Journal of Philosophy* **87**: 27–51.

Dennett, D. (1994). Get real, *Philosophical Topics* **22**: 505–568.

Ellis, B. (2001). *Scientific Essentialism*, Cambridge University Press, Cambridge.

Glymour, C. (1992). Invasion of the mind snatchers, *in* R. Giere (ed.), *Cognitive Models of Science*, Vol. XV of *Minnesota Studies in the Philosophy of Science*, University of Minnesota Press, pp. 465–471.

Levin, J. (1988). Must reasons be rational?, *Philosophy of Science* **55**: 199–217.

Lewis, D. (1972). Psychophysical and theoretical identifications, *Australasian Journal of Philosophy* **50**: 249–258.

Lewis, D. (1994). Reduction of mind, *A Companion to the Philosophy of Mind*, Oxford University Press.

McLaughlin, B. and O'Leary-Hawthorne, J. (1994). Dennett's logical behaviorism, *Philosophical Topics* **22**: 189–258.

Mölder, B. (2010). *Mind Ascribed: An Elaboration and Defence of Interpretivism*, John Benjamins, Amsterdam.

Putnam, H. (1975). The meaning of "meaning", *in* K. Gunderson (ed.), *Language, Mind, and Knowledge*, Vol. VII of *Minnesota Studies in the Philosophy of Science*, University of Minnesota Press, pp. 131–193.

Schiffer, S. (2003). *The Things We Mean*, Clarendon Press, Oxford.

Thagard, P. (2000). *Coherence in Thought and Action*, MIT Press, Cambridge, MA.