Review of Philosophy and Psychology (2023) 14:1309–1328
https://doi.org/10.1007/s13164-022-00643-z

Attributing Psychological Predicates to Non-human Animals: Literalism and its Limits

Andrés Crelier[1]

Accepted: 25 April 2022 / Published online: 20 May 2022

Abstract
In this essay, I deal with the problem of the attribution of psychological predicates to non-human animals. The first section illustrates three research topics where it has become scientifically legitimate to explain the conduct of non-human animals by means of the attribution of psychological predicates (mind-reading in apes, episodic memory in rats, and the feeling of regret in rats and mice). The second section discusses several philosophical objections to the legitimacy of such attributions provided by central thinkers from the last decades (like Malcolm, Stich, Davidson, Dummett, McDowell, and Brandom). I try to show that these objections —which are related among other questions to the holism of the mental, the indeterminacy of the attributions, and the strangeness of animal concepts— can be alleviated. In the third section, I propose to adopt a literalist view of the attributions in the sense articulated by Figdor (2018). At the same time, I argue that one must draw limits to the conceptual change forwarded by her literalist view, taking into account holistic considerations and the fact that the psychological concepts must retain their core notes.

Keywords Psychological predicates · Non-human animals · Intentional attitudes · Literalism

1 Introduction

In the past decades, a growing amount of evidence about animal behavior and cognition is being interpreted using a mentalist vocabulary. More generally, in research areas such as cognitive ethology, animal psychology, and comparative animal cog-

✉ Andrés Crelier
pcrelier@mdp.edu.ar; andrescrelier@gmail.com

[1] National Scientific and Technical Research Council, University of Mar del Plata, Mar del Plata, Argentina

nition, it has become scientifically legitimate to explain the conduct of non-human animals (from now on 'animals') by means of the attribution of psychological predicates (Andrews 2009; Figdor 2018). In opposition to this view, there is an important philosophical tradition, featuring central thinkers from the last decades, which favors skepticism about the legitimacy of these kinds of attributions Malcolm 1972; Stich 1979; Williams 1973; Davidson 1982; Dummett 1993; McDowell 1996, 2009; Brandom 1994). As a result from different arguments, these authors see the attributions of thoughts to animals as a *façon de parler* or an instrumentalist interpretation of the conductual evidence (Davidson); the attributed states are considered *proto*-thoughts (Dummett) or belief-like states (Stich); or they are assigned a derivative status dependent on the human linguistic communities (Brandom).[1] Although from very different perspectives, these authors promote a reinterpretation of a growing number of mentalist hypotheses regarding animal cognition, which should thus be seen as systematically untrue, about a different topic or have circumscribed legitimacy inside their respective fields of research.

There are at least two different ways to deal with this situation. The first one is to insist that there is no common ground between the two perspectives —i. e. between the research areas where cognition is interpreted using a mentalist vocabulary and the important philosophical line of arguments that I mentioned. Since they would be talking about different topics, the price to pay is to give up a potentially fruitful dialogue between them.[2]

The second path, which I propose to follow, avows that there is a common ground between both perspectives and recognizes that there is a corresponding tension. This leads to an initial dissatisfaction since if the particular arguments coming from the above-mentioned tradition (which I will discuss in section 2) are right, the mentalist attributions in a growing number of research programs (as illustrated in section 1) would be illegitimate. However, recognizing this initial tension does not imply that it cannot be reduced.

In this paper, I will try to show how it can be done, i.e., how to respond and mitigate the force of the particular arguments of this philosophical tradition. Accordingly, my general hypothesis reads as follows: it is possible to assume a common ground between both fields of inquiry and to alleviate the philosophical objections regarding the attributions of psychological predicates to (some) non-human animals, and the resulting viewpoint can be expressed as a moderate "literalism".

The structure of the discussion is as follows. I review recent cases of attribution of intentional attitudes in three research programs (I); I then examine and alleviate

[1] This general viewpoint can also be couched in ontological terms as 'anti-realism or 'non-realism' on animal thoughts. However, some of these authors may also be seen as encouraging an anti-realist viewpoint regarding the attribution of thoughts to humans. Therefore, the ontological discussion —which deserves its own space— may obscure the distinction I make between admitting and denying the legitimacy of the attributions in the realm of animals.

[2] As a reviewer of this paper suggested, this can be considered an unsatisfying situation but not necessarily an unsatisfying analysis, which may be accomplished for example by distinguishing cognition and mind (cf. Keijzer 2021). In this paper, I do not aim at a direct analysis of the situation, but I intend to show that assuming the existence of a common ground helps to understand, with philosophical tools, the attributions that are made in different research fields. Ultimately, this strategy also helps to view the limits of this common ground, as I intend to show in section 3.2.

the force of some philosophical objections to the attribution of attitudes (II); I finally defend a 'literalist' (Figdor 2018) extension of the psychological vocabulary outside the human realm and point out some of its limits (III).

Two further remarks are in place in this introduction. Firstly, in order to develop my view, I will line up both perspectives in the following way. On the one side, the research topics that I take into account in cognitive ethology and animal psychology are primarily focused on the attribution of cognitive *capacities*, leaving the thought's contents on a second place (as it is illustrated in section 1); and, on the other side, the philosophical arguments draw their conclusions paying special attention to the *contents* of thoughts, not to the attitudes themselves. To align both perspectives, I will apply the philosophical arguments to the psychological predicates, which I will refer to as intentional attitudes.[3]

Secondly, since many of the arguments that I will put forward are not new, I would like to indicate the aimed contribution of this paper. As mentioned above, whereas the philosophical debate that I take into account is mostly concerned with the thought's contents, I will focus instead on the intentional *attitudes*. At the same time, the review of the scientific research will be applied to a debate that usually employs everyday examples; by using scientific cases, I will put the corresponding disciplines in dialogue, which may help overcome some skeptical perspectives on animal thoughts. Finally, the overall conclusions concern the scope of 'literalism,' which is a recent proposal that expresses to a great extent the viewpoint that I favor.

[3] The usual idiom used in philosophy to describe a mental state takes it to be a 'propositional attitude,' that is, a psychological stance towards a 'content' that can be expressed by means of a proposition composed of concepts. Since not all attitudes have propositions as contents, I will use the expression 'intentional attitude' (Glock 2010, p. 13). A reviewer of this article pointed out that the philosophical arguments in the debate are usually arguments against attitudes by making a point about content. I agree with the remark, but I would formulate it differently. I would affirm that the conclusions of the arguments that concern the thought's contents are usually extended to include the attitudes themselves. For example, the holistic constraints on the attribution of thoughts (considered as contents) are taken to include the attribution of attitudes (Davidson 1982). However, I believe that this traditional approach does not take into account the intentional attitudes in a more direct way, and that there is room to inquire if the arguments primary aimed at contents have the same implications when applied to the attitudes themselves. This is a central motivation of this paper (specially dealt with in section 2). My approach is thus compatible with revisionary perspectives such as the one recently put forward by Newen and Starzak 2020a, b> (I thank the reviewer for the reference). More precisely, the orthodox view assumes that the consideration of contents has a priority in the debate, and one main reason is that without the possibility of attributing specific contents, it has no sense to attribute attitudes. Contrary to this perspective, Newen and Starzak plausibly claim that there are behavioral and cognitive criteria that permit the attribution of attitudes independently of the attribution of specific thoughts. In agreement with them, I believe that even if attitudes require contents (as I argue in Section 2.2), they have their own conditions of attribution. (Other authors that favor the idea that behavioral and cognitive sophistication are sufficient to attribute beliefs-like representations that avoid the requisite of determinacy are Stich 1979, Godfrey-Smith 1999; Sterelny 2003 and Millikan 2005, see the discussion in Glock 2020).

2 Samples from Ethology: Mind-reading, Remembering and Regretting

2.1 Chimpanzees Understand what Others Believe

The generic label 'mind-reading' alludes to the capacity of understanding, predicting, and acting according to (an understanding of) the mental states of others. The starting point of this line of research in animals dates back to Premack and Woodruff (1978), who asked if nonhuman primates understand false beliefs. A recent experiment with the participation of two leading researchers, Call and Tomasello, defends the view that apes distinguish between true and false beliefs in the context of social interactions (Buttelmann et al. 2017). The researchers adopted a test designed for human infants between 1 and 2 years old, where the anticipatory look and the helping conduct show an understanding of the goals and (false or true) beliefs of an actor.

As in the classic false belief tests, the infants initially watch an experimenter placing a toy into a box. In the false belief condition, the experimenter leaves the room and an assistant moves the toy to a different box. The experimenter then returns and struggles to open the (now empty) box where he had originally put the toy. The infants, who are given the possibility to intervene, help them to accomplish what they consider to be her goal: to retrieve the toy (which has been moved to another box). In the true belief condition, the experimenter *watches* that the toy is moved to the new location. When she then tries to open the wrong box (the empty one), the infants do not try to retrieve the toy (as in the false belief situation) but help them to accomplish what they now consider to be her goal: to open the box. This would demonstrate that infants can distinguish between different goals (to open the box or to get the toy), and the true or false beliefs associated with these goals.

The replication of this experiment involved 34 great apes (chimpanzees, bonobos, and orangutans) and showed a similar pattern of conduct since in the false belief situation the apes seem to understand the false belief of the experimenter. They accordingly tugged more often (than in the true belief situation) at the box where the toy has been moved (unknowingly to the experimenter), that is, not at the box where the experimenter was actually looking for the toy. Since the only difference between the two conditions is the fact that the experimenter *believed* different things regarding the place of the toy, this is already evidence that the apes understood her goals and beliefs.

However, the researchers consider the alternative hypothesis that they may be understanding not a false belief but the experimenters´ state of ignorance. Therefore, they designed a follow-up test with two conditions to discern between understanding a false belief and ignorance. Whereas in one of them the experimenter mistakenly thought that the toy was in a certain box, in the other one she had no clue of the possible location. The results showed a different behavior on the part of the apes since they tended to look for the toy on the false belief condition and behaved randomly on the ignorance one (which thus functioned as a control condition).

From both studies, the researchers concluded that apes can understand the goals of others, as it is shown by their helping behavior. Since these goals are partly understood in relation to beliefs, they 'may have a basic understanding of others´ false

beliefs' (Buttelmann et al. 2017, p. 1, 13). As it is known, this kind of optimistic conclusion is challenged by a deflationary explanation, according to which these abilities amount to understanding the behavior of others, not their psychological states. Apes and other species would merely be understanding regularities (or rules) of conduct in order to get what they want (Lurz 2018).

I do not need to go into the respective discussion on mind vs. behavior-reading, since the attribution of the ability to understand the conduct of others is already a mentalistic ascription. In the complex scenarios that I described, understanding a behavior amounts to understanding its goal, which –as noted above– implies understanding some of the beliefs regarding these goals. In any case, as Heyes stresses, the proposers of behavior-reading (against mind-reading) hypotheses make use of categories such as person, object, and location, so that the subjects under investigation are still taken as intentional systems that interpret others.

There are also-lower level explanations like those formulated in terms of 'submentalizing' abilities (which according to Heyes should have more place in the research program on animals; Heyes 2014; 2015), or tracking abilities couched in quasi-mentalist terms (Butterfill and Apperly 2013). My view is that these hypotheses are not currently in a position to debunk the higher-level hypothesis. However, one may be cautious and examine other research fields, as I do in the following two sub-sections.

2.2 Rats Remember an Order of Events

All animals have memory in a broad sense of the term, but the way to explain this fact can avoid the ascription of the intentional attitude of remembering, appealing, for instance, to behaviorist explanations in terms of 'instrumental conditioning' or 'habituation'. By contrast, in recent decades a research program inspired by human psychology focused on 'episodic memory' and made a central use of the intentional idiom. In general terms, an animal possesses episodic memory if it can remember an earlier event or episode of its own life and act upon this memory.[4] Then again, since the tasks manifesting this ability could also be solved by assessing the familiarity of an item or event, episodic memory is confronted with alternative explanations. According to the latter, an animal can solve memory problems by comparing the strengths of different memory traces, which are a function of the elapsed time (Crystal 2018, p. 106).

In a recent study, a group of researchers defended the view that rats possess the ability to remember episodic events, more specifically, that 'rats rely on episodic memory replay to remember the order of events rather than relying on non-episodic memories.' (Panoz-Brown et al. 2018, p. 1628). That is, instead of searching for evidence of the capacity to remember multiple episodes, they looked for the ability to 'replay' a unique order of events. The setup of the experiment took advantage of the fact that rats can discriminate and remember a large number of different odors and trained them to identify the second and the fourth to last items in a series of 5

[4] The definition is broad since I focus on a particular experiment. There is an ongoing debate on how to understand this kind of memory and how to prove this capacity in non-humans (cf. Hoerl & McCormack 2018; Crystal and Suddendorf 2019).

to 12 odors. They tested this ability in a series of trials with lists whose length was variable and non-predictable. After the rats were exposed to a sequence, they had the opportunity to choose between two odors from the list, but they were only rewarded for choosing the second (or the fourth) to the last.

The results showed a high accuracy of correct choices, which served as initial confirmation of the hypothesis that rats replay episodic memories. However, the researchers were aware of the alternative explanation of the same data in terms of non-episodic memory. The trace strength of each item in a list decay over time, which may serve as a cue to order the items in the sequence: 'an animal may obtain high accuracy by selecting the item that matches the typical memory trace strength without replaying episodic memories' (p. 1629). To discard this hypothesis, the experimenters produced atypical intervals among the last items of the list and even placed foil odors in the second and fourth to last position according to a memory strength ordering. If rats relied on the internal cues provided by the memory traces, this would have led them to make incorrect choices, that is, to behave below chance. As it happened, they continued to behave above chance.

2.3 Rats and Mice Regret a Bad Choice

For Darwin, 'A moral being is one who is capable of comparing his past and future actions or motives, and of approving or disapproving of them. We have no reason to suppose that any of the lower animals have this capacity' (Darwin 1871/1982, p.136). Although this differentialist approach to morality has been challenged (De Waal 2000 on reconciliation practices in primates), neither the moral feeling of remorse (a deep regret or guilt for a wrong committed) nor the kindred notion of repentance (loaded with religious resonance) seemed accessible objects for ethological research. However, the research on animal psychology has been recently exploring a related notion, regret, in mammals such as rats and mice.

In the experiments that I briefly review, the notion of regret is taken from a human research paradigm, where it is defined as 'the subjective experience of recognizing that one has made a mistake and that a better alternative could have been selected.' (Sweis et al. 2018, p. 1). These studies tried to identify this mental state in rats by means of regret inducing situations in the laboratory. The so-called 'Restaurant row' task, as described by Steiner and Redish (2014) consists in a series of successive zones where the rats have to decide, during an hour a day of foraging within the task, whether to stay or to skip an indicated delay for a particular food. The rats have stable preferences for different foods, which have been independently ascertained, and are given an acoustic cue that indicates the delay they have to wait in each zone to get the reward. They know which is the kind of food they are waiting for, and they have a time threshold that they are prepared to wait to get each kind of reward.

Once the animals are familiar with the task, the researchers randomly introduce situations that are apt to show evidence of regret. The rats are firstly given a low-cost delay, where they should wait for less than their own threshold for waiting for a certain (high value) reward. Sometimes, they skip these low-cost offers, and after that, they are presented with a regret inducing situation. They are given a high-cost delay, that is, they are informed that they should wait for the reward (this time a low-value

one) beyond the point at which their own threshold for waiting would have been exceeded. The rat becomes aware that it has made an economic mistake; if it had waited for the previous reward, it would have obtained a more valuable result.

The evidence for regret is behavioral since after the regret inducing situations the rats paused and looked back to the previous reward site (among other relevant conducts); and it is also neurophysiological since the researchers found neural activity patterns (in the orbitofrontal cortex and the ventral striatum), recorded during the 3 s after the studied reward delivery, that was more consistent with a representation of the previous reward than the present one, and also consistent with human reports of 'might have been representations'.

The situation induces regret because the bad outcome is the product of a decision of the agent itself. It differs from situations that induce disappointment, where the agent does not make a mistake. To control for the latter emotional state, the researchers designed two conditions that should induce disappointment but not regret. In the first one, the sequence was similar to the condition that induced regret but the rat *waited* for the first low-cost reward, i.e., it didn't skip it after knowing about the kind of reward and amount of delay to get it. Since the next offer implied a high-cost, it was apt to induce disappointment, but not regret (since taking the first offer was a good choice). In the second one, the rat encountered two high-cost offers, and it skipped the first one. This induced disappointment in the second offer because both offers were high-cost, but not regret (since skipping the first one was also a good choice in this case).[5] In the two control situations, the rats did not pause and look backward, in fact, they tended to look toward the next zone. An analysis of the neuronal activity after the relevant choices also showed a difference from the activity during regret.

Sweis et al. (2018) hold that this experiment demonstrates that rats are responsive to the immediate effects of regret but not to its possible long-term consequences. They accordingly designed an experiment to show that mice can learn from regret situations and change the decision-making strategies to prevent future regret. In this variant of the Restaurant Row test, the result was that, according to the researchers, the mice avoided the scenarios that may lead to regret by learning to plan ahead. That is, the evidence amounts to a behavior that is coherent over a longer period of time and encompasses different actions, elections, and strategies.

Sweis and colleagues stated some of the results using a vocabulary that includes the notions of changing one mind, evaluating the outcome of an action, learning, adopting decision strategies, deliberating, planning, and avoiding entering certain scenarios that would provoke an undesirable mental state (p. 11).

3 Mitigating Philosophical Objections

Even if the mentalist attributions play a central role in some research programs, one can press the point in a skeptical tone: can we legitimately say that apes read minds,

[5] Note that in each control condition the correct decision is not relevant to disappointment itself —that is, it does not cause this negative emotion- but it is relevant to contrast the sequence with the regret condition.

rats remember the order of presentation of a list of odors and mice regret? My central aim in what follows is not to discuss whether the attributions that I reviewed are the best explanations of animal conduct. I only state that according to the current research, where they play a central explanatory role, they possibly *are* the best explanations –which may be turn out to be false any scientific hypothesis. My concern rather relates to the tension between these attributions and diverse philosophical considerations that put into question their legitimacy.

3.1 Holistic Considerations

It is plausible to claim that to attribute an attitude presupposes the attribution of a list of other (actual or possible) attitudes to the same creature and that each one of them must be manifestable in conduct (Searle 1994, p. 210). The way one represents the kind of relationships that the members of the list must have among themselves may put restrictions to the attributions generally, and with particular force in the non-human case. Therefore, I will spell out three ways to understand the relationships.

According to the first one, the meaning of an attitude depends on the network in which it is placed. But I think one should be cautious with this general notion. Although it makes sense at the level of content –for instance at the level of a conceptual holism, whose extreme variants have nonetheless been criticized (Glock 2018, pp. 95–97; cf. Newen and Starzak 2020a, b)–, it is not the case that the meaning of an attitude is completely dependent on a network of other attitudes. The mental state of remembering something can be described as a relation between an agent and a represented past event, independently of a description of the act of perceiving something or regretting a choice. In other words, two creatures capable of remembering particular episodes of the past may be very different as regards other mental capacities.

It is more plausible to uncover the conceptual links demanded by a particular attitude, in the sense that the attribution of an attitude usually presupposes the ascription of others (and their possible related contents). In the context of the experiments above: a rat cannot *regret* X if it did not *remember* that p and *believe* that q; a chimp cannot *understand the goal* of X if it did not *believe* that this goal is the result of certain action; a chimp cannot *expect* that the experimenter will open the left box without *remembering* that she (the experimenter) did not watch the toy being moved to the box on the right.

Finally, the relationship between attitudes may also be understood in terms of the opposition between them. But it is usually difficult to find out which are the adequate relationships: whereas remembering can be opposed to forgetting, the experience of regret may be opposed to the feeling of frustration only along certain notes (regret is specially linked to actions done by the agent herself). And even if we find a par of opposite attitudes —say regret vs. pride— it does not follow that an agent must be capable of both of them to be capable of experiencing only one of the par.

The lesson from this section is that each ascription has meaning inside a limited set of attitudes, where presupposition relationships are to be found. This imposes some conceptual demands on our ascriptions (to animals and humans alike); for instance, we cannot attribute to a creature the experience of regret if we do not, at the same time, attribute to it the capacity of episodic memory. However, these holistic con-

siderations are no obstacle to the attributions made above, all of which can be manifested in conduct.

3.2 Attitudes Require Contents

A further condition is that the attribution of an attitude requires the attribution of adequate content. And it may be the case that certain attitudes require contents that cannot be manifested in non-linguistic conduct. Interpreting Wittgenstein, Savigny spells out this restriction in terms of content details, in the sense that, at a certain point, the attribution of content details cannot be justified if the creature does not possess the richness of expressive resources that only a language provides (Savigny 1995, p. 44).

But this is disputable since one can add further details to a content specifying, for example, the time and space of a remembered location, without setting it outside the grasp of a non-linguistic creature: a rat may remember that the fourth to the last odor of the list is this same one. Similarly, the fact that a chimp can understand the false beliefs of others does not imply that the contents of those beliefs must be simple (that needs further empirical research).

It is more fruitful to examine this restriction in relation to the *kinds* of content. It surely lies outside the realm of a non-linguistic creature the comprehension of the English Bill of Rights or hexadecimal arithmetic. But there are abstract notions that are not clearly beyond the reach of non-linguistic creatures. With reference to a time scheme, Savigny imagined a complicated way in which a dog could manifest that his master will beat him the day after tomorrow, which Wittgenstein famously declared outside its capacities (Savigny 1995, 43–44). As with the content details, the research shows that the issue should be discussed in relation to the empirical evidence —for example, on planning for the future in great apes.

Finally, even if the requirement of being capable of opposite attitudes (in order to attribute one of the attitudes) can be disregarded, the situation is different in the case of contents. Given a content that is of the adequate kind, if S believes that *p*, she should be able to believe that *not p*, or, to avoid the problem of negated contents, she should be able to believe that *q*, i.e. a content different from *p* that is in opposition to it. I think that the experiments we have seen put the subjects in situations where they can behaviorally manifest these capacities. For example, a chimp may manifest the belief that the toy is inside the box, and perhaps that it is not the case that the toy is in inside the box; or, to avoid attributing a negated content, that the box is empty or that a banana is inside it.[6]

[6] The attribution of negatable contents is a controversial issue since some authors deny that non-linguistic animals can have thought contents involving negation (Bermúdez 2007; Millikan 2007). For reasons of space, I will only suggest that the problem that this generates for mental holism in the animal case may be avoided in at least three different ways: (i) one can defend the thesis that a mental holism of beliefs does not require negatable contents; (ii) one may pursue an anti-realist strategy, maintaining the legitimacy of such attributions on instrumentalist grounds; (iii) one can defend the possibility of attributing negatable contents (and thus contents involving negation) to non-human animals. My personal view is that the last option is worth pursuing.

3.3 Patterns of Behavior

Is there a general way to justify the attributions of intentional attitudes? To answer this question we must keep in sight the relation between mind and conduct, which leads to a kind of hermeneutical circularity. A legitimate *particular* attribution must be manifested in conduct, but we must *interpret* conduct in order to ascribe a psychological predicate. A way out of the circle is to consider that the attributions must be grounded on a coherent array of behavioral manifestations, which functions as independent criteria to test the single attributions (in the next section, I return to the circularity reproach that still lurks here). I believe that this requirement may be expressed by the metaphor of 'pattern,' which plays a central role in the way Wittgenstein explains the ascription of psychological predicates (Wittgenstein 1967, PU II, i, p. 209; Savigny 1995, pp. 53–56).

As Savigny puts it, we do not infer the presence of a pattern but we see it (Savigny p. 54). Taking some distance from the social conditions highlighted by Savigny, I will employ the notion of a 'pattern of behavior' in a liberal way to stress the following aspects: Firstly, a pattern is not to be inferred from instances of behavior, since a pattern permits one to interpret them. The hypothesis that a chimp understands the false belief of an experimenter cannot be exclusively grounded on the conduct of trying to open a particular box, but it has to be interpreted in a setting where the chimp acted and perceived a sequence of events. Similarly, the evidence of regret in a rat cannot be reduced to its looking back at a certain point in the experiment, but it has to be related to a past action in which the rat did not take a low-cost treat, subsequently finding a higher-cost offer. Note that the patterns of behavior highlighted in the research are very complex without involving language possession.

Secondly, although the patterns are not inferred from the observed conduct, their attribution is confirmed by conspicuous features. The rats look back after they, by hypothesis, have made a bad choice; the chimps approach the food when they, by hypothesis, believe that the dominant is not knowledgeable. These conspicuous elements pertain to the pattern, but their role as signs of its presence is contingent. If rats look back, we can take this fact as a reliable sign of the regret they are supposedly experiencing, but it may be the case that rats do not look back when experiencing this attitude towards a past action, displaying in its place other signs of regret.

Thirdly, the instances of behavior conforming to a pattern must be coherent. Rats not only look back at a crucial moment, but they change their foraging conduct maximizing the food intake and, in the case of mice, avoiding regret situations. Interestingly, this coherence includes non-behavioral elements such as the neurophysiological process previously typified.

Fourthly, the primacy of the pattern for the interpretation of conduct means that patterns are concept-dependent; we can articulate them because we have the appropriate psychological predicates ultimately taken from human contexts. This gives rise to the objection of anthropomorphism since the projection of our concepts would never suffice to objectively grasp the animal mind.

Before addressing this reproach, I want to stress that the Wittgensteinian notion of a pattern (which I construed liberally) helps conceptualize some requirements discussed above. The attribution of a pattern implies the possible ascription of attitudes

and contents that cohere, but it does not force the adoption of extreme holistic perspectives; it depends on understanding psychological human concepts, yet allows for objective evidence that confirms or disconfirms their application outside the human sphere; finally, it suits the way in which ethologists and psychologists develop the hypotheses, design the experiments and debate the conclusions.

3.4 Our own Psychological Concepts

As stated above, the researchers do not approach animal behavior without some previous understanding of how to identify an action or a pattern of behavior, which they have gained from the human case (Hacker 1990, p. 147). Our ascription of attitudes to animals may thus be making an unwarranted projection of our own conceptual scheme.

Although in the attribution of attitudes we do not suppose that the creature possesses the concepts by which we understand these attitudes, we assume that these concepts are *applicable* to identify mental states beyond the human case. Furthermore, the ascription of many attitudes requires the ascription of related contents, some of which are of a conceptual nature; then again, these concepts are taken from our own conceptual repertoire. Not only are we using our notion of regret to interpret the rat conduct, but we are also attributing to it the capacity to conceptually understand different kinds of entities and events.

Some authors concede this point and try to make up for this anthropomorphism, claiming that animals can be in *similar* mental states to the ones we are in (Rowlands 2009, p. 195). As regards attitudes, rats do not remember episodes exactly as we do but they do it very similarly, apes do not understand false beliefs but have a similar cognitive strategy. This is possible, but I would also like to cast doubts on a general recourse to similarity. Firstly, it is unwarranted to affirm right from the start that our concepts (or, to put it more mildly, their relevant notes) do not apply to the creatures studied (or eventually that they do not possess them). All we have is evidence to judge case by case what we have (and do not have) in common with the interpretees, even in the human case. Secondly, similarity implies difference, which alters the truth conditions of the attributions, as Routley argues (Routley 1981, pp. 410–411). In our example, if we affirm that the rats experience 'something like regret,' the conditions under which the attribution is true may include the experience of disappointment, among other feelings that are close enough to regret. But disappointment is a different attitude from regret and has different truth conditions —which were recreated in the control tasks of the experiments to discard this alternative attribution. In the same way, non-episodic memory in rats may render very similar actions to the ones expressing episodic memory, but it also consists of a different cognitive ability, which careful lab experiments may distinguish.

It is thus sensible to accept, with a critical eye that leaves room to similarity, that our own concepts are sometimes applicable, such as they are, outside their original realm (I return to the issue in section 3). But what are our own concepts anyway? As Routley points out, the notion of 'our concepts' is something of a myth, since they already vary enormously among humans (Routley 1981, p. 390). Does 'regret' have a necessary connection with 'remorse' or only an accidental one? Is the notion of

'memory' conceptually linked to 'planning'? There is room to flexibly adapt our concepts when we are interpreting others, which we already do in the human case across different cultures and times, without their ceasing to be our own ones. I believe that the debate on the attribution of thoughts to animals is sometimes oriented by what one may call a 'Procrustean-bed' picture. This view rightly assumes that the animal attitudes and concepts must somehow fit the conceptual scheme of the interpreter, i. e. in our own conceptual scheme, but it wrongly implies that this scheme is rigid, as the picture of a bed suggests.

3.5 Indeterminacy in the Attribution of Attitudes

Applying some objections originally aimed at the attribution of contents (Davidson 1982; Stich 1979), it can be claimed that, in the case of animals, there are always indefinite many alternative attributions among which we can not conclusively decide, i.e., our ascriptions are in principle underdetermined.[7] Note that these kinds of doubts usually arise in a stage of the debate where animals are contrasted with humans, regarding whom this problem does not arise, mainly because of their linguistic capacities.

If we apply it to our examples, the objection takes the following form: how do we determine if rats regret, are frustrated but expectant, or speculate about a future reward? How do we decide if chimps understand the beliefs of others, deliberate on them, have doubts about their goals, or speculate about getting the food later on? Especially as regards contents, such worries provided a skeptical incentive to question our attributions to animals (Stich 1979; Davidson 1982; Williams 1973; Dummett 1993), and they have been challenged on different grounds Routley 1981; Marcus 1990; Glock 2000; Rowlands 2009; among others). However valuable and applicable to the case of attitudes these rebuttals are, I think that the problem cannot be easily solved since the objector can easily adopt the role of a skeptic that insists on the indeterminacy of our favored attributions —and of the attributions that help explain the attributions, etc.

In spite of this, one can take some lessons from the debate to show how indeterminacy can be alleviated, at least in the following two ways: (a) calling into question some of the sources of indeterminacy; (b) showing that the methodology of animal research can reduce indeterminacy.

(a) A possible source of indeterminacy comes from holistic considerations. If we concede that the criteria of identity of an attitude come from the network where it is embedded, it will be impossible to precisely define an attribute that pertains to a network that is largely different from ours, as it is presumably the case regarding the mental life of many non-human species. But I have already argued that the holistic requirements are minimal with respect to attitudes since each of them only presupposes a limited set of particular attitudes (independently of the rest of the network).

[7] In section 2.2, I admitted that attitudes require contents, which is in line with Wittgenstein's and Davidson's holism, and I argued that the attribution of contents is in principle solvable. In this section, and in this paper generally, I deal with the attitudes themselves, which is not Davidson's primary concern (see also footnote 3). At the same time and for reasons of space, I do not deal with the interesting question whether the ascription of attitudes is priori to the ascription of contents or viceversa.

However, we still need to assume a shared area of attitudes that would allow for a sufficient specification. This leads to the problem of the relevant kind of evidence necessary to make plausible ascriptions. Some authors think that only a speaking creature can provide this evidence, in the absence of which we are unable to make legitimate ascriptions of thought (cf. Savigny 1995). Since animals do not manifest such distinctions in conduct, no amount of knowledge arising from animal behavior would be enough to this end.

Now, the claim cannot be based on the idea that the data is *scarce*, besides being *in principle* insufficient. If it is indeed scarce, it can be augmented and it is thus not insufficient in principle. So the problem lies in the kind of evidence, and specifically in its non-linguistic nature. But non-linguistic evidence can lead us to the same interpretations of thought as its linguistic analogs Routley 1981; Bermúdez 2003, p. 100; Beck 2013, p. 530). A person that *tells* us that she remembers the order in which four visitors arrived is not making use of a different capacity —and thus expressing a different intentional attitude— than the rat that *picks up* the odor corresponding to the second to last of a list. Both actions show the ability to *remember* a sequence of events. Similarly, the person that *tells* us that the toy is on the left box is not expressing something different than the chimp that *tugs* at the left box to help the experimenter open it. Both are expressing their *understanding* of another's goal.

The growing field of research in animal cognition has gathered an impressive amount of evidence by way of improving the design of the experiments and refining the hypotheses. The main source of indeterminacy cannot, therefore, be the chronic and/or principled lack of evidence coming from the research in non-linguistic animal cognition. We may apply here, somewhat ironically, this quote from Davidson: 'Success in interpretation is always a matter of degree (…) It is always possible, of course, to improve one's understanding of another, by enlarging the database, by adding another dose of sympathy or imagination, or by learning more about the things the subject knows about.' (Davidson 1996, p. 232).

One must concede that natural languages are invaluable tools to articulate thoughts. But even this may be qualified. Some think that natural language is not the primary criterion for the attribution of thought (Routley 1981, p. 406) since many ascriptions are based on non-linguistic evidence. Relatedly, linguistic competence is not a necessary condition for attributing and specifying attitudes, for (as we have seen) we usually make use of non-linguistic evidence. Finally, even though language is generally a sufficient condition, a particular statement asserted by a person may be undermined by their action, for speakers may be subject to self-delusion manifested in their conduct (Routley 1981,p. 406).[8] All in all, these considerations bring the animal and human cases closer to each other.

(b) Even if cognitive ethology has not 'opened a window' into animal minds (Griffin 1978) —which is a moot Cartesian metaphor— it has permitted us to 'make inroads' into the determinacy problem in the animal case (Glock 2020). How deep

[8] A further indeterminacy source is related to the diverse representational formats that some authors postulate to explain different mental capacities. According to them, when we translate the content of a non-linguistic format into a natural language we cannot avoid indeterminacy, even when the target thought is a human one (Beck 2013).

the research may go is largely an empirical matter; the hypothesis that rats and mice experience regret was perhaps not imagined some time ago has recently gathered plausible (even if fallible) evidence.

The research incorporates cognitive distinctions not previously made, and it may be the case that it gains no evidence to apply these distinctions more broadly. Certain mammals may not meet all the conditions for ascribing the concept of regret, and some reptiles may not even meet any of them (like the capacity of episodic memory possessed by birds and mice). In certain cases, we may assume that the subjects have indeterminate contents somehow related to regret, and in other cases that the attribution of regret is simply false. How much determinacy we can legitimately assume and where to give up the quest for more specific attributions, are in part empirical questions.

The inference to the best explanation is perhaps how the hypotheses are being proposed. This way, the behavior and the neurophysiological processes which were present just after the rats made the bad choice could initially be interpreted as regret or as disappointment. And we saw that the attribution of disappointment was discarded after the control conditions. Similarly, the attribution of episodic memory was tested against the hypothesis that these creatures were familiar with a sequence of events that left a trace in non-episodic memory. In general, the attributions have gained accuracy after some hypotheses were discarded and others provisionally confirmed. A survey of these lines of research thus shows how the control over the evidence is constantly improved by the design of the experiments, with the subsequent gain in attribution determinacy; and how this process is enhanced by the debate on the hypotheses and the results.

4 A Literalist Interpretation of the Attributions and its Limits

4.1 The Literalist View

So can we legitimately say that rats remember the order of particular events related to odors and regret having made bad choices, that chimps understand the goals of others and predict their conduct accordingly, either reading their minds or their conduct? For all we (fallibly) know in the context of the respective research areas, we may affirm that they do. And the philosophical objections that I have considered fail to draw an apriori and demanding differentialist line of division between humans and non-humans in the ascriptions of attitudes, for example with the possession of a natural language and the correspondent capacities.

I used a general strategy to mitigate the skeptical force of some philosophical arguments, which consisted in bringing the human and the non-human cases closer. This strategy was present early in the debate, for example in Routley: ‘The central behavioral criteria for the attribution of beliefs are common to humans and (other) animals.’ (1981, p. 405). I believe that the motivation for adopting this kind of strategy has increased over time, encouraged by diverse research programs in animal cognition. A similar perspective is expressed recently by Figdors` ‘literalism,’ which

is concerned with the attribution of psychological predicates as used in sciences such as biology and ethology (Figdor 2018, p. 11).[9]

In this section, I will interpret my conclusions so far as a version of her literalism (as in Figdor 2018), which I construe by means of two central theses that mirror the first two sections of this paper.[10] Firstly, Figdor makes a diagnosis of the state of the art in the 'life sciences,' which can be labeled her 'extension thesis.' In a similar spirit to section 1, she aims to show that science is using a psychological vocabulary to describe phenomena in unexpected non-human domains. According to her, this extension is taking place by means of qualitative and quantitative analogies (cf. Chapters 2 and 3). Notice that, since she understands science as a fallible enterprise, this descriptive claim does not imply that the psychological ascriptions are actually true (pp. 9–10).

The extension thesis faces the objection that the use of the psychological vocabulary outside the human realm can only be considered legitimate if the meaning of the psychological predicates is not taken literally. These kind of objections motivate the development of Figdor's 'semantic thesis,' according to which the meaning of the psychological terms used in the extension must be taken *prima facie* in their literal meaning. In her words, 'the semantic contribution of psychological predicates to truth conditions is the same in statements about the relevant nonhumans as it is in statements about humans' (p. 61).

In a similar way to section 2, her defence of this thesis takes into account several objections which come mainly from philosophers (like Bennett, P. M. S. Hacker, Searle, Dennett, Brandom, and Davidson, among others). The critiques are of a varied nature but they all point out in a similar direction: the interpretation of the relevant terms cannot be literal since their extended application to many non-human organisms (or parts of organisms) do not meet important requirements, for example, holistic constraints as the ones put forward in the philosophy of Brandom. Figdor develops an extensive argumentation against actual and possible objections to conclude that her literalist view should be seen as the default one to understand what the extension thesis foregrounds.

Both theses can be formulated as a comprehensive viewpoint: the use of a psychological vocabulary to describe and explain mental phenomena (as shown in behavior)

[9] As a reviewer of this paper pointed out, one can construe Figdor's position as primary concerned with cognition (as set by the discourse in the cognitive sciences) and not with intentional states (as discussed in philosophy). However, I think it is also possible to understand her literalism as an extensive discussion with philosophy, therefore assuming a common ground between both fields of inquiry. For one, she accepts that her theory can be understood in a philosophical vocabulary: in 'the traditional philosophical framework (…) thoughts are analyzed as attitudes towards propositions (…). In these terms, Literalism is a theory about attitude ascriptions, not content ascriptions.' (Figdor 2018, pp. 10–11). Furthermore, Figdor (2018) develops arguments against several philosophers: Sellars, McDowell, and specially Brandom are discussed in relation to their 'space of reasons' and its implications for a literalist view (Chap. 4); the Wittgensteinian objections coming from Bennett and Hacker have a central role throughout the debate; and a variety of other prominent philosophers are given voice in the dialectics (like Searle, Grice, and Dennett). Of course, one could argue that the assumption of common ground is in her case only for the sake of the debate. For my present purposes, I only claim that Figdor's view can also be construed as concerned with the intentional attitudes that philosophers discuss.

[10] For reasons of space, my reconstruction will leave aside certain aspects of Figdor's literalism, for example the ones concerning the metaphysical and ethical consequences of her position.

of many non-human organisms (and parts of organisms) has been extended by the scientific research in unexpected ways, and the best way to understand the meaning of the terms that are being used in this extension is the literal one. Since in many cases this leads to what seems an inadequate application of psychological terms (like flies *prefer;* neurons *decide*), there is growing pressure, in many research fields, to modify the meaning of some psychological predicates (p. 5).

This conclusion is similar to my claim regarding the flexibility of our mental concepts. The pressure on the revision of the concepts is partly caused by the possibility of employing them in a broader range of cases, which also suggests that something 'real' is being tracked. Against the 'Procrustean bed' picture of the attribution of intentional attitudes to animals, we can modify, revise, and specify our own concepts (Figdor 2018, p. 58). This change is a decision product of pressure that, according to Figdor, is currently happening with the psychological predicates in different research areas.

4.2 The Limits of Literalism

As I reconstructed Figdor's view, the acceptance of both the extension and the literalist theses puts pressure on conceptual change, notably when the extension includes non-human organisms not closely related to human beings, such as flies and bacteria, or parts of organisms, such as neurons. Figdor claims there are no apriori limits to conceptual revision ('the barn door [to a supposedly linguistic abuse] is *always* open' (Figdor 2018, pp. 84–85).

Without dealing with the metaphysical background of literalism, I will argue that there are limits to the revision of our psychological vocabulary. I think that the philosophical arguments that I discussed in section 2, even if mitigated, retain a considerable force. For reasons of space, I will only highlight two constraints arising from them.

The first one is represented by the holistic requirements of the attribution of psychological predicates. I have claimed (in section 2.1) that the attribution of a single intentional attitude (as expressed by a psychological predicate) implies the possibility of attributing a web of other attitudes. And we need a pattern of behavior sufficiently rich as evidential support for such attributions. I argued that this demand can be met by different research programs: one can say that rats regret since (among other reasons) one also finds consistent evidence that they can believe and know things, they have food preferences, they can remember events, and they can feel frustrated.

But our attributions seem to lose legitimacy as soon as the similarity between the non-human and the human conduct (and neurophysiology) vanishes. Figdor believes that this difficulty can be overcome by means of quantitative analogies, whose application to unexpected domains forces us to extend the use of our psychological vocabulary (Figdor 2018, Chap. 2). I doubt that this can be done without damaging the holistic constraints that I put forward. Even if the conduct of bacteria is remarkably flexible and can be partly explained by the same mathematical model as the one we use to explain certain human mental phenomena, one still needs a pattern of behavior that expresses a holistic network of propositional attitudes in order to attribute a particular psychological concept. If we attribute to bacteria the capacity to remember

something specific, we must also attribute to them the capacity to have beliefs. In the absence of a web of attributable attitudes, it is a moot question whether we can make sense of the idea of their being in a particular mental state at all.[11]

Whereas the first limit is set by 'external' holistic considerations (the network where the concepts are embedded), the second one is related to the 'internal' conditions of identity of our psychological predicates (their core notes). At this point, my critical remarks on Figdor will situate me closer to an anthropocentric perspective. I agree with her literalism that our concepts are more flexible and less attached to their original language game than it is sometimes assumed, but I think that each one of them has semantic notes that cannot be removed without the concept losing its identity.

Figdor seems to accept the notion that concepts have essential notes, at least for the sake of the discussion (see for example Figdor 2018, pp. 73–74), and she thinks that these notes can be discerned by means of empirical research, where the application of mathematical models to non-human domains plays a central role. One may raise the objection that a mathematical model is a structure that is interpreted to indicate what it represents (p. 34), so it cannot bring about a complete change of the concepts on which it depends. However, Figdor does not affirm that these models can *cause* a conceptual change by themselves but that they may *motivate* a conceptual revision. In her words, they "generate significant semantic pressure on our interpretation of psychological predicates" (Figdor 2018, p. 55).

As I understand her position, a revision amounts to a debate that includes at least two kinds of considerations: some of them are related to the success of the mathematical model to fit the data, and others concern the notes of a given concept that may (or may not) be abandoned. This is illustrated by Figdor several times: for example, the application of a mathematical model to non-human domains motivates the view that the concept of "surprise" does not necessarily include the note of "conscious affect" (Figdor 2018, pp. 55–57). The conceptual revision has for Figdor no apriori constraints, but one can suspect that a complete change of the core notes of a particular concept amounts to the formation of a different one. As mentioned above, in the debate Figdor makes use of the notion that a concept has essential notes. But whereas she believes that the discovery of those notes is an open empirical question, which may lead to a radical conceptual revision, I think that it partly (but crucially) depends on our capacity to understand and use the concept in question, to recognize it as the same one.

Putting both kinds of limits together, in cases where holistic and identity constraints are at stake, the pressure for a conceptual change loses its force. And if con-

[11] Figdor tries to sidestep these holistic considerations for the case of the propositional attitudes (not for the intentional contents) (pp. 77–78). As I already argued, I find this possibility unconvincing, since a mental life that consists in having only one kind of intentional attitude is not what we would recognize as a mental life at all. In the wake of her discussion with Brandom, she also follows another strategy (p. 78 ss). She argues that even if we concede that mental life must be embedded in a holistic normative dimension (the 'space of reasons'), we have no reasons to deny the possibility of finding a non-human normative dimension. But I think that this strategy does not work either. It is very improbable to find normative conditions as rich as the ones put forward by Brandom in relation to human language outside the human realm. Furthermore, even if it is possible to imagine a deflationated normative dimension, it is implausible to attribute it in the domains where bacteria and neurons display their complex behavior.

ceptual revision is not possible, one has to abandon one of the two initial theses of literalism for the particular case: either the *extension* is not legitimate in a literal sense, or the extension is legitimate but we cannot accept that the meaning of the psychological terms has to be taken *literally*.[12] As a result, the extension of a particular concept may have limits such as the human sphere or the mammalian sphere. Perhaps we may extend 'regret' to rats but not to bacteria, or we may not ascribe to rats the possession of something like a 'language.' The application of predicates must be pondered case by case, but I believe that a literalist view should be on guard against the attribution of concepts which lose their identity conditions and are not embedded in a proper holistic network of mental capacities.

5 Concluding Remarks

In this paper, I intended to show that it is possible to assume a common ground between two different fields of inquiry, represented by the scientific research (as illustrated in section 1) and the philosophical conceptual arguments (as reconstructed in section 2). I argued that the skeptical objections coming from the latter field can be mitigated and that there is room for conceptual change regarding the psychological predicates that are being attributed to non-human animals (in section 2). I then attempted a reconstruction of Figdor's literalism by means of two main claims that express a viewpoint similar to my own (in section 3).

Finally, I argued that the philosophical arguments, which I mitigated in section 2, put limits to literalism, more precisely, to the extent to which one can revise and apply our psychological concepts in non-human domains. I highlighted two constraints related to holistic considerations and to the conditions of identity of our psychological concepts. I thus defended a modest version of literalism as a plausible philosophical frame in which to understand the attributions of psychological predicates outside the human sphere.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1007/s13164-022-00643-z.

Acknowledgements I am grateful to Hans Johann Glock's research team at the University of Zurich (Switzerland) and to Thomas Muller's research team at the University of Konstanz (Germany) for helpful remarks, and to Laura Danón and Daniel Kalpokas for valuable comments and suggestions. The anonymous reviewers of this paper also helped me to improve the paper in several impor-tant ways.

[12] When the literal interpretation and the conceptual revision evince these limits one may dissociate the terms employed in the empirical field of cognition, on the one hand, from the mentalist vocabulary, on the other (Keijzer 2021 generalises this strategy, that I would only adopt where literalism falls short). Along the same lines, it may be objected —with Figdor- that even if the revised concepts are beyond recognition according to our folk psychological intuitions, they may still be part of the conceptual apparatus of valid scientific research. This may be the case, but it would come close to a 'technical view,' which was extensively criticised by Figdor, who affirms that 'technical terms often retain important semantic links to their non-technical cousins' (p. 72). I have no space to deal with the complex problem of the relation between folk psychology and science in Figdor's theory. But I would like to suggest that, since Figdor defends a literalist view according to which the concepts that are used in unexpected domains (flies *prefer*, neurons *decide*, etc.) should be interpreted as having the same meaning as they have in the human domain, one could expect them to retain some of their relevant everyday notes.

Authors' Contributions Not applicable.

Funding Not applicable.

Availability of Data and Material (data transparency) Not applicable.

Code Availability Not applicable

Declarations

Conflicts of Interest/Competing Interests Not applicable.

Ethics Approval Not applicable.

Consent to Participate Not applicable.

References

Andrews, K. 2009. Politics or metaphysics? On attributing psychological properties to animals. *Biology & Philosophy* 24 (1): 51–63.

Butterfill, S., and I. Apperly. 2013. How to construct a minimal theory of mind. *Mind and Language* 28 (5): 606–637.

Beck, J. 2013. Why We Can't Say What Animals Think. *Philosophical Psychology* 26 (4): 520–546. https://doi.org/10.1080/09515089.2012.670922.

Bermúdez, J. L. 2003. *Thinking Without Words*. Oxford: Oxford University Press.

Bermúdez, J. L. 2007. Negation, contrariety and practical reasoning: Comments on Millikan's Varieties of meaning. *Philosophy and Phenomenological Research* 75 (3): 663–669.

Brandom, R. 1994. *Making it explicit: Reasoning, representing, and discursive commitment*. Cambridge, Mass.: Harvard University Press.

Buttelmann, D., F. Buttelmann, M. Carpenter, J. Call, and M. Tomasello. 2017. Great apes distinguish true from false beliefs in an interactive helping task. PLoS One 12(4). https://doi.org/10.1371/journal.pone.0173793.

Crystal, J. D. 2018. Animal Models of Episodic Memory. *Comparative Cognition & Behavior Reviews* 13: 105–122. https://doi.org/10.3819/CCBR.2018.130012.

Crystal, J. D., and T. Suddendorf. 2019. Episodic memory in nonhuman animals? *Current Biology* 29 (24): R1291–R1295.

Darwin, C. 1982 (1871). *The Descent of Man and Selection in Relation to Sex*. Princeton, NJ: Princeton University Press.

Davidson, D. 1982. Rational animals. *Dialectica* 36 (4): 317–327.

Davidson, D. 1996 (1994). Donald Davidson. In: *A Companion to the Philosophy of Mind.* S. ed. S. Guttenplan S, 231–236. Blackwell, Oxford.

De Waal, F. B. M. 2000. Primates –A Natural Heritage of Conflict Resolution. *Science* 289 (July 2000): 586 – 90. https://doi.org/10.1126/science.289.5479.586.

Dummett, M. 1993. *Origins of Analytical Philosophy*. London: Duchworth.

Figdor, Carrie. 2018. *Pieces of Mind: The Proper Domain of Psychological Predicates*. Oxford: Oxford University Press.

Glock, H.-J. 2000. Animals, thoughts and concepts. *Synthese* 123 (1): 35–64. https://doi.org/10.1023/A:1005295521736.

Glock, H.-J. 2010. Can Animals Judge? *Dialectica* 64 (1). https://doi.org/11-33.10.1111/j.1746-8361.2010.01227.x.

Glock, H.-J. 2018. Animal rationality and belief. In *The Routledge handbook of philosophy of animal minds*, eds. K. Andrews, and J. Beck, 89–99. London & New York: Routledge.

Glock, H.-J. 2020. Determinacy of Content. The Hard Problem about Animal Intentionalit. *The Harvard Review of Philosophy* 27: 101–120. DOI https://doi.org/10.5840/harvardreview202072631.
Godfrey-Smith, P. 1999. Genes and Codes. In *Where Biology meets Psychology*, eds. D. J. Buller, and V. Hardcastle, 305–332. Cambridge, MA: MIT Press.
Griffin, D. 1978. Prospects for a cognitive ethology. *Behavioral and Brain Sciences* 1 (4): 527–538. https://doi.org/10.1017/S0140525X00076524.
Hacker, P. M. S. 1990. *Wittgenstein. Meaning and Mind*. Oxford: Basil Blackwell.
Heyes, C. 2014. Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science* 9 (2): 131–143. https://doi.org/10.1177/1745691613518076.
Heyes, C. 2015. Animal mindreading: what's the problem? *Psychonomic bulletin & review* 22 (2): 313–327. https://doi.org/10.3758/s13423-014-0704-4.
Hoerl, C., and T. McCormack. 2018. Animal Minds in Time: The question of episodic memory. In *The Routledge handbook of philosophy of animal minds*, eds. K. Andrews, and J. Beck, 56–64. London & New York: Routledge.
Keijzer, F. 2021. Demarcating cognition: the cognitive life sciences. *Synthese* 198 (1): 137–157. https://doi.org/10.1007/s11229-020-02797-8.
Lurz, R. 2018. Animal mindreading: the problem and how it can be solved. In *The Routledge handbook of philosophy of animal minds*, eds. K. Andrews, and J. Beck, 229–237. London & New York: Routledge.
Malcolm, N. 1972. Thoughtless Brutes. *Proceedings and Addresses of the American Philosophical Association* 46: 5–20. https://doi.org/10.2307/3129585.
Marcus, R. B. 1990. Some Revisionary Proposals about Belief and Believing. *Philosophy and Phenomenological Research* 50: 133–153.
Newen, A., and T. Starzak. 2020a. How to ascribe beliefs to animals. Mind & Language, 1–19. https://doi.org/10.1111/mila.12302.
McDowell, J. 1996. *Mind and World*. Cambridge, MA: Harvard University Press.
McDowell, J. H. 2009. *Having the World in View: Essays on Kant, Hegel, and Sellars, Cambridge*. Cambridge, MA: Harvard University Press.
Millikan, R. 2005. *Language: A Biological Model*. Oxford: Oxford University Press. https://doi.org/10.1093/0199284768.001.0001.
Millikan, R. 2007. Reply to Bermúdez. *Philosophy and Phenomenological Research* LXXXV (3): 670–673.
Newen, A., and T. Starzak. 2020b. How to ascribe beliefs to animals. Mind & Language 37: 3–21. https://doi.org/10.1111/mila.12302.
Panoz-Brown, D., V. Iyer, L. M. Carey, C. M. Sluka, G. Rajic, J. Kestenman, … K. G. Tucker. 2018. Replay of episodic memories in the rat. *Current Biology* 28 (10): 1628–1634. https://doi.org/10.1016/j.cub.2018.04.006.
Premack, D., and G. Woodruff. 1978. Does the chimpanzee have a theory of mind? *Behavioral and brain sciences* 1 (4): 515–526.
Routley, R. 1981. Alleged problems in attributing beliefs, and intentionality, to animals. *Inquiry : a journal of medical care organization, provision and financing* 24 (4): 385–417. https://doi.org/10.1080/00201748108601945.
Rowlands, M. 2009. *Animal rights: Moral theory and practice*. New York: Palgrave Macmillan.
Savigny, Eike v. 1995. Keine Hoffnung für Hunde. In *Wittgenstein über die Seele*, ed.O. Scholz O & E. von Savigny E., 41–56. Frankfurt am Main: Suhrkamp.
Searle, John. 1994. Animal minds. *Midwest Studies in Philosophy* 19 (1): 206–219. https://doi.org/10.1111/j.1475-4975.1994.tb00286.x.
Steiner, A., and David Redish. 2014. Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nat Neurosci* 17 (7): 995–1002. https://doi.org/10.1038/nn.3740.
Sterelny, K. 2003. *Thought in a Hostile World*. Oxford: Blackwell.
Stich, S. P. 1979. Do animals have beliefs? *Australasian Journal of Philosophy* 57 (1): 15–28.
Sweis, B. M., M. J. Thomas, and A. D. Redish. 2018. Mice learn to avoid regret. PLoS Biol 16(6). https://doi.org/10.1371/journal.pbio.2005853.
Williams, B. 1973. *Problems of the Self*. New York: Cambridge University Press.
Wittgenstein, L. 1967. *Philosophische Untersuchungen*. Frankfurt am Main: Suhrkamp.