

Some More Thoughts About Thought and Talk:

Davidson and Fellows on Animal Belief

1. One tenet of Donald Davidson's account of radical interpretation is the interdependence of utterance interpretation and belief ascription. Assigning meanings to a speaker's utterances requires interpreters to make concomitant determinations of their beliefs. But at the same time, one cannot fix the contents of a subject's beliefs without identifying the meanings of their utterances. So belief specification and meaning determination go hand-in-hand. Now if Davidson is right, and one cannot attain a purchase on a subject's beliefs without also interpreting their utterances, then what about our nearly irresistible urge to attribute beliefs to creatures that lack linguistic competence? If belief ascription requires one to interpret a subject's utterances, then it would seem that no matter how useful or fruitful it was to treat them *as if* they had beliefs, the attribution of beliefs to non-linguistic animals is, strictly speaking, unwarranted.

Davidson defends this unfriendly attitude towards animal belief in a couple of short articles.¹ However, the arguments to be found there are, by nearly all accounts, overly compressed and less than fully convincing.² In a recent issue of *Philosophy*, Roger Fellows attempts to defend Davidson's position by providing these arguments with some much-needed unpacking.³ In this paper, I'll briefly explain why I fail to find Fellows' refinements of Davidson's line of reasoning wholly persuasive.

2. In a preliminary remark, Fellows reminds us that while his view is unfriendly to the idea that non-linguistic animals could have beliefs, that needn't make it unfriendly to animals themselves.

While they might lack the discursive or deliberative capacity to qualify as moral *agents*, that doesn't rule out their status as moral *patients*. In this respect, they are like human children-appropriate objects of moral concern but not condemnation. And so his position doesn't necessarily license us to behave beastly towards the beasts. More significantly, Fellows doesn't deny that animals could have pre-linguistic intentional states or "quasi-beliefs." So what prevents such states from counting as bona-fide beliefs? Fellows' answer rests upon an appeal to what he dubs 'Principle P': 'A creature can have the belief that P ascribed to it, only if it possesses those concepts mastery of which is required in order to have the belief that P.' (p. 591) So presented, this principle is vacuous; *of course* believers must have whatever conceptual mastery is required to have beliefs. Nevertheless, Fellows' intent is clear. The capacity for belief requires at least some conceptual mastery. Perhaps it would be better to say that to have the belief that P one must have mastery over the concepts required to *express* that belief. While we might find it useful to attribute to non-linguistic creatures intentional states the contents of which we employ our own concepts to pick out, it isn't as if they need to be in possession of those very same concepts, or indeed any concepts at all. On this view, beliefs must be conceptually articulated. At a minimum, this means that beliefs have to be part of a sufficiently rich network of inferentially connected beliefs, through which (and only which!) individual beliefs acquire a determinate content. That would explain our reluctance to credit dogs with beliefs that essentially involve the concept of BONE, for such beliefs implicate a host of other, inferentially related beliefs that we find it hard to imagine a dog having. So to credit dogs with beliefs that are (in a de dicto sense) about bones or about the president threatens to inflate their true cognitive capacities, for it is unclear that creatures like dogs are able to assemble inferential networks that are sufficiently rich to sustain beliefs about bones or presidents.

Clearly, one response available to the friend of animal belief is to reject the idea behind Principle P, namely that beliefs require conceptual mastery. After all, the sin of deflating a creature's cognitive capacities ought to be every bit as worrisome as that of inflating them. And plenty of folk are prepared to countenance doxastic-like intentional states with a "mind-to-world" direction of fit that aren't conceptually articulated. To deny that such states qualify as genuine beliefs begins to look like a terminological quibble. Indeed, principle P sits uneasily with Fellows' further preliminary remark that belief is a basic propositional attitude which "lies at the centre of the cognitive domain." (p. 589) For the more conditions one piles into the notion of belief (such as conceptual articulation), the less plausible it is to think of it as a ground-level intentional state, one that must be available to a creature for it to have any intentional states whatsoever.

3. Still, the divide between concept-mongering creatures and those capable only of primitive, non-conceptual intentional states is one worth marking, whether or not one chooses to draw it in terms of the ability to have propositional attitudes like belief. It is this gulf that Davidson (and Fellows) doesn't think can be crossed without the aid of language. So let's turn now to the case Davidson builds against animal belief, as well as Fellows' elaboration of that case.

As Fellows observes, Davidson's argument breaks down into two separate subconclusions.

First subconclusion: One cannot be a believer without having the concept of belief.

Second subconclusion: One cannot have the concept of belief without the ability to interpret the utterances of others.⁴

Together these two clearly entail that belief requires linguistic ability. Fellows subscribes to the same overall strategy, endorsing both of these subconclusions, but providing somewhat distinct arguments for them. I'll address each of the subconclusions in turn.

The first subconclusion is in my opinion the most controversial. It asserts that to have any beliefs at all, one must not only possess the concepts contained in that belief, one must also possess the concept of belief itself. The concept of belief, then, is a ground-level concept, in the sense that it must be grasped, if any concepts are to be grasped at all. Now this privileged status certainly seems implausible, especially if it is taken to mean that a creature must have a sophisticated theory of mind for it to have any beliefs at all. So what does Davidson have to say on its behalf? Here Davidson claims that the ability to have beliefs demands that a creature be able to recognize when circumstances are not as the creature has taken them to be. Believers must be capable of what he terms "surprise."⁵ But, he continues, this capacity for surprise in turn demands an awareness of the possibility that one might be mistaken, and so logically requires creatures to have the capacity for second-order thoughts or beliefs about their beliefs. In particular, a believing creature must be capable of having the belief that one of its own beliefs fails to be true, or that one of its own beliefs fails to be as it should, and so ought to be revised. On a reasonable reading of Principle P, then, such second-order beliefs require one to have some grasp of the concept of belief. And so the first subconclusion is secured: belief requires the concept of belief.

The first step in this argument seems reasonable enough. The requirement that a creature be capable of surprise amounts to the requirement that believers must somehow acknowledge that their beliefs are governed by the norm of truth. The capacity for surprise implies an appreciation that the way one takes the world to be need not line up with the way it actually is; surprise requires a sensitivity to the fact that one's beliefs can be mistaken. To be a genuine believer, then, one must be sensitive to the demand that one's beliefs are beholden to the way things are, and so subject to revision. Since they fail to rationally regulate their behavior around any such norm, we are thus inclined to withhold attributions of belief to instinct-bound or inflexibly tropistic creatures. They simply fail to "get the point" of having beliefs at all. However, the trouble with Davidson's argument lies in the step from this capacity for surprise to the requirement that one be capable of having beliefs that are about beliefs. Here Davidson overlooks the possibility that the contrast between truth and error might be grasped implicitly in a creature's self-corrective behavior. As Peter Smith has pointed out, far from obviously requiring believers to have any conscious understanding that one's beliefs might be false, an appropriate awareness of the possibility of error might simply require one to be "sensitive to the demands of truth in the sense that he holds his beliefs open to revision in the light of experience."⁶ Believers then would not have to make any use of the concepts of belief or truth in their own thoughts. That type of understanding would emerge only after one consciously reflects upon practices of self-correction that are already in place.

Davidson's argument thus turns on the idea that belief revision, as such, must be reflective activity. The revision of one's beliefs must be governed by the thought that one's beliefs ought to be changed. But this surely renders the process of belief revision much too

intellectual. We would do better to regard belief revision, not as activity that itself needs to be guided by thoughts, but rather as activity through which we can begin to discern thoughts in the first place.

Fellows offers a somewhat different defense of the first subconclusion, one that doesn't focus upon a creature's capacity for surprise. Instead, he argues that for a creature to possess the conceptual capacity demanded by Principle P for it to be ascribed any beliefs at all, it must be capable of counterfactual reasoning. In order for a creature to master a certain concept, not only must it know what falls into that concept's actual extension, it must also know what falls in its possible extension. Otherwise, we wouldn't be able to distinguish the creature's deploying that concept, as opposed to one that is extensionally equivalent to it. As Fellows puts it:

So if C [a creature] is to have the belief that there is a bone in front of it, it must possess the concept BONE. If C possesses this concept, it knows not only the actual extension of BONE, but also which possible objects fall within its extension or are excluded from it. If C failed to be able to distinguish the actual from the possible extension of the concept BONE, it could not have beliefs about bones. (p. 598)

But, Fellows continues, in order to have beliefs about possible objects, a creature must be capable of counterfactual thought. And such counterfactual reasoning requires a creature to make judgements about what would be the case, if certain conditions it doesn't believe to be the case, actually turned out to be true. A reasonable interpretation of Principle P demands that for a creature to have such complicated thoughts, it must then have mastery of the concepts of BELIEF, as well as those of TRUTH and CONSISTENCY. And so once again, the first subconclusion is secured.

Does this defense of the first subconclusion fare any better than Davidson's? I think it falls prey to much the same kind of complaint as before: that Fellows over-intellectualizes concept-possession much like Davidson has an overly reflective account of belief revision. Fellows begins his argument with an importantly true observation. Counterfactual considerations are indeed important for the proper attribution of beliefs and concept mastery. In order to license our attribution of a certain concept to a creature in favor of one that is extensionally equivalent to that concept, we need to know what, given certain contrary-to-fact circumstances, a creature would believe. That is, we need to justify statements of the form:

1. Were certain circumstances to be the case, a creature C would believe such 'n such.

But Fellows slides from the justification of statements of this form to the justification of statements such as:

2. C believes that were certain circumstances to be the case, such 'n such.

But this is an evident non-sequitor. It trades on illicitly importing counterfactual conditions into the scope of a creature's beliefs. In order for us to sustain attribution of beliefs involving certain concepts to a creature, all we need to know is what it *would believe* of possible objects. But that doesn't require that the creature itself has to conceive of those objects *as possible*. And so I conclude that Fellows' route through counterfactual reasoning offers little improvement over Davidson's defense of the first subconclusion.

4. The second subconclusion asserts that creatures must have linguistic capacity before they can possess the concept of belief or entertain beliefs about beliefs. While its converse is widely accepted, by no means is this claim self evident. Those who point to instances of apparently deceptive behavior on the part of non-linguistic animals to show that they have second-order

beliefs or rudimentary "theories of mind" don't commit any egregious conceptual errors. To be sure, establishing that non-linguistic creatures are sensitive to the putative states of mind of others is profoundly difficult. In fact, Daniel Dennett (one of Fellows' explicit targets) has argued that one will probably not be able to discern second-order intentionality in the absence of linguistic behavior, for the conditions under which second-order intentionality is most likely to evolve are the same as those in which linguistic behavior is likely to emerge.⁷ But this *contingent*, empirical connection between second-order intentionality and language is certainly not the tight, *conceptual* kind of connection that Davidson is looking for.

Nor is Davidson (or Fellows) likely to embrace the kind of defense Wilfrid Sellars provides for the second subconclusion. In his "myth of Jones", Sellars aims to show how our concepts of inner intentional states like belief could be *modeled upon* antecedently intelligible concepts of overt linguistic utterances.⁸ So the idea of belief is available only to creatures who are already engaged in a practice of making such utterances to one another. However, insofar as it asks us to envision just what Davidson and Fellows want to deny- namely creatures engaged in thoughtful linguistic activity, but who utterly lack the concept of belief - this line would seem to be unavailable to them. It evidently flies in the face of the first subconclusion.

Concerning the second subconclusion, Davidson writes:

We have the idea of belief only from the role of belief in the interpretation of language, for as a private attitude, it is not intelligible except as an adjustment to the public norm provided by language. ('Thought and Talk', p. 170)

This is quick, indeed, and in desperate need of reconstruction. Davidson's reasoning, as best I can make out, goes something like this. In order to grasp the concept of belief one must

understand that beliefs are beholden to the objective world. That is, one must grasp the idea that believers are rationally obliged to revise their beliefs when evidence overwhelmingly suggests that their beliefs are false. As Davidson sometimes puts it, believers must have a grip on the "subjective-objective contrast."⁹ But one can actually witness beliefs coming apart from the way things are (or observe patently false beliefs) only in others, for one's own grasp of the way things objectively are necessarily coincides with one's beliefs about the way things are. So the norm of correctness for beliefs is intelligible only for those who are engaged in the essentially social activity of interpreting the utterances of others. Hence second-order intentionality, or the capacity to have beliefs about beliefs arises only for those capable of language.

However, if this accurately portrays Davidson's reasoning, then it can be challenged at virtually every step. John Bishop, for instance, reminds us that even though the notion of beholdeness to an objective world might figure in any adequate analysis of the concept of belief, it would be unreasonable to require subjects to grasp that analysis before they could have the concept of belief.¹⁰ Otherwise, several humans, including those that seem to converse, would come out lacking the concept of belief. Moreover, it's unclear why the phenomenon Davidson appeals to in support of the first subconclusion - the sense of surprise that drives belief revision - wouldn't be *another* place where subjects encounter the subjective-objective contrast in experience. After all, it's highly implausible that I always need to turn to others in order to grasp that one of my own beliefs is mistaken. So why couldn't the concept of belief also emerge for those creatures who simply try to make sense of their own self-corrective practices, without any regard for others? Finally, even if it were true that the concept of belief emerges only for those

capable of interpreting the behavior of others, it doesn't necessarily follow that they'd have to be interpreting their *utterances*.

Fellows' defense of the second subconclusion is almost as compressed as Davidson's. He begins as Davidson does, with the claim that those with the concept of belief must have the concept of a state which aims at objective truth. This means that to have the concept of belief, one must grasp the distinction between true belief and false belief. "What follows from this," as Fellows sees it, "is that [a creature] C must be able to entertain the thought that its own belief about a particular matter changes while the world remains unchanged." (p. 597) For a creature cannot draw the distinction between true and false beliefs if it cannot distinguish between changes in mind from changes in the world. And this, according to Fellows is precisely what language allows us to achieve.

A language L enables C to distinguish changes in belief from changes in the world. The that clauses in L, which fix the contents of C's beliefs, have (relatively) stable meanings, which enable C to determine the congruence between C's own beliefs and the beliefs of others, on the one hand, and the world itself, on the other. (p. 597)

And later, while restating this last step in a summary:

Language-learning, which is community based, provides for the fixation of belief, in so far as language enables a creature's beliefs to be made manifest and hence held up to scrutiny by itself and by others of its kind. (Ibid)

The central idea, then, is that to have the requisite grip of the subjective-objective contrast, a creature needs to find some way to fix, or to stabilize the contents of a creature's beliefs. Beliefs are individuated by their contents; so if a creature cannot distinguish between the contents of various beliefs, then it couldn't very well be said to have the concept of belief. Language is what allows creatures to make the contents of their beliefs appropriately stable and manifest, so that they can become appropriately accessible to both public and private survey.

This defense of the second subconclusion is interesting, and more developed than Davidson's. Still, it is not above challenge. Fellows asserts that language is *one* route by which creatures can come to be aware of the distinction between changes of mind and changes in the world, but he needs to argue that this is the *only* way. Otherwise, he has slipped in a sufficient condition where a necessary one is called for. But it's a mystery how Fellows would go about arguing for this stronger claim, for it's far from clear that the concept of belief really does require that belief contents be stabilized in a way that could only be provided by language. One doesn't need to make the contents of one's own beliefs explicit in a public language before one is able to revise those beliefs in the face of countervailing evidence and recognize that that is what they have done. Though it falls far short of a conclusion demonstration that one must be a language-user to have second-order intentionality, Fellows' concerns nevertheless succeed in shifting the burden of proof onto those who would reject the second subconclusion. So perhaps the most charitable thing to conclude is that Fellows has issued an important challenge to those defending the idea that one could have the concept of belief in the absence of linguistic capacity. Their task is to show Fellows how creatures could make the specific contents of beliefs manifest to themselves without availing themselves of the expressive resources of a public language.

5. Despite my objections to Fellows' reconstruction of Davidson's argument, I'm no fan of the idea that the intentional capacities of non-linguistic animals are continuous with our own. As linguistic beings, we evidently engage in communicative behavior that is *meaningful*, in the sense that it can be evaluated as true or false with respect to standards maintained and enforced by others in our linguistic community. It is an interesting philosophical question how creatures

could ever come to institute such *social* norms of correctness. By no means, however, does this recognition of the special status of linguistic intentionality militate against the possibility of animal belief. Indeed, one might want such a concept in order to tell a satisfying story about how linguistic activity could have emerged. And it does make sense to say that animals can make mistakes about the way things are in their surroundings. Insofar as they engage in patterns of behavior that is discernibly rational, they can be said to have states of mind with a mind-to-world direction of fit. Whether or not we dignify those states with the term "belief" is, as I hinted at earlier, a semantic squabble of no great philosophical significance.

Instead, my chief complaint with Davidson and Fellows is with their overall strategy of attaching special significance to the concept of belief. For it suggests that we must be able to speak about beliefs before we can speak about anything at all. And it reinforces the broadly Gricean idea that second-order intentionality is intimately bound up in our capacity to engage in linguistic communication. Requiring second-order intentionality (such as in the form of Gricean communicative intentions) for linguistic meaning threatens to excavate an unacceptably wide gulf between us and those from whom we presumably evolved. It seems most likely to me that those engaged in the most primitive forms of linguistic communication - those whom Sellars might have referred to as "our Rylean ancestors" - didn't have the conceptual resources to construct either thoughts or talk that were about thought. Their initial stock of concepts probably did not include anything so intangible as that of BELIEF. Full-blown second-order intentionality (that involving the concepts of BELIEF and other states of mind) only emerged later, after they discovered how to think reflectively about their own already extant linguistic activity. Instead, this initial stock most likely consisted of concepts for everyday external things (berries, rabbits,

etc) and mundane activities (eating, sleeping, and so forth), none of which would seem to be ground-level, mastery of which is required for a creature to have any mastery of concepts whatsoever. Indeed, I doubt that there are any such concepts at all.

¹ D. Davidson, 'Thought and Talk', *Mind and Language*, S Guttenplan (ed.) (Oxford: Clarendon Press, 1975), 7-23 and 'Rational Animals', *Actions and Events*, E. LePore and B. McLaughlin (eds.) (Oxford: Basil Blackwell Press, 1985), 473-80.

² See, for instance, J. Bishop, 'More Thought on Thought and Talk', *Mind* LXXIV (1980), 1-16, P. Smith, 'On Animal Beliefs', *Southern Journal of Philosophy* 20 (1982), 503-12, and A. Ward, 'Davidson on Attributions of Belief to Animals', *Philosophia* 18 (1988), 97-106. Another reconstruction of Davidson's argument, similar in spirit to Fellows' project, can be found in P. Moser, 'Rationality without Surprises: Davidson on Rational Belief', *Dialectica* 37 (1983), 221-6.

³ R. Fellows, 'Animal Belief', *Philosophy* 75, No. 294 (2000), 587-598.

⁴ See Fellows, p. 595. Fellows presents and addresses these subconclusions in the reverse order, but this ordering seems more natural to me, and it follows Davidson's presentation of the argument in 'Rational Animals', p. 478.

⁵ See 'Rational Animals', p. 479.

⁶ See Smith, 'On Animal Beliefs', p. 507.

⁷ D. Dennett, *Kinds of Minds* (New York, NY: Basic Books, 1996), chapter 5.

⁸ See W. Sellars, 'Empiricism and the Philosophy of Mind', section XII, reprinted in *Knowledge, Mind, and the Given*, W. deVries and T. Triplett (eds.) (Indianapolis, IN: Hackett, 2000), 258ff.

I should probably add that Sellars need not maintain an unfriendly attitude towards animal belief. Just because the concept of belief is *modeled* upon overt linguistic performances, it doesn't follow that that concept couldn't be properly extended to creatures who don't engage in linguistic activity. All that follows is that the ideas of inner intentional states wouldn't occur to those who weren't already engaged in linguistic activity.

⁹ 'Rational Animals', p. 479-80.

¹⁰ Bishop, 'More Thought on Thought and Talk', pp. 9-10.