Animal Understanding and Animal Self-Awareness

Peter Woodford

Abstract Theorists arguing that non-human animals (simply animals from this point forward) are self-aware often make the case on the basis that non-human species *understand* aspects of themselves and the world, and these forms of understanding indicate self-awareness. But the notion of understanding in this context is often taken for granted. This article aims to analyse the nature of animal understanding to clarify the kind of understanding that matters for discussions of self-awareness, namely, self-understanding. I proceed by drawing on discussions of understanding offered in contemporary epistemology, and then by discussing the relevance of the concept of self-understanding here for discussions of animal self-awareness. I argue that the kind of self-understanding relevant to discussions of animal self-awareness is specifically an animal's understanding of its own causal influence on the world and on others.

Keywords Animal minds · Animal cognition · Comparative psychology · Philosophy of biology · Self-awareness

1 Introduction

While recent work in philosophy and the sciences has defended the notion that non-human animals and pre-verbal infants understand the world, the idea that animals understand themselves seems to appear less plausible (Grimm 2016; Baumberger et al. 2017). The aim of this article is to show that there is a defensible concept of self-understanding that allows for the recognition of non-human forms of self-understanding. I do not aim to go all the way to showing that non-human animals *do* understand themselves, but rather to articulate a concept of self-understanding that opens space for the possibility of non-human self-understanding. Moreover, I argue that this is the form of understanding that is relevant to the case for

P. Woodford (✉)
London, UK

J. M. Viejo, M. Sanjuán (eds.), *Life and Mind*, Interdisciplinary Evolution Research 8, https://doi.org/10.1007/978-3-031-30304-3_2

animal self-awareness. The article proceeds in a number of steps. First, I analyse the nature of animal understanding by drawing on recent discussions in epistemology. Next, I defend a concept of animal self-understanding that is based on these discussions of animal understanding. I then show that recent discussions of animal self-awareness, which have highlighted the importance of various forms of agency – particularly bodily and social – in the concept of animal self-awareness, can inform and further develop the concept of animal self-understanding. Finally, I show how the concept of self-understanding defended here fits into recently defended evolutionary views of the emergence of sentience and self-awareness.

2 Understanding and Animal Understanding

What is understanding? The philosophical literature on understanding is vast, and the landscape of debate shows disagreement over fundamental conceptual questions. Nonetheless, some consensus seems evident around the notion that understanding is something more than knowledge, if knowledge consists in the justified affirmation of true beliefs. A prevalent conception of understanding is that it is a grasp of causal relations and dependencies that hold within various arenas of the world (Zagzebski 2001). To take an example borrowed from a recent article, one might know *that* a house caught on fire due to faulty wiring, without *understanding* how a house might catch on fire due to faulty wiring (Grimm 2016). Thus, unlike justified affirmation of propositions that are true, understanding involves knowledge of how things *work*; someone who understands can grasp the patterns and causal dependencies in the world that, at various levels, undergird the facts we might affirm. Another common example concerns the familiar experience of taking a test. Some students might answer questions correctly on an exam by rote memorisation of a list of true facts, affirmed through justified trust in the epistemic testimony of educators, without really *understanding* the reasons why these facts obtain. So, I can pass the test in biology without really understanding much about the way the biological world works. According to this line of thought, understanding can involve propositional knowledge, but it is better thought of as a grasp of how things *work* in a given region of world.

Given this account, it might appear that the case for animal understanding is not very strong. There are two reasons this might be the case: first, understanding causal patterns is often thought of as *scientific* understanding, and, second, such understanding still requires propositional knowledge 'that' certain facts obtain. Let us consider these in turn. It might appear that true understanding of causal relations is involved in *scientific* understanding, or the kind of knowledge of nature's causal structure that has been extremely hard won through the advancement of the natural sciences. To *really* understand the event of a house catching on fire, we need to know something about physics, chemistry, and thermodynamics. Or, to understand why a lake is frozen, we need to know about geological patterns, meteorology, and the earth's orbit. If such sophisticated forms of culturally preserved and transmitted

causal understanding are required to *really* understand, then the case for understanding in animals appears a non-starter. However, some epistemologists have recently pushed against an account of understanding that sets the bar at such fine-grained, extensive knowledge of causal relations (Grimm 2016; Baumberger et al. 2017). For example, if a plastic bag is rolling along the ground and flying up in the air, I understand that it is the strong wind that is moving it. If a lake is frozen, I understand that it is due to the cold. These explanations may be a coarser, less fine-grained understanding of causal patterns than many today have come to expect, but it is the basic understanding that sciences then go on to deepen and clarify. The point here is that my understanding need not extend as far as the causes of the wind itself, or of the cold, for me to properly understand what is going on with the lake or the plastic bag. Still, both mechanistic, fine-grained causal understanding of nature in the sciences and more 'everyday' understanding of how the world works suggest that my understanding is a matter of grasping causal patterns and relations of dependency among phenomena. As some epistemologists have argued recently, understanding is a matter of grasping how some region of the world works (Grimm 2016).

The second aspect of the problem might appear more difficult. If understanding of causal relations requires mastery of the concept of *causation*, or, further, of propositions asserting various states of affairs and causal dependencies between them, then it might also appear unlikely that animals possess understanding. Here we encounter something of a Hume versus Kant problem. Can an understanding of causal relations and dependencies be indicated by the mere lack of surprise that one event has followed another? Can it be indicated by the mere confident *expectation* that some event will occur after another? Expectations can, of course, involve concepts, but does understanding require grasp of concepts, for example the concept of one event making another event happen? There is not space to settle this debate here, but let me offer something of a compromise to get the rest of the inquiry going. We can grant that there is certainly a difference between 'Kantian' conceptual understanding that involves the concept of causation and the understanding indicated by 'Humean' expectation; nevertheless, 'Kantian' understanding involving the concepts need not be thought as the measure of understanding *tout court*.

It is certainly difficult to accept the notion that animals understand causal patterns in the world at the level of the natural sciences. But there are well known examples of animal behaviour that support the ascription of such 'Humean' understanding to animals. Experimental research on rats (Blaisdell et al. 2006), corvids (Taylor et al. 2012; Jelbert et al. 2019), primates (Völter et al. 2016), and on pre-verbal children (Gopnik et al. 2001; Kushnir and Gopnik 2005) suggests that non-verbal animals do indeed track accurately the causal relations and patterns between events. To take a recent representative example, Jelbert et al. 2019 tested whether New Caledonian crows could infer whether an object was light or heavy based on whether it could be moved by a breeze generated by a fan. Crows were trained to receive a reward after dropping either a light or a heavy object into a dispenser. Then, birds were able to watch how two suspended objects behaved in front of a fan. They were able to pick out the light or heavy object accurately (whichever they were trained to expect a reward from) 73% of the time, and did no better than chance in control trials without

the fan on. Researchers argue that the crows were rightly inferring the weight of objects by watching how they behaved in front of a fan (Jelbert et al. 2019). Other experiments with crows suggest that they can also infer *agential* causes behind the observation of some phenomena, which has also already been shown in pre-verbal children (Saxe et al. 2005; Saxe et al. 2007; Taylor et al. 2012).

Such experimental results suggest a new possibility. We can recognise that scientific understanding is more extensive and detailed than 'everyday' understanding, and that humans may possess a form of Kantian conceptual understanding that Humean animals lack. Nonetheless, 'everyday' Humean understanding can be recognised *as* understanding of causal relations, dependencies, and patterns, and this gives us what we need to make the case for animal understanding (Grimm 2016). Some might argue that experiments like the one cited above show that animals do possess non-linguistic concepts as well – even if the concept of causation is not necessarily one of these concepts. This is a debate worth having, but my focus at this stage of the argument is only to defend the plausibility of a conception of understanding that involves only the ability to track causal patterns accurately and to develop expectations about them.

3 Understanding and Self-Understanding

If the development of expectations about patterns in the world that do, indeed, match such patterns counts as a basic form of understanding, we can also suggest some implications for the notion of self-understanding. Divergences in the concept of self-understanding mirror those of the concept of understanding. For example: do I need to have scientific understanding of causal patterns in my body, or about how I will perform in psychological experiments to have self-understanding? Just as in the case of a frozen lake, we can distinguish between more or less fine-grained knowledge here. Let us say I do not like cilantro, and I do not like it because it leaves a soapy taste in my mouth. I understand that I have no desire for cilantro because every time I taste it, it leaves a soapy taste in my mouth. I do not need to understand biochemistry or the physiological mechanisms of taste sensation, or even that a gene has been isolated that appears to explain variation in the taste of cilantro (Eriksson et al. 2012). I understand that it tastes bad, and so I avoid it; this too constitutes a form of 'everyday' understanding of causal patterns involving the traffic between my sensations and the world. Of course, we can admit that there are more or less sophisticated forms of self-understanding, but such an 'everyday' understanding of how the world occasions my own sensations is understanding nonetheless.

Do animals have this 'everyday' sort of self-understanding? To stick with the example above, the difficulty here is to determine whether a behavioural response, such as avoiding eating something that an animal has tasted, involves an awareness of a causal pattern. Experimental research on the phenomenon of *conditioned taste avoidance* (CTA) suggests that avoiding food that has been manipulated to taste bad or toxic is widespread across a variety species including mantises, blue jays, slugs,

and molluscs (Bures et al. 2002; du Toit et al. 1991; Parker et al. 2008; Reilly and Schachtman 2008). The survival value of striking an optimal balance between openness to novel foods and avoidance of toxic foods is thought to explain the prevalence of this phenomenon and its primitive neurological underpinnings. Sensations that are experienced after eating a food lead a creature to develop expectations that steer them away from foods in the future. Examples of CTA can involve classical conditioning in which an animal is trained to expect a desired food or to exhibit vomiting in the expectation of a toxic food in response to an arbitrary stimulus (such as ringing a bell); however, CTA is often studied as a more particular phenomenon in which an internal state of the organism – such as pleasure, equilibrium, or disequilibrium – serves as either the reward or punishment for ingesting a food source (Bures et al. 2002).

Studies of CTA, I think, suggest that animals can develop an understanding of how *they* work through an understanding of causal patterns involving their own sensations. Just as I arrive at a form of self-understanding by tasting cilantro, so might animals come to understand the traffic between their bodily responses and the world. We can ask about the relation between awareness and understanding for such cases as well. Here, it also appears that awareness is a necessary, but not sufficient, condition for self-understanding; furthermore, understanding is a sufficient, but not necessary, indicator of awareness. If a creature is not aware of its own bodily states, psychological states, or behaviours, it cannot be said to understand causal patterns that involve them.

Just as animals need to be aware of some features of the external world to understand causal patterns therein, so too do they need to be sentient and aware of the stuff in the world that occasions their own sensations. These need not be complex sensations; it seems that basic experiences of pain and pleasure are enough. Philosophers working on the nature of animal minds have argued that the most basic form of self-awareness is *bodily self-awareness* (DeGrazia 2009; Bermudez 1998). While bodily self-awareness is thought to involve more than just sentience (more on that in a moment), the ability to experience basic sensations such as pleasure or pain is required. For example, David DeGrazia argues that bodily self-awareness can involve anything from temporally immediate experiences of pain, pleasure, thirst, hunger to more complex emotions such as fear in anticipation of danger, or excitement in anticipation of something pleasurable. These more complex forms of awareness involve memory and/or projection, and some examples of conditioned taste avoidance may belong in this category insofar as they involve anticipation or expectation. The main claim here is that self-understanding ought to be understood as the ability to track accurately causal relations involving one's own sensations and the world.

4 Self-Understanding and Agency

The previous sections made the case that animals should be said to understand themselves if they develop accurate expectations about causal patterns involving the traffic between the world and their own bodily sensations. Yet, one might worry that only the presence of immediate sensations and the avoidance of noxious stimuli is not enough to constitute self-understanding. After all, such avoidance behaviour might be a simple, automatic response to an environmental stimulus, no different from a thermostat. Examples of conditioned taste avoidance show that organisms can learn to act appropriately – adaptively – as a result of their own sensations. But, given the evolutionary primitiveness and prevalence of this phenomenon across species, including molluscs, something more seems necessary. In addition to having rudimentary forms of subjective experience such as pleasure and pain, as in examples of taste, self-understanding seems to involve the capacity to behave flexibly in response to learning how phenomena 'out there' both affect one and *are affected by one*.

Consider a more sophisticated form of self-awareness that in turn suggests a more sophisticated form of self-understanding. Experiments on 'self-agency' tested whether or not captive, trained rhesus macaques could distinguish a computer icon that they were controlling with a joystick from an icon that was moving randomly on the screen (Couchman 2012, 2015). Couchman et al. define 'self-agency' as the awareness that some actions and consequences are self-generated rather than the result of external forces (Couchman 2012, 2015; Hoffman et al. 2018). These studies found that captive and trained macaques could identify icons that they were controlling with the joystick with the same success rate as humans. Their ability to do so appears to show that they track the difference between events *caused by them* and events that were simply happening, but not caused by them.

Couchman et al. argue that 'self-agency' results from an integration of cognitive information involving sensory-motor cues, prior expectations about the effects of one's actions, and perception of the outcomes of one's actions. The explanation Couchman et al. offer for the evolution of self-agency is that is important for an animal to be able to behave flexibly in uncertain situations in which a habitual response is not adaptive (Couchman et al. 2009). They argue that an awareness of one's own agency affords an animal greater ability to control the outcome of actions in situations that present novel or unfamiliar features. Interestingly, this understanding of why self-agency evolved matches some accounts of the origins of consciousness itself. As DeGrazia reports, Cabanac et al. argue that consciousness arose from the ability to integrate information from multiple senses (internal and external) and to respond flexibly rather than automatically (Cabanac et al. 2009; DeGrazia 2019). While Cabanac et al. argue that such integration and flexibility emerged in amniotes, early land-dwelling mammals, it is at least clear that it is quite developed in the self-agency displayed by rhesus macaques. The possibility of distinguishing between more or less 'automatic' behaviour in animals – more or less ability to control and flexibility in one's response to a stimulus – allows for the recognition of degrees of

animal self-understanding, just as it does for degrees of animal self-awareness, that are indexed to the flexibility and control animals are able to exert over their own behaviour in various domains.

David DeGrazia recently amended his own concept of bodily self-awareness to foreground the additional criterion of agency. *Bodily agential self-awareness*, he argues, is indicated not only by presence of sensations or by automatic pursuit and avoidance behaviour, but by the ability act flexibility in response to stimuli as a result of learning (DeGrazia 2019). This further criterion allows us to avoid the thermostat problem in the case of self-understanding. We should be convinced that the animal already possesses more rudimentary awareness of its own bodily states, for example that it experiences pain, pleasure, hunger, and other forms of bodily sensations. Agential self-awareness should be thought of as developing out of the ability to behave flexibly in response to awareness of such sensations. It too involves the development of expectations about causal patterns, but here these are causal patterns that exist between such subjective senses and the external world. DeGrazia argues that such agential awareness exists across many reptile, mammal, and bird species, and there is some indication that it may also exist in insects (Barron and Klein 2016). He also argues that such agential awareness involves an ability to form some kind of spatio-temporal 'map' of the world (DeGrazia 2019). Wherever we draw the line, such abilities indicate not only the presence of self-awareness, but also of self-understanding, in which an animal has oriented itself within some set of expectations about its own influence on the world.

5 Social Self-Understanding, Social Agency and Social Self-Awareness

The previous section described agential self-understanding as understanding of causal patterns that involve an animal's own actions and the effects of an animal's own actions. A fuller concept of self-understanding can now be given: self-understanding requires more than sentience or taste avoidance behaviour. It requires that an animal can use information about its own actions and capacities for action to perform behaviours that are not merely 'automatic'. Flexible behaviours suggest that animal assess their own capacities and limitations for action in various aways: Can I make it across the gap? Can I catch the prey? Is the prey worth the effort? Self-understanding, then, is present if an animal develops accurate expectations of causal patterns involving its own actions.

With this concept of self-understanding in hand, we can identify additional forms of self-understanding that can arise when animals that live in stable groups and develop forms of complex social interaction with the same individuals over time. In animal groups that persist over time, in which animals recognise one another, there are often social restrictions on what one can and cannot do. These often take the form of aggressive behaviour directed at individuals who pursue a valued resource, or

who occupy subordinate positions, especially those who transgress their 'rank'. There are also aspects of social structure that permit individuals to 'get away with' certain actions that others might not get away with.

There is extensive evidence that non-human animals understand how social relationships work in the groups that they live in. For example, Dorothy Cheney and Robert Seyfarth's pioneering field experiments with baboons (*Papio Ursinus*) used playback experiments to show that baboons track relationships among conspecifics (Cheney and Seyfarth 2008). In other words, baboons understand who is dominate or subordinate to whom, who is related to whom, and which individuals are likely to help one another if a conflict arises. Baboons demonstrated such knowledge in Seyfarth and Cheney's experiments through clear reactions of surprise when hearing various types of vocalisation that break expectations related to the current standing of relationships between individuals in the group (Cheney and Seyfarth 2008). To take one telling example, baboons react more strongly – they show more signs of surprise and even distress – to rank reversals *between* kin groups than *within* kin groups. This makes sense given that rank reversals between kin groups have more potential to disrupt the social organisation and dominance hierarchy in a baboon troop.

In addition to the concept of *bodily agential self-awareness*, David DeGrazia also introduced the concept of *social self-awareness* into philosophical discussion of animal minds. He defines it as 'awareness of oneself as part of a social unit with differing expectations attaching to different positions' (DeGrazia 2009). DeGrazia argues that many group-living animals including baboons, great apes, dolphins, elephants, and wolves and domestic dogs are socially self-aware on the basis that they demonstrate *social understanding* (DeGrazia 2019). In other words, DeGrazia concludes from studies like Cheney and Seyfarth's that baboons, for example, demonstrate social self-awareness because they understand general social dynamics and relationships in their group. This is an important point, but just as with bodily agential self-awareness, it is also crucial here to include the criterion of flexible social *agency*. In other words, social self-awareness does not only involve an understanding of the group structure, but also an ability to use such knowledge to act successfully in the group.

An example from hamadryas baboons (*Papio hamadryas*) that has been cited regularly in philosophical literature is illustrative here (Bermúdez 2007; DeGrazia 2009). Kummer (1982) originally reported in a discussion of tactical deception (an example initially regarded as evidence of 'theory of mind') that an adult female spent 20 min gradually shifting her seating position over a distance of two meters to a place behind a rock where she began to groom a sub-adult male follower of the group (not one ordinarily belonging to the group) – an interaction not tolerated by the adult male (Kummer 1982). The adult male could see her, but not that she was grooming another male. Whiten and Byrne argue that the female *understood* that the harem male leader could not see that she was grooming another male (Whiten and Byrne 1988).

Such behavioural inhibition in the presence of a dominant individual seems to show that the female baboon had expectations of how the social world *works*, and

that the accuracy of these expectations meant either 'getting away' with what she wanted to do, or being the recipient of an aggressive attack. She understood what would happen if the resident adult male were to see her grooming the sub-adult male. These forms of social expectation and self-understanding have also been studied experimentally in captive rhesus macaques (Drea and Wallen 1999). In one experiment, macaques were taught to solve a simple colour-association task to learn the location of boxes baited with peanuts. The monkeys were tested in two social situations: as a complete social group and as a 'split' group, where half of the troop – either the dominant or subordinate matrilines – were removed from the testing area. In both conditions, the dominant individuals retrieved the food from the baited boxes. In contrast, the subordinate individuals retrieved the food correctly only when in the split condition. Because the subordinates performed well in the split condition, their performance in the combined condition suggests that they inhibited expressing their knowledge in the presence of dominant individuals. Again, this behaviour suggests that they had some expectation about what would have happened had they not inhibited their knowledge in this way. Just as understanding that the wind is causing the bag to fly in the air, this form of social understanding is a form of causal understanding of how social interaction works. Yet, like conditioned taste avoidance, it involves expectations about causal patterns involving an animal's own actions.

These examples suggest that we should amend the concept of social *self*-understanding in a way parallel to our amendment of the concept of understanding in general. Baboons might understand how the social world works, but we cannot yet speak of social *self*-understanding until there is evidence that they can flexibly use information about the relationships between others, and their own relationships to others, to perform successful actions. Fortunately, there is also abundant evidence that many animals can do this, but we will continue with examples from baboons to fill in our picture of this highly social and socially intelligent species. In an example drawn from ongoing research at the Tsaobis Baboon Project in Namibia, experiments are performed in which individuals are given the opportunity to explore small novel foods (Carter et al. 2014). Not all individuals are as willing as others to explore let alone consume the novel foods, but some individuals learn quickly the value of the foods. One particular example is demonstrative (Carter, personal correspondence). An adult female, 'Yaoundé', was presented with and quickly ate a slice of apple dyed red. Later, her sub-adult son, 'Okavango', was given the same stimulus. Although he was interested in the apple, picking it up and exploring it, he did not eat it. Yaoundé saw Okavango with the apple and approached him, at which point he turned and walked away from Yaoundé, preventing her from acquiring the food. Yaoundé approached Okavango to groom him, which he allowed. After several minutes of intense grooming, Okavango relaxed to the point that he dropped the apple, at which point Yaoundé snatched up the novel food, ate it and walked away. The observer had the impression that the grooming was a ruse to acquire the apple piece.

Examples like this one show that non-human animals can understand where they 'stand' in relation to other individuals and in relation to the general social dynamics

of the group. They can use this understanding to perform goal-directed, successful behaviours that take advantage of such information. One knows what one can do, and cannot do, by understanding something about the general causal patterns that make up baboon social life. From a metaphysical perspective, the social understanding and social self-understanding of animals are interesting because they do not involve knowledge of a world wholly external to the animal agents themselves. In other words, it is, in part, baboon social understanding and self-understanding that makes the social life of the group unfold in the way it does. There is a feedback loop between expectations of social patterns and the existence of those patterns in way that has been insightfully analysed in philosophical work on social ontology (Haslanger 2013; Hacking 1995).

6 Awareness and Understanding

Inquiry into the nature of animal understanding and self-understanding affords an opportunity to inquire into the relationship between understanding and awareness in general. The relationship between understanding and awareness is a central issue for familiar thought experiments such as John Searle's 'Chinese room' or the Turing test, and other examples involving automata and thermostats. A familiar line of thought is that while thermostats, computers (up to now), or translation rooms might be able to track changes in environmental 'inputs', and deliver the appropriate 'output' in response to changes in 'inputs', they are not aware of their own states, computations, the inputs and outputs themselves, or the causal relationship between inputs and outputs. In Searle's translation box, the idea is that neither the box as a whole, nor any of its parts, understands the meaning of the words it is correctly translating. While this example involves linguistic understanding, the others raise the question of whether systems that respond appropriately to stimuli – in some cases flexibly – can be said to understand the world or themselves if they are not aware of the world or themselves.

Intuitions about the general relationship between understanding and awareness are involved the way we interpret non-human minds. Forms of life can be more like unaware thermostats, but the way they track of features of the world may also approach the awareness and expectation of causal patterns and dependencies required for understanding. It seems possible for a creature to be aware in some way – to have some form of subjective experience or 'what it is like' to be it – and yet not to understand the world. In other words, if primitive subjective 'experience' has the structure of 'white noise' as Peter Godfrey-Smith suggests, it probably does not track causal relationships in a way necessary to develop expectations about such relationships (Godfrey-Smith 2017). Like Cabanac et al., referenced earlier, Godfrey-Smith looks to evolution of sensory-motor feedback loops to find the origins of consciousness. As organisms begin to track not only what they are doing, but how their experience of the world changes as a result of what they do, their experience moves from undifferentiated 'white noise' to more structure and

integration. Godfrey-Smith favours the view that consciousness arose with the cognitive integration of information from both internal and external sources, and that such integration occurred fairly early on in the evolution of animal life.

Such an evolutionary view supports a certain conception of the relationship between awareness and understanding. While we cannot infer the ability to understand from the presence of awareness or some form of subjective experience, the reverse seems to be a justified inference. If an animal indeed understands the world, or itself, then it must be aware of the world, or itself. It seems right, then, to design experiments that test various forms of understanding – such as the experiments on causation with crows, on 'self-agency' in macaques, and on social knowledge in baboons cited above – and to conclude from evidence that animals understand causal patterns in the world and are able to act flexibly on the basis of such understanding that they are aware of the world in various ways. The inference from understanding to awareness, or self-understanding to self-awareness, (but not the reverse inference) is supported by the conception of understanding as a set of expectations involving causal patterns between external events, or between one's own actions and their effects (both internal and external).

Of course, such an inference is not widely accepted by empirical researchers investigating awareness or self-awareness. Baboons, for example, are often thought to lack self-awareness because they have not to date 'passed' the classic mirror test (Carter, personal communication). That is, they do not appear to be able to recognise their reflection in a mirror *as* an image of their own body. Since the 'mirror test' has been taken as the gold-standard test for self-awareness in animals, many have assumed that species lack self-awareness if they do not perform self-directed behaviours in response to their reflection. Animals from diverse taxa appear to 'pass' the test by performing such self-directed behaviours, such as inspecting their bodies or touching marks placed on their bodies (Gallup 1970). Nonetheless, mirror tests have had controversial results, and the findings are often not as clear-cut as they are sometimes presented. For example, chimpanzees are often claimed to 'pass' the mirror test to the exclusion of all other non-human species (Gallup and Anderson 2018); however, while some chimpanzees undeniably 'pass', others do not (Swartz and Evans 1991). This indicates that mirror self-recognition is not a universal trait in chimpanzees, and may even be learned. Further questions arise when we consider that wild chimpanzees do not respond to a mirror in a similar manner to captive individuals (Anderson et al. 2017). At the other end of the scale, small cleaner wrasse have passed the test by rubbing their body against a rock only when it was injected with a dye and placed in front of a mirror, but this has been denied as evidence of self-awareness due to incredulity that fish possess a physiological architecture complex enough to support self-awareness (de Waal 2019; Kohda et al. 2019).

If we bring philosophical reflection on the nature of understanding and scientific work on awareness and self-awareness together, as I suggest we do, then a wider range of experimental evidence might support inferences regarding awareness and self-awareness. For example, experiments on understanding of 'self-agency' in rhesus macaques ought to have the same status as the mirror test in discussions of self-awareness, and studies of social self-understanding involved in behaviours such

as deception or third-party reconciliation in baboons ought to justify inferences of social self-awareness. The inference from understanding to awareness is sound, and this is because understanding requires not only that animals track causal patterns in the world and in the traffic between the world and their own actions, but also that they act successfully and adaptively on the basis of such understanding. Recognition of various forms of understanding across a range of species can help us appreciate a greater variety of evidence that might be available for awareness and self-awareness.

7 Conclusion

This article has defended a notion of self-understanding as an understanding of causal patterns involving the traffic between one's sensations, one's actions, and the effects of one's actions on the outside world. If understanding involves the ability to track causal patterns in the world, and this has been shown in pre-verbal infants and a variety of bird and mammal species, then understanding does not require propositional knowledge, mastery of the concept of causation, or linguistic comprehension. I have argued that self-understanding does not require these capacities either. A conception of self-understanding that does not involve these capacities allows for the possibility that non-human animals can understand themselves. I do not claim to have gone so far as to have proven that self-understanding is present in the examples given here, but merely that they are good candidates for satisfying the conception of self-understanding given here. More empirical work would be required to demonstrate conclusively that animals accurately track causal patterns involving their internal sensations, actions, and the effects of these actions.

In closing, let me gesture briefly to some broader implications of the concepts of understanding and self-understanding given here. One is that animals and pre-verbal children make sense of the world in a 'first-order' manner before they are able to reflect on how they make sense of the world in a 'second-order' manner. Animals may, then, develop expectations about causal patterns without, apparently, the ability to reflect on what or how they understand. The forms of understanding and self-understanding that may exist in non-human animals have likely been shaped through an evolutionary history of causal traffic between subjective responsiveness and what exists in the outside world. Of course, pre-linguistic, 'everyday' forms of understanding and self-understanding are not 'worldviews' in the sense of ideas about the fundamental nature of things, or of what is good and just, nor are they fine-grained, full causal pictures like what we expect in the natural sciences. But, these more lofty and rigorous forms of understanding seem to presuppose the 'everyday' capacities to discern with how things work in the world that this article has aimed to bring into view.

Acknowledgements and Funding I would like to thank Dr. Alecia Carter (UCL) for discussion of these issues and the anonymous reviewers for helpful comments on earlier drafts. I would like to

thank the Templeton World Charity Foundation for funding that allowed for the time to complete this article (TWCF0502).

References

Anderson J, Hubert-Brierre X, McGrew W (2017) Reflections in the rainforest: full-length mirrors facilitate behavioral observations of unhabituated, wild chimpanzees. Primates 58:51–61. https://doi.org/10.1007/s10329-016-0574-7

Barron A, Klein C (2016) What insects can tell us about the origins of consciousness. Proc Natl Acad Sci U S A 113:4900–4908. https://doi.org/10.1073/pnas.1520084113

Baumberger C, Beisbart C, Brun G (2017) What is understanding? An overview of recent debates in epistemology and philosophy of science. In: Grimm S, Baumberger C, Ammon S (eds) Explaining understanding: new perspectives from epistemology and philosophy of science. Routledge, New York, pp 1–34

Bermudez J (1998) The paradox of self-consciousness. MIT Press, Cambridge

Bermúdez J (2007) Thinking without words. Oxford University Press, Oxford

Blaisdell A, Sawa K, Leising K, Waldmann M (2006) Causal reasoning in rats. Science 311:1020–1022. https://doi.org/10.1126/science.1121872

Bures J, Bermudez-Rattoni F, Takashi Y (2002) Conditioned taste aversion: memory of a special kind. Oxford University Press, Oxford

Cabanac M, Cabanac AJ, Parent A (2009) The emergence of consciousness in phylogeny. Behav Brain Res 198:267–272. https://doi.org/10.1016/j.bbr.2008.11.028

Carter AJ, Marshall HH, Heinsohn R, Cowlishaw G (2014) Personality predicts the propensity for social learning in a wild primate. PeerJ 2:e283. https://doi.org/10.7717/peerj.283

Cheney D, Seyfarth R (2008) Baboon metaphysics: the evolution of a social mind. University of Chicago Press, Chicago

Couchman J (2012) Self-agency in rhesus monkeys. Biol Lett 8:39–41. https://doi.org/10.1098/rsbl.2011.0536

Couchman J (2015) Humans and monkeys distinguish between self-generated, opposing, and random actions. Anim Cogn 18:231–238. https://doi.org/10.1007/s10071-014-0792-6

Couchman J, Coutinho M, Beran M, Smith JD (2009) Metacognition is prior. Behav Brain Sci 32:142. https://doi.org/10.1017/S0140525X09000594

DeGrazia D (2009) Self-awareness in animals. In: Lurz R (ed) The philosophy of animal minds. Cambridge University Press, Cambridge, pp 201–217

DeGrazia D (2019) Animal self-awareness: types, distribution, and ethical significance. In: Fischer B (ed) The Routledge handbook of animal ethics. Routledge, New York, pp 71–82

Drea CM, Wallen K (1999) Low-status monkeys "play dumb" when learning in mixed social groups. Proc Natl Acad Sci U S A 96:12965–12969. https://doi.org/10.1073/pnas.96.22.12965

Eriksson N, Wu S, Do C, Kiefer A, Tung J, Mountain J, Hinds D, Francke U (2012) A genetic variant near olfactory receptor genes influences cilantro preference. Flavour 1:22. https://doi.org/10.1186/2044-7248-1-22

Gallup G (1970) Chimpanzees: self-recognition. Science 167:86–87. https://doi.org/10.1126/science.167.3914.86

Gallup G, Anderson JR (2018) The "olfactory mirror" and other recent attempts to demonstrate self-recognition in non-primate species. Behav Process 148:16–19. https://doi.org/10.1016/j.beproc.2017.12.010

Godfrey-Smith P (2017) Other minds: the octopus and the evolution of intelligent life. William Collins, London

Gopnik A, Sobel D, Schulz L, Glymour C (2001) Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation. Dev Psychol 37:620–629

Grimm S (2016) Understanding and transparency. In: Baumberger C, Grimm S, Ammon S (eds) Explaining understanding: new perspectives from epistemology and philosophy of science. Routledge, New York, pp 212–229

Hacking I (1995) The looping effects of human kinds. In: Sperber D (ed) Causal cognition: A multidisciplinary debate. Clarendon Press, New York, pp 351–394

Haslanger S (2013) Resisting reality: social construction and social critique. Soc Theor Prac 40(1): 145–152

Hoffman M, Beran M, Washburn D (2018) Rhesus monkeys (Macaca mulatta) remember agency information from past events and integrate this knowledge with spatial and temporal features in working memory. Anim Cogn 21:137–153. https://doi.org/10.1007/s10071-017-1147-x

Jelbert S, Miller R, Schiestl M, Boeckle M, Cheke L, Gray R, Taylor A, Clayton N (2019) New Caledonian crows infer the weight of objects from observing their movements in a breeze. Proc Royal Soc B: Biol Sci 286:20182332. https://doi.org/10.1098/rspb.2018.2332

Kohda M, Hotta T, Takeyama T, Awata S, Tanaka H, Asai J, Jordan A (2019) If a fish can pass the mark test, what are the implications for consciousness and self-awareness testing in animals? PLoS Biol 17:e3000021. https://doi.org/10.1371/journal.pbio.3000021

Kummer H (1982) Social knowledge in free-ranging primates. In: Griffin D (ed) Animal mind - human mind. Springer, Berlin

Kushnir T, Gopnik A (2005) Young children infer causal strength from probabilities and interventions. Psychol Sci 16:678–683. https://doi.org/10.1111/j.1467-9280.2005.01595.x

Parker L, Rana S, Limebeer C (2008) Conditioned nausea in rats: assessment by conditioned disgust reactions, rather than conditioned taste avoidance. Can J Exp Psychol 62:198–209. https://doi.org/10.1037/a0012531

Reilly S, Schachtman T (2008) Conditioned taste aversion: neural and Behavioral processes. Oxford University Press, Oxford

Saxe R, Tenenbaum J, Carey S (2005) Secret agents:inferences about hidden causes by 10- and 12-month-old infants. Psychol Sci 16:995–1001. https://doi.org/10.1111/j.1467-9280.2005.01649.x

Saxe R, Tzelnic T, Carey S (2007) Knowing who dunnit: infants identify the causal agent in an unseen causal interaction. Dev Psychol 43(1):149–158. https://doi.org/10.1037/0012-1649.43.1.149

Swartz K, Evans S (1991) Not all chimpanzees (Pan troglodytes) show self-recognition. Primates 32:483–496. https://doi.org/10.1007/BF02381939

Taylor A, Miller R, Gray R (2012) New Caledonian crows reason about hidden causal agents. Proc Natl Acad Sci U S A 109:16389–16391. https://doi.org/10.1073/pnas.1208724109

du Toit JT, Provenza FD, Nastis A (1991) Conditioned taste aversions: how sick must a ruminant get before it learns about toxicity in foods? Appl Anim Behav Sci 30:35–46. https://doi.org/10.1016/0168-1591(91)90083-A

Völter C, Sentís I, Call J (2016) Great apes and children infer causal relations from patterns of variation and covariation. Cognition 155:30–43. https://doi.org/10.1016/j.cognition.2016.06.009

de Waal F (2019) Fish, mirrors, and a gradualist perspective on self-awareness. PLoS Biol 17: e3000112. https://doi.org/10.1371/journal.pbio.3000112

Whiten A, Byrne RW (1988) Tactical deception in primates. Behav Brain Sci 11:233–244. https://doi.org/10.1017/S0140525X00049682

Zagzebski L (2001) Recovering understanding. In: Steup M (ed) Knowledge, truth, and duty: essays on epistemic justification, responsibility, and virtue. Oxford University Press, Oxford, pp 235–252