Animal Agency

Do any non-human animals have agency? The answer, one might think, unless one is using the term 'agency' in an unusually inflated sense, is obviously 'yes'. Agency is the power to act – and surely animals act in various ways? Some animals hunt; some build dams; some collect and store food for the winter. Many animals play, and fight, and tend to their young. Some groom one another; some engage in complex courtship rituals. Some can be trained to help with human tasks such as rounding up sheep (dogs), or retrieving fish (cormorants); and arguably the most intelligent of all, the chimpanzees, have even been trained to communicate quite complex messages by use of a rudimentary sign language. In view of these facts, it is surprising, then, that animals have figured so little in the philosophical discussions of the nature of agency which have taken place in the last century or so; and indeed that their status as agents continues to be denied (e.g. Stoecker, 2009). Of course, human agency is distinctive in numerous respects – many of which are associated with other important philosophical concerns: e.g. with practical reason, with morality, with self-control – which helps to explain the focus of philosophers on specifically *human* action. But as I shall make clear below, the neglect of animal agency by philosophers has not only prevented us from having a clear view of the animal phenomenon; it has obscured our view of the metaphysics of the human one as well. For we humans – let it not be forgotten – *are* animals. Human agency is, first and foremost, a species of animal agency (though admittedly a very special one) and attempts to say what agency is which entirely ignore that fact are likely to be doomed to failure.

(i) The Causal Theory of Action

According to what has arguably been the dominant theory of action over the past 50 years or so, the Causal Theory of Action, or CTA, actions are merely events - generally speaking, bodily movements, such as my right arm's rising, or my left eye's closing. What marks actions out from events which are *not* actions, on this view, is a distinctive sort of causal provenance. My arm's rising counts as an

action if and only if it is brought about by the right sorts of things and in the right sort of way[i] - for example, perhaps if it is brought about by my *desire* that my arm should rise (which perhaps is in turn brought about by my *desire* to ask a question in a seminar, together with my *belief* that I can signal my desire to ask a question by raising my arm), these beliefs and desires together constituting the reasons for which I act (Davidson, 1963). Other philosophers have adopted slightly different, but still basically causal/functionalist strategies, in which the distinctive precursors of actions are not beliefs and desires, but rather *other* sorts of mental state and event, for example, intentions (Bratman 1987; Mele, 1992) or volitions (Locke, 1975 [1689]; Mill, 1970 [1872]). Fundamentally, though, the idea behind all these views of action is the ultimately dualistic one that an action is essentially a physical thing which is produced by a mental thing, or set of mental things (or else perhaps the ultimately not very different view that we should think of an action as a *complex* entity composed of both the mental cause(s) and the physical effect together). The question whether animals act, then, given a theoretical background of this kind, becomes the question whether the purposive bodily movements which we undoubtedly observe taking place in their limbs, wings, legs, fins, and so on, have the appropriate sorts of mental cause.

Do these bodily movements have the appropriate sorts of mental cause? This is a difficult question to answer. For there have always been many people who have doubted the wisdom of ascribing mental states to animals. Some, for example, have argued that animals cannot have beliefs (Davidson, 1975 and 1982; Stich, 1979). Beliefs are propositional attitudes, ascribed by way of conceptual content – but how do we know whether e.g. to ascribe a dog the belief *that the cat is up the tree* when it isn't at all obvious that the dog can distinguish cats from a range of other small animals that it also likes to chase (e.g. squirrels), or whether it makes any distinction between trees and bushes (to use an example taken from Stich)? And if we have no idea how to provide content for beliefs of a sort that would fit the doggy mind, how sensible is it to think of dogs as having beliefs at all? It might seem less easy to see what is wrong with supposing that animals have *intentions*, but doubt here has also been expressed. One might wonder, for example, whether we can really be sure

that animals *know what they are doing* in the way that is arguably required to justify the attribution of intentions to them. Birds certainly behave purposively when they build nests, for example, choosing moss and grasses carefully and weaving the whole together so as to produce astonishingly well-crafted receptacles for their eggs and for the subsequent rearing of chicks. But do they have any idea why they are building the nests? Do they have any conception of what is to take place, shortly, within them? – particularly during the first year of breeding, when they have had no previous experience? Or is the whole process merely instinctive, done without any sense of the end to which the nest is a means? And if the answer is that the process *is* instinctive, and comes with no knowledge or understanding of its point or purpose, can it be correct to suppose that the bird really intends to build a *nest*?

One person's modus ponens, however, is another person's modus tollens. If a certain philosophical view of action leads to the conclusion that animals can't act, perhaps it is that view of action that should be questioned? Harry Frankfurt is one of the most influential philosophers to have pointed out how wrong the conclusion that animals aren't agents seems. Frankfurt (1978) discusses the difference between what happens when a spider moves its legs as it travels across the ground, and what goes on when its legs move in exactly the same way because they are being manipulated by a child who has attached strings to them and is moving the spider like a puppeteer. Frankfurt notes that in the former case, unlike the latter, the movements are attributable *to the spider* as the agent of the movements, and that this would seem to be just the very same distinction as the distinction between a person's actively raising their arm and a person's arm going up because someone else has raised it. Spiders, then, would seem to be agents too. Moreover, it seems that we can know that spiders are agents in this sense, without knowing anything about whether they have mental states it is worth calling beliefs or intentions. Even if agency *does* demand the existence of certain forms of mentality, might it not be overkill to assume that it requires these comparatively sophisticated states? Perhaps if animals can, for example, *see* certain objects; *want* certain things; and *try to get them* (as might seem much more plausible than that they have beliefs or intentions),

these capacities would be sufficient to constitute appropriate precursors for exercises of something we could recognise as agency?

What kind of theory of action might make possible such an extension of agency to animals? One possibility, as suggested above, might be simply to downgrade the sophistication of the mental states required to serve as causes of action to things which might be simpler than beliefs and intentions. Another, more radical suggestion might be that we should try a different way of privileging as definitional the feature of agency to which Frankfurt's argument draws attention – the fact that we attribute certain of the movements which go on in the world to *agents.* Our dualistic heritage encourages to think about the production of effects in the world by agents as though that were equivalent to the production of effects in the world by *minds*, by mental states and events – but this may not be the only way in which one might be able to think about what it means to attribute some happening to an agent as its source. The idea of an entity as a thing which *has* a body, rather than just *being* a body is an interesting one – and even if a thing's having a body entails that it has some mental capacities (as seems likely), it might not follow that each occasion on which an agent controls its body in the way we think distinctive of action is an occasion on which the agent's mental states are directly involved in the exercise of that control. To see this, it helps to think about types of human activity which are a bit less deliberative and intentional than the sorts of examples that philosophers tend mainly to choose – in particular, it can be useful to consider what have sometimes been called 'sub-intentional actions', such as absent-mindedly scratching an itch or distractedly fiddling with one's jewellery. Are these actions? I think so. Like the movements of the legs of Frankfurt's spider, they are due to the agent, even if they are not intended, or even wanted by that agent. But if they are actions, we are going to need a different and less mentalistic way of thinking about what is essential to agency than the one which the CTA and its descendants have suggested to us, one which might in turn make it seem much more straightforward to accord agency to at least some animals.[ii]

(ii) Which animals?

As that last phrase 'at least some' suggests, though, it is not yet clear how widely even a more generous and flexible conception of agency can be extended across the animal kingdom. Are all animals agents? One might worry that extending the attribution *too* widely risks losing sight of the nature of the distinctive phenomenon we are trying to capture. The movements of some very simple creatures, for example, seem explicable as fairly simple mechanisms, and one might think that the attribution of agency is excessive in such cases. Take a paramecium, for example. A paramecium is a single-celled creature which moves through water by means of cilia on the outside of its body. But is it really right to say that *the paramecium* is beating its cilia? One reason for thinking the answer might be 'no' might be the thought that no consciousness, no thought, no real mentality could really be supported by a single cell. Another, though, might relate to the idea that in a case like this, an entirely mechanistic explanation for the relevant motions threatens to make the invocation of agency redundant. There is detailed research, for example, showing that the direction of movement of a paramecium depends on a fairly small number of factors, including water temperature and pH value (Glaser, 1924; Chase and Glaser, 1930). I think it would not be surprising if it were to turn out that the motions of paramecia are pretty much entirely predictable given knowledge only of a fairly manageable number of environmental variables. But if this did turn out to be the case, one might think that the paramecium seems to fall into the same category as a robot or a self-driving car, something whose movements are settled by a complex interaction between its (in this case, biological) programme and the environment, but perhaps not one we need to think of as agential.

This raises the important question, though, of what creature-environment interactions have to be like in order to warrant the attribution of agency to a creature. Some philosophers would doubtless argue that the movements of *all* animals (including even those of humans) are *also* settled by complex interactions of a basically mechanistic sort, differing from those we can observe in the paramecium only in degree of complexity. One interesting and extremely influential manifestation of

this view is encapsulated in Daniel Dennett's idea that we adopt an 'intentional stance' in order to explain the behaviour of certain kinds of systems – humans, at least some animals, and potentially also some artificial systems, such as chess computers and thermostats – and that there is nothing more, or less, to the question whether a system has intentional states like beliefs and desires than there is to the question how well the intentional stance works to explain and predict the behaviour of those systems (Dennett, 1987). Dennett himself is mainly interested in the question what is required for the ascription of intentional mental states, rather than in *our* question, which is what *agency* requires. But one can see that a similar strategy might work in the case of agency too. Perhaps agents are those things which it proves fruitful to treat *as* agents – that is to say, with respect to which it pays off to consider what informational state the creature may be in and what its goals are, in order to decide what it may be about to do. Of course, as Dennett himself points out, one can in a sense treat almost anything as an agent – for example, he notes that it would be possible to predict the behaviour of a lectern by ascribing it the belief that it is currently located at the centre of the civilised world and desiring above all else to remain at that centre. We'd then predict that it wouldn't do anything – and hey presto! – we'd be right. But in this case, as Dennett says. 'we get no predictive power from [the intentional strategy] that we did not antecedently have. We already knew what the lectern was going to do – namely nothing' (p. 23). Perhaps, then, the mark of a true agent is that what one might call the agential strategy, a strategy which assigns the system or creature in question such things as perceptual knowledge and a range of goals, gives us predictive power that we did not have available before. It is not that we could not *in principle* predict the motions of an agent from what Dennett calls 'the physical stance', a stance which involves determining the physical constitution of a system and the physical nature of the impingements upon it and using the laws of nature to predict the result. But that would be an immensely impracticable task, because the calculations would be so astronomically large and complex. Where agents are concerned, the agential strategy simplifies matters for the purposes of explanation and prediction.

This Dennettian view is designed to be compatible with the view that ultimately, everything is mechanism – and if there is any kind of differentiation to be made amongst the various varieties of mechanism, the differences will be matters of degree, not of kind. Because the idea is that whether or not a system is an agent is a matter of how fruitful it is for prediction and explanation to apply the agential stance, the line between the agents and the non-agents will not be sharp. An alternative view, though, might be that agency is distinctive, in part, because of its modal features – that an exercise of agency is always such that it *does not have to happen*, at least not in quite the way or at quite the time that it does happen, so that agency is an essentially indeterministic phenomenon. If the motions of the paramecium are all strictly determined by such things as the temperature and pH of the water, and so on, then intuitively, nothing really remains to be settled *by the paramecium* – and this might be thought to justify the view that the paramecium does not count as an agent in respect of those motions – any more than I count as an agent of my motions when my unconscious body is flung from the top of a building and I fall to the ground. But when a cow wanders across a field, say, perhaps we might be more inclined to believe that it makes motions happen in its body that need not have happened in precisely those ways and at precisely those times, even holding fixed all prior circumstances. Of course, animals act in accordance with their instincts – and so there are *limits* to what any given animal might possibly do on any given occasion. Cows are not free to decide not to graze, for example. But where we attribute agency, this view would have it that we attribute at the same time a certain freedom – a movement only really being worth attributing *to a creature* (rather than, say, to antecedent events and interactions going on in its parts) if the creature has the power to *settle* something of what results. Of course, much more needs to be said about what kind of metaphysical picture of the world could allow for such settling – how certain results can obtain neither because they are determined to do so, nor by chance. But if such a metaphysical picture is available, this view might hold out the promise of respecting Frankfurt's point that we need an account of agency which explains properly what it means to attribute some movement *to a*

creature, and which does not simply identify that with attributing the movement to certain mental states and events.[iii]

Even if some animals do have agency, it has to be admitted, of course, that human agency is a distinctive phenomenon. In the next two sections, I want to turn to look at two important potential differences between animal and human agency.

(iii) Rationality

Are animals rational? That depends, of course, what is meant by 'rational' – a matter on which philosophers do not agree. It seems fairly clear that animals can do certain sorts of *reasoning* task. Some crows, for example, can solve puzzles of quite considerable complexity to release tools by means of which to extract pieces of food which they can see, but which are inaccessible except by way of those tools.[iv] Some have suggested that since animals can reason, since they seem to be able to make inferences, that that settles the question in favour of their rationality. But this is not really clear. The concept of rationality has a long history in philosophy – and one very powerful conception of rationality connects rationality with *normativity*. As Robert Brandom puts it, 'To be a rational being in this sense is to be subject to a distinctive kind of normative appraisal: assessment of the reasons for what one does … Rational beings are ones that *ought* to have reasons for what they do and *ought* to act as they have reason to' (Brandom 2009: 2-3).

Ought animals to have reasons for what they do? It seems quite natural to suppose that sometimes at least, some animals *do* have reasons for what they do – that a cow might run at a dog because she wants to protect her young. But in many ways, it seems strange to bring animals within the fold of our 'ought' judgements. We can perhaps make sense of a quasi-normative 'ought' attached to an imputed teleology: 'if that nest is to stay in the tree, the blackbird ought to choose a stronger branch'; 'the salmon ought to swim up the fish ladder if they're to reach their spawning grounds',

etc. There are better and worse ways for animals to achieve what we take to be their ends, and so we can make some kind of sense of the idea that an animal ought to do something in one way rather than another. But other parts of the structure that generally pervades the 'space of reasons' (Sellars, 1956) are manifestly lacking in the case of non-human animals. Animals are not answerable to anyone for their mistakes; we do not (because we cannot) demand justifications, excuses or explanations from them. One can train a dog but arguably one cannot *get it to see* that there are better ways for it to do things.

One line of thinking which I find helpful when considering the question whether any non-human animal is rational draws on what can reasonably be supposed to be an Aristotelian thought. Aristotle is well known for his claim that human beings are rational animals – and moreover that it is in some sense part of the *essence* of a human being to be rational. But of course, if this idea is not to fall at the first hurdle, it needs to be explained why it is not disproved by the existence of manifestly non-rational human beings – humans, for example, who are unable to speak or think and who do not show any evidence of the powers associated with reason or reflection. One response is that the Aristotelian account does not intend to assert that rationality is part of the essence of any *individual* human beings; it is rather part of the essence of the *kind*, human being, in that in some sense or other it characterises what human beings are like when development follows a certain teleologically structured path in the way it is supposed to. Moreover, Aristotle clearly thought that 'rational' was a *special* differentia of the genus 'animal' - unlike 'mammal' or 'two-footed', which arguably are also parts of the essence of the kind human being (Boyle, 2012). But in what way special? Boyle argues, to my mind very convincingly, that rationality was thought of by Aristotle not merely as a power or collection of powers which is possessed uniquely by human beings – but rather as a distinctive form of *psyche*, or soul, as important and as fundamental a kind of soul as the Aristotelian vegetative and animal souls. Just as animality might be thought of as a way of being which fundamentally builds on, but also utterly transforms, the kind of life found in plants, so he suggests, rationality can be thought

of as a way of being which fundamentally builds on, but also utterly transforms the nature of the animal life to which it belongs.

Different participants in the debate about animal rationality, then, may simply be talking past one another. Some participants think of rationality as an individual power, or perhaps a connected set of powers, which can be possessed to a greater or lesser degree – and according to this conception, there can be no very hard and fast divides between the rational and the non-rational, and everything argues for respecting the continuities in nature. For others, though, to call a being 'rational' is not merely to indicate that it is possessed of particular powers, but rather to indicate that the whole form of its life is of a specially transformed kind – a kind which involves the ability to respond to reasons as such (McDowell, 1994); to justify one's actions; and to operate on principles.

(iv) Moral Responsibility

Another interesting (and related) question is whether animals can ever be morally responsible for what they do. Although I suspect most of us think the answer to this question is 'no', it is harder than one might think to justify this answer. One traditional route to the view that animals are not morally responsible agents, of course, goes via the idea that animals are not agents at all – that they have none of the requisite freedom to do otherwise – but as I have tried to suggest above, and have argued in more detail elsewhere (Steward, 2012), that may not be correct. Moreover, it does not seem to be true that no non-human animals have systems that might count as moral *codes*. Some animal societies appear to be based upon certain pro-social norms, the violation of which attracts considerable opprobrium, and which members of the animal society make the appearance of having internalised. Moreover, some animals seem to be motivated by what might be argued to be moral reasons – they feel sad, for example, at the suffering of others and make attempts to ameliorate it (Rowlands, 2013). It might be pointed out that non-human animals have no capacity to scrutinise or discuss their motivations; but perhaps it might be reasonably be replied that there are many

perfectly good – even virtuous – human beings who do little or no such scrutinising either. Might it not be overkill, then, to demand a practice of scrutiny in a morally responsible agent?

Some animals, then, may be involved in a kind of norm-following which might count as a primitive morality; and some may be responsive to moral reasons. But I think it continues to seem implausible that animals are morally *responsible* for their actions. It has been suggested that the key factor which distinguishes what we might call (following Rowlands) morally sensitive *subjects* from morally responsible *agents* concerns the capacity to intervene with respect to one's potentially morally sensitive habits and behaviours, in the light of such things as reflection, or the acquisition of new information (Rowlands, 2013; Musschenga, 2015). On Rowlands' view, 'What demarcates moral subjects from moral agents … is a kind of level of understanding' (Rowlands, 2013, p.239). It is a lack of understanding, though, rather than a lack of the capacity for agency, which appears to be the obstacle to the assignment of responsibility.

It has long been recognised, of course, that animals should be thought of as moral *patients* – that because of their capacity to undergo suffering, we have duties to them. I should like to end on the thought that if we are prepared, in addition, to recognise the status of at least some of them as *agents* (albeit not morally responsible ones), as I believe we should, we might need to give a different and more radical account than is sometimes given of where those duties end. Agents make choices, have ends and live lives – and prima facie, we have a duty to agents to permit them, so far as is consistent with other reasonable moral demands, to exercise that agency in the course of those lives, in the ways that they wish, or may be presumed to wish. We therefore have prima facie duties not merely to avoid causing animal agents pain, but also to refrain from curtailing or interfering unduly with their natural lives. These are duties we have not thus far sufficiently recognised.

RELATED TOPICS: Agency and Causation; Agency and Agents; Rational Agency; Agency and Responsibility; Agency and Morality

Bibliography

Boyle, M. (2012).). 'Essentially Rational Animals', in G. Abel and J. Conant (eds.) Rethinking Epistemology. Berlin: Walter de Grutyer Verlag.

Brandom, R. (2009). *Reason in Philosophy*. Cambridge: Harvard University Press.

Bratman, M. (1987). *Intention, Plans and Practical Reason.* Cambridge: Harvard University Press.

Chase, A. and Glaser, O. (1930). 'Forward movement of paramecium as a function of the hydrogen ion concentration'. *Journal of General Physiology* 13: 627-36.

Davidson, D. (1963). 'Actions, reasons and causes'. *Journal of Philosophy* 60, pp. 685-700, repr. in Davidson *Essays on Actions and Events*. Oxford: OUP, 1980: 3-19.

Davidson, D. (1973). 'Freedom to Act'. In T. Honderich (ed.) *Essays on Freedom of Action*. London: Routledge and Kegan Paul, 137-56. Repr. in Davidson (1980): 149-62.

Davidson, D (1975). 'Thought and talk'. In S.D. Guttenplan (ed.) *Mind and Language* (Oxford: Clarendon Press): repr in Davidson *Inquiries into Truth and Interpretation* (Oxford: OUP):

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

Frankfurt, H. (1978). 'The problem of action'. *American Philosophical Quarterly* 15, repr. in his *The Importance of what we Care About* (Cambridge: CUP): 69-79.

Glaser, O. (1924). 'Temperature and forward movement of paramecium'. *Journal of General Physiology*, 7: 177-88.

Locke, J. (1975)[1689]. An *Essay Concerning Human Understanding* ed. P. Nidditch. Oxford: OUP.

Mele, A. (1992). *Springs of Action*. Oxford: Oxford University Press.

Mill, J.S. (1970) [1872] *A System of Logic*. 8th edn. London: Longman.

Musschenga, Albert (2015). 'Moral Animals and Moral Responsibility', Les ateliers de l'éthique 10(2): 38-59.

Rowlands, Mark. (2013). Can Animals Be Moral? Oxford: Oxford University Press.

Steward, H. (2012). *A Metaphysics for Freedom*. Oxford: OUP.

Stoecker, R. (2009). 'Why Animals Can't Act'. *Inquiry* 52: 255-71.

Stich, S. (1979). 'Do animals have beliefs?' *Australasian Journal of Philosophy* 57: 15-28.

Further Readings

Boyle, Matthew (2012). 'Essentially Rational Animals', in G. Abel and J. Conant (eds.) *Rethinking Epistemology*. Berlin: Walter de Grutyer Verlag.

In this paper, the author attempts to argue the case for thinking that being an animal is fundamentally transformed when the powers distinctive of rationality are present, in a way that justifies the idea that humans are animal agents of an utterly distinctive kind, in many ways non-continuous in virtue of that transformation, with other animal agents.

Frankfurt, Harry. (1978). 'The Problem of Action'. *American Philosophical Quarterly* 15: repr. in his *The Importance of what we Care About* (Cambridge: CUP): 69-79.

In this paper, Frankfurt argues against the causal approach to making the distinction between what an agent does and what merely happens to him; and makes out the case for the view that the conditions for attributing the guidance of bodily movements to a whole creature, rather than to a mechanism within one, obtain outside the human arena.

Hurley, Susan and Nudds, Matthew (2006). *Rational Animals*. Oxford: Oxford University Press.

A useful collection of articles, which asks interesting questions about what it might mean to attribute rationality to a creature and brings to bear on the issues many empirical studies of the kinds of tasks animals are able to perform and psychological theorising concerning the processes which might underlie them.

Rowlands, Mark. (2013). *Can Animals Be Moral*? Oxford: Oxford University Press.

This book argues, against centuries of tradition, that animals can act for moral reasons and thereby count as what Rowlands calls 'moral subjects'.

Steward, Helen. (2012). *A Metaphysics for Freedom* Oxford: Oxford University Press, Ch 4.

This chapter defends the view that many animals are agents; and also considers the difficult question of how it is to be decided which animals are agents, and why.

---

[i] 'In the right kind of way' because it is accepted by the proponents as well as the detractors of the CTA that certain sorts of counterexample show conclusively that more is needed for an action to have taken place than merely the causation of the appropriate sort of bodily movement by the appropriate sort of mental cause; see Davidson (1973) for a discussion of the so-called 'deviant causal chains' problem as it applies in the case of action.

[ii] See my (2012) for a detailed discussion of how this might be done by appeal to the idea of top-down causation.

[iii] See my (2012) for an attempt to develop the wanted metaphysics.

[iv] See e.g. BBC, 2014, accessed 8th January 2019 <https://www.wimp.com/a-crow-solves-an-eight-step-puzzle/ to see a crow in action.