Interpretivism and "Canonical" Ascriptions

Henry Jackman

Department of Philosophy, York University

The following paper investigates the crucial notion of a "canonical ascription statement" in Bruno Mölder's *Mind Ascribed*, and argues that the reasons given for preferring the book's approach of canonicality to a more common understanding of canonicality in terms of the ascriptions we would "ideally" make are not only unpersuasive, but also leave the interpretivist position more open to skeptical worries than it should be. The paper further argues that the resources for a more compelling justification of Mölder's conception of canonicality are already in Mölder's book itself.

Keywords: meaning, truth, interpretation

Bruno Mölder's *Mind Ascribed* (2010)[1] is the most thorough and impressive defense of interpretivism that I have seen. Indeed, its so much so that I find very little to disagree with. Still, while I am sympathetic with both the larger project, and with Mölder's account of the crucial notion of a "canonical ascription statement," I think that Mölder does himself a disservice with the reasons he gives for preferring his account of canonicality to a more common alternative. Not only are his reasons unpersuasive, but they also make the interpretivist position seem more open to skeptical worries than it is. Further, the resources for a more compelling justification are, as I will argue, already in Mölder's book itself.

Mölder presents the following account of what it is to be in particular mental state (where "'X' stands for a subject of mentalistic ascription, 'f'-ing specifies a mental state and 'p' gives its content"):

> *Possession condition* for X f's that p: f's that p is canonically ascribable to X" (171)

A lot of work is done here with the term 'canonically,' since if the possession condition for X f's that p, were merely that 'f's that p is ascribable to X,'

the account would be too forgiving. It would leave no gap between actually having a mental state and merely *seeming* to an ascriber to have one, and without such a gap, ascriptions that we intuitively take to be mistaken would instead have to be treated as constituting attitudes of the purported believer. Indeed, 'misascription' would, on such a view, turn out to be a contradiction in terms. As Mölder puts it:

> If we explain the possession of mental states in terms of their ascribability, we need a criterion for the kind of ascription that the interpretivist account can utilize for its purposes. The reason for this is that not every particular ascription is up to the required level. Often we make ascriptions on the basis of insufficient data, or we do not pay attention to the different aspects of the situation. Thus, if the theory were to claim that X has all those mental states that one could ascribe to X, it would hardly be believable. What we need is to fix a standard for ascriptions that can play the constitutive role that is designated for them in the ascription theory. I shall call the ascription that conforms to this standard a *canonical* ascription. Using this term, we can say that one has a mental state provided that it is canonically ascribable to one. This notion can also be used to draw the distinction between the correct and the incorrect ascription. Hence the distinction is an interpretivist version of the is/seems difference. (170)

Mölder explicates this crucial notion of canonical ascription as follows:

> *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. (175)

'The ascription statement *As*' is meant to stand here for statement like 'X *f*'s that p' (175.) and, crucially, 'ascription sources $c_1,\ldots, c_4$' stand for (1) "The ... *other mental states with contents that are ascribable*," (2) "a subject's *behaviour*," (3) "a subject's *environment*, or more exactly, the environmental *stimuli* that have an effect on the subject," and (4) "a subject's *personal background*, which involves facts about the subject's personal history and dispositions, their language community and idiosyncratic language understanding" (161-162).

Mölder's decision to explain canonical ascriptions this way raises the question of why he does not cash out the seems/is in a more familiar fashion. In particular, it raises the question of why the Interpretivist should not explain the non-canonical/canonical difference in terms of the ascriptions that we actually make and the ascriptions that we would *ideally* make if our ascriptions were *maximally* coherent with complete knowledge of all of the ascription sources. Mölder, however, explicitly rejects this "ideal" route, and his reasons for doing so are worth quoting at some length:

perhaps we can specify the canonicality condition through the notion of an ideal ascriber. We might say that one has such mental states that an ideal ascriber would ascribe to one. The ideal ascriber is omniscient with respect to the sources and can assess the coherence of the information better than anyone else. However, this account has a serious deficiency. Namely, it makes all of our ordinary ascriptions noncanonical, for ordinary folk cannot instantiate the ideal ascriber. It follows that we are always mistaken about others' mental states, and, what is worse, we are also mistaken when we engage in self-ascription. No one can be the ideal ascriber, even in respect of himself, for no one has privileged access to the ascription sources. Note that the possession condition is outlined in terms of what is ascribable rather than in terms of what is in fact ascribed. If we specify the canonicality condition through the ideal ascriber, the consequence from these two conditions is that in each case of actual ascription, there is some mental profile ascribable by the ideal ascriber which may differ from any profile that could actually have been ascribed.

Hence it is better not to fix the canonicality condition to the notion of an ideal ascriber. Since we routinely ascribe mental phenomena in everyday life, it would be astounding if the canonical ascription were in principle inaccessible to us. (172)

Mölder seems to make three crucial assumptions in the passage above:

1. "[O]rdinary folk cannot instantiate the ideal ascriber."

2. If we cannot instantiate the ideal ascriber, then canonical ascriptions are "in principle inaccessible to us."

3. If canonical ascriptions are inaccessible to us, then "we are always mistaken about others' mental states."

While I agree with the first of these assumptions, I would like to call in to question the second and third and thus the main reasons Mölder gives for rejecting the idealized notion of canonicality.

After all, if canonicality were spelt out in terms of ideal conditions, it would look something like the following:

As is a *canonical ascription statement* $=_{df}$ *As* is a statement that is maximally coherent with complete knowledge of the ascription sources $c_1, \ldots, c_4$.

On such an understanding of canonicality, it would certainly be the case that we would not be canonical *ascribers*, but it does not seem to follow from this that there would be no sense in which we could make canonical *ascriptions*. An ascription, *As*, that we made could be deemed canonical if it was also

an ascription that a canonical ascriber would make, even if the grounds we had for making it ourselves are not themselves canonical. It would seem to follow then that, contrary to assumption 2 above, canonical ascriptions are, in fact, accessible to us.

Now one might argue that while the canonical ascriptions are accessible to us in the sense that we can make an ascription, *As*, that also happens to be a canonical one, they are not accessible to us in the stronger sense that we can be the ones to make the ascriptions canonically. But why should that matter? Insisting on such a stronger notion of accessibility would salvage assumption 2, but only at the cost of undermining assumption 3. It is only the weaker notion of accessibility that is needed for our ascriptions to be *true*, and so it is only this weaker notion of accessibility that is required to put to rest the worry, expressed in the third assumption, that if canonical ascriptions are inaccessible to us, then our ascriptions will be "always mistaken." We make a correct/true ascription if it corresponds to a canonical one, even if we cannot ascribe it canonically ourselves.

This would leave us with a position on ascription statements very close to a traditionally popular reading of Peirce's famous claim that "the opinion which is fated to be ultimately agreed to by all who investigate, is what we mean by the truth" Peirce (1992b). On this popular reading, a statement is true if it would be believed at an idealized "end of inquiry" in which we had taken our investigation into every question to its ideal limit under epistemically ideal conditions. It is not a consequence of such a view that all of our beliefs must be false just because they are held under less than ideal conditions. We are not, and probably never will, reach this "end of inquiry" but our current beliefs can still count as true if they correspond to beliefs that would be held at that ideal end point.

The idea that a true ascription need only correspond to the ascription that would be canonically ascribed, rather than be canonically ascribed itself, is a natural one, and seems to be endorsed by Mölder himself when he admits that many, if not most, of our ascriptions do not meet even his own "relaxed" (173) conditions of canonicality.

> A consequence of this account is that it can happen that we do not recognise the canonical ascription when it is actually made. The reason for this is that although we can in principle instantiate canonical ascribers, we need not always be sufficiently well-placed in respect of the sources and the coherence condition. We do not know whether we have taken all relevant information into account. But this characterises our epistemic condition in general and it is not a deficiency in the present account, which does not purport to compile a practical vade-mecum for everyday ascribers. Instead, it attempts to provide a metaphysical condition for the possession of mental states. (174-175)

> The theory is given a weaker formulation in terms of ascribability, compared to the stronger but implausible formulation in terms of actual ascription. For, surely, the examples of actual ascription in everyday life that satisfy the canonicality requirements are not that numerous. (177)

Of course, a possibly relevant difference here is that on Mölder's account our ascriptions are *typically* made in a non-canonical fashion, while on the ideal understanding of canonicality, they *never* are made in a canonical fashion, so for Mölder, *As* will always be, in principle, ascribable *by us*.[2] Perhaps it is this in principle accessibility that Mölder takes to be essential, and in the following he stresses the importance of the fact that, on the relaxed conception of canonicality, we can *actually* ascribe canonically.

> In saying that the canonical ascription must be such that it can be actual, I mean that it should rely on the sort of information and involve the sort of computational processes that are in principle accessible to ordinary folk. We *can* be canonical ascribers according to this standard of canonicality. This implies that canonical ascriptions are to be found in a world like ours, and the threat from omniscient ascribers is thereby blocked. (174)

But what exactly is the "threat" that the omniscient ascriber would really present if canonicality were understood in ideal terms?

At times it seems like the underlying worry is epistemic. On both the relaxed and ideal accounts an ascription is true if it corresponds to a canonical one, but on the more relaxed understanding of canonicality, it might seem that we could, in principle, *tell* whether an ascription was canonical, while on the ideal understanding, one could never *definitively* put to rest the worry that a particular ascription failed to be canonical, and this leads Mölder to worry that, "If we reach the canonical method at the end of inquiry, then we are faced with the consequence that we presently have no idea what mental states we now "really" have." (173)

However, if the worry is that we can never know whether a *particular* ascription of ours is not mistaken, it is not clear that even the relaxed version canonicality is free from trouble. Relaxed canonicality may be achievable, but one can never tell with certainty that one has achieved it (since the epistemic possibility that one's ascriptions may need to be revised can never be

---

[2] Or, possibly, even if they *never* were on Mölder's understanding of canonicality, such ascriptions *could*, in principle, be made by us, while on the ideal understanding they *could not*, in principle, be actually made by us. As he puts it, "It is true that the act of ascription need not be actual, but all that matters is that canonical ascribers can be actual, that is, ordinary interpreters can instantiate the (pleonastic) property of being a canonical ascriber" (178).

definitively ruled out). That an account of ascriptions should entail a certain amount of fallibilism about the ascriptions we make is not a bad thing, and perhaps the worry is that the ideal account of canonicality pushes one from fallibilism to skepticism, and that the ideal conditions may be so far from what we actually do that there is a real worry that *most*, or even *none*, of our ascriptions will correspond to the canonical ones. If the idealized notion of canonicality entailed skepticism rather than just fallibilism, then that would be reason to reject that account of belief.

Such worries would be similar to those of someone who worried that Peircean accounts of truth, particularly those that tied truth to what was believed at some sort of idealized "end of inquiry," left one open to skepticism, since it seems possible that none of our beliefs would still be endorsed by any mythical community that reached such an ideal stage. Peirce (1992a,b) himself had little time for such doubts, and there are reasons for thinking that they are especially misplaced when the topic is our ascriptions of mental states.[3]

Davidson himself gives such reasons in his discussion of the "omniscient interpreter" in his "A Coherence Theory of Truth and Knowledge" (1986) where he argues that any interpreter, including an omniscient one, would have to treat most of our beliefs, including our beliefs about what others believe, as true. Of course, Mölder does not endorse every aspect of the Davidsonian theory, in particular, its crucial appeal to the Principle of Charity,[4] but in the absence of any concrete reason why we should expect such a divergence from the ideal interpreter, it seems doubtful that we should need anything like an *a priori* guarantee that we will match such ascriptions. Further, if an omniscient interpreter really did come up with radically different interpretations than we did, its hard to see why that would not give us reason to revise our own, which would undermine their canonicality in even the relaxed sense. If an ideal interpreter had an interpretation different than our own, presumably that difference could be explained either in terms of

[3] Some (say, Churchland 1981) might argue that all of our ascriptions could turn out to be false because an ideal ascriber would use a completely different taxonomy of mental states, but such an ascriber would be relying on evidence beyond that available in c_1-c_4, and in any case, Mölder argues persuasively that such a theory would be working with "successor concepts of folk notions" (132) and would not belie the truth of the claims made with the folk notions themselves.

[4] See, for instance, "Perhaps the most crucial difference between Davidson's approach and the ascription theory is that the ascription theory is not presented as a theory of meaning. In Davidson's approach, both a subject's beliefs and the meanings of his utterances are fixed in the process of (radical) interpretation. It is assumed that the subject's beliefs conform to the principle of charity and this allows for the assignment of interpretive truth conditions to the subject's sentences. For several reasons, I do not think that this can provide a complete account of meaning." (124).

their being aware of some fact that we were unaware of, or in terms of their being able to tell a more consistent story about those facts than we were, and either of those scenarios would give us reason to revise our position.

It seems, then, that the reasons Mölder presents for rejecting the ideal understanding of canonicality are not necessarily compelling. If there could be an ideal interpreter, it is hard to see why an interpretivist should not understand our mental states in terms of what they would ascribe to us.

That said, I think that something like Mölder's notion of canonicality is preferable to the idealized notion, but this is not because the idealized notion raises too many worries that there may be "some mental profile ascribable by the ideal ascriber which may differ from any profile that could actually have been ascribed" (172), but rather because of doubts about whether there really could be a determinate ideal interpreter.

To see why, consider again the parallels with the reception of the Peircean account of truth. As mentioned above, an initially popular conception of Peirce's account tied truth to what we would believe at the "end of inquiry," where the end of inquiry was understood as the ideal epistemic state where one's community had taken its inquiry into every question, under ideal epistemic conditions, as far as it could go. The fact that we have not actually reached such a stage was not a problem for such an account, since our beliefs can count as true if they correspond to beliefs that *would be* at such a stage. However, explaining what these "ideal epistemic conditions" and "end of inquiry" were presented serious difficulties for the theory, especially if these later two notions were supposed to be explained without any prior reliance on the concept of truth. Indeed, many critics came to suspect that we could make no coherent sense of those notions.[5]

Such problems have led some of Peirce's defenders[6] to argue that the Peircean account need not posit any sort of "end of inquiry" where all of our questions are settled, but rather only needs to state that a belief is true if it would continue to be held through successive improvements of our epistemic position. Whether there could be a single stage where *all* of our questions were settled, or any single set of epistemic conditions that are best described as "ideal," is irrelevant to whether our beliefs can be true on such an account. Such a position hopes to capture what was appealing about the Peircean view (such as connecting our notion of truth to that of rationality and inquiry, and ruling out the possibility that even our best possible

---

[5] At least partially due to the more "formal" problems that seemed to arise when truth was defined in terms of an epistemic state that was meant to incorporate the justification for every truth at once (see, for instance, the discussion of the relation of Peircean truth to Gödel's theorem in Johnston 1993).

[6] See, for instance, (Misak 2013).

inquiry into a question could be mistaken) without needing to make sense of the idea of a global "end of inquiry." Contemporary Peircean's are thus inclined to replace:

> P is a true $=_{df}$ P is a statement that we would endorse when we reached the ideal "end of inquiry."

With

> P is a true $=_{df}$ P is a statement that would continue to be endorsed no matter how much further we were to investigate and debate.[7]

Rather than appealing to anything substantial like "ideal epistemic conditions" such an account of truth explains truth with the comparatively modest base of a belief's surviving successive improvements to our epistemic position. It thus does not have the explanatory burden of explaining what the ideal epistemic conditions would be, and does not face any problems if the phrase 'ideal epistemic conditions' failed to have a (unique) referent.

I think that Mölder's account of canonicality has similar advantages over its more idealized rival.[8] The problem with the idealized conception of canonicality is that it presupposes that we have a clear sense of what it would be for us to have "maximally coherent beliefs." However, as Mölder discusses in the book's fifth chapter, there are good reasons[9] to think that rationality

---

[7] See, for instance, (Misak 2013, 37). The Peircean notion of truth thus turns out to be something very much like Crispin Wright's notion of "superassertability" (Wright 1992, 48).

[8] Indeed it was not entirely clear to me why one could not even replace Mölder's

> *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. (175)

With the simpler:

> *As* is a *canonical ascription statement* $=_{df}$ if *As* were in fact ascribed, it would not require revision.

Of course, the simpler version seems to be little more than the application of an epistemic conception of truth to the ascription of mental states, and would seem compatible with views about the ontology of belief that tied them more closely to brain states, etc. Part of what makes that account an interpretivist one is precisely that the relevant information is limited to "the ascription sources $c_1,\ldots, c_4$," and not, say, the readings of an FMRI machine. Still, that sort of restriction on what type of evidence would be relevant for revision need not be built explicitly into the definition of canonical ascriptions itself, though perhaps without it, one could no longer say that "the canonical ascription supplies the condition that fixes the obtaining of such facts" (171).

[9] Most famously associated with (Davidson 1980), but see also (McDowell 1985), and (Child 1994).

is "uncodifiable" in a way that prevents there from being a determinate answer to the question of which set of ascriptions best fit the ascription sources and thus be in a position to claim to be "maximally" coherent with them. If rationality is, "uncodifiable" in this way, then there may be no determinate answer to the question of what set of ascriptions would be maximally coherent with the ascription sources, and thus no determinate answer to the question of what mental states an ideal ascriber would ascribe to us. If rationality is uncodifiable, the ideal conception leaves facts about our mental life indeterminate in the way that Mölder's relaxed conception does not. As Mölder put it, "the [relaxed] notion of canonical ascription helps us to get rid of the indeterminacy at the level of possession of mental states. When it is allowed that all variant canonically ascribable mental states are possessed, then the variation in interpretations does not entail the indeterminacy of the mental." (182)

Indeed, this might be a way to make sense of the worry that that, on the idealized conception of canonicality, it could very well turn out that none of our actual ascriptions are true. Mölder initially presents this worry as if it were about the canonical ascriber making a *different* set of ascriptions than we would ("there is some mental profile ascribable by the ideal ascriber which may differ from any profile that could actually have been ascribed"), which is why we might be "always mistaken about other's mental states." On this line of thought, by contrast, the worry is that, on the ideal conception of canonicality, *no* ascriptions are canonical, since there is no set of ascriptions that are maximally coherent. If no mental state is canonically ascribable, then the possession conditions for being in a particular mental state are never satisfied, and thus any attribution of mental states to others will be mistaken.[10]

Of course, arguments about the uncodifiability of rationality are always controversial, but on the approach suggested here, if there does turn out to be a set of ascriptions that is "maximally coherent" with the ascription sources, then one should revise one's ascriptions to match it, leaving the "ideal" and "relaxed" conceptions of canonicality functionally identical. The ideal conception may thus give the same results as the relaxed conception if rationality turns out to be codifiable, while it commits us to having no mental states at all if rationality is uncodifiable. Compared to the relaxed conception, the ideal conception of canonicality has very little benefits while running the risk of encouraging serious costs. The relaxed conception thus seems prefer-

---

[10] In much the same way, one could argue that the "traditional" reading of the Peircean account of truth could entails that no belief is true, since if the "end of inquiry" fails to pick out even a *possible* state, there would be no beliefs that would be held at the end of inquiry.

able to the ideal conception even if it is merely an open question whether or not rationality can be codified.

I must confess that I worry that I may be missing something here, since I think that this alternate defense of relaxed canonicality is at least implicit in Mölder's book already, and I will look forward to hearing from Mölder just why anything more should be needed to rule out the ideal conception of canonicality.

Bibliography

Child, W. (1994). *Causality, Interpretation and the Mind*, Oxford University Press, Oxford.

Churchland, P. (1981). Eliminative materialism and the propositional attitudes, *The Journal of Philosophy* **78**: 67–90.

Davidson, D. (1980). Mental events, *in* D. Davidson (ed.), *Essays on Actions and Events*, Oxford University Press, pp. 207–225.

Davidson, D. (1986). A coherence theory of truth and knowledge, *in* E. Lepore (ed.), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Blackwell.

Johnston, M. (1993). Objectivity refigured: Pragmatism without verification, *in* J. Haldane and C. Wright (eds), *Reality, Representation, and Projection*, Oxford University Press, New York, pp. 85–130.

McDowell, J. (1985). Functionalism and anomalous monism, *in* E. Lepore and B. McLaughlin (eds), *Actions and Events*, Blackwell, Oxford.

Misak, C. (2013). *The American Pragmatists*, Oxford University Press, New York.

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

Peirce, C. S. (1992a). The fixation of belief, *in* C. S. Peirce (ed.), *The Essential Peirce*, Vol. 1, Indiana University Press, Indianapolis.

Peirce, C. S. (1992b). How to make our ideas clear, *in* C. S. Peirce (ed.), *The Essential Peirce*, Vol. 1, Indiana University Press, Indianapolis.

Wright, C. (1992). *Truth and Objectivity*, Harvard University Press, Cambridge, MA.