BELIEF, ANIMALS, ROBOTS

PIERPAOLO MARRONE
Università Trieste
Dipartimento di Studi Umanistici
marrone@units.it

ABSTRACT
In this article I deal with the problem of the attribution of belief, its potential for the normative problem of the definition of person, the implications of the apparent ability of some machines to form concepts.

KEYWORDS
Beliefs, animals, robots.

It would seem that the ability to form concepts is a prerequisite for rational action. Whatever the definition of rationality we want to use, the ability to generalize experiences in order to orient oneself in the future and to correctly interpret the past is an essential component. This component seems to be particularly necessary for a rationality conceived instrumentally, that is understood as the ability to choose the appropriate means in view of the achievement of specific ends. It would appear that animals are indeed capable of instrumental rationality (Andrews, 2015; Glock, 2000; Glock, 2009; Oslo, 2010). The evidence is the surprising adaptive capacities of living organisms and their survival during the evolutionary process. But possessing and exhibiting behaviors that can be interpreted according to the criteria of instrumental rationality does not seem to be equivalent to the ability to form concepts. Certain plants have developed adaptive mechanisms to attract insects for use as pollinators, but it seems risky to say that this has to do with plants' ability to form concepts. So it could be inferred that it is one thing to define a behavior as rational and another thing to define an agent as rational. There may be behaviors that can be interpreted according to some definition of rationality (for example the instrumental one), but they do not at all define the agent who has that behavior as rational. In fact, if we excessively extend the concept of rationality of an action according to the definition of instrumental rationality, we will make it coincide with that of sufficient reason ("since everything that happens happens for a reason, then everything hap-

pens as it should happen") , but this would make it useless for understanding what it means to be a rational agent.

One can therefore act rationally without being rational agents in the full sense. This immediately raises the question of what makes an agent rational (Jago, 2014; Roelofs, 2017). The first answer that perhaps intuitively many would come to mind is that the agent must exhibit some form of intentionality (Davidson, 1980; Dennett, 1989). How this form can be ascertained is a controversial question. Perhaps in animals an indirect index can be considered the ability to lie. If I can deceive another animal, then this indicates the presence of a complex thought, something that could be schematized as "wanting someone to want something and induce a behavior to do so". When my cat, meowing and heading towards the balcony where the bowl with her food is, causes me to get up from my armchair on which she then goes to sit, I think she has managed to make me do something that I would not have wanted to do for achieve its goal more easily. If she manages to do it more than once I can begin to think that this did not happen by chance and that my cat has an ability to plan future actions with a certain awareness, thanks to the adoption of recurring patterns like "if I meow insistently heading towards the balcony, then he will get up from my favorite chair and free it ". Of course this is a schematization that could be completely inappropriate. It is not necessary that my cat has the concept of "balcony" or "favorite chair" or "insistently". However, something like this has to happen in her thinking for her to actually be able to plan to get me out of my favorite chair that is also hers (Oliver, 2009). Now, the question is what exactly needs to happen.

The omnipervasiveness of machines in our lives is a phenomenon that I do not think even needs to be emphasized to understand that it raises many fundamental questions for those involved in philosophy (Cingolani, 2019). Perhaps even more existentially pressing questions arise from the forthcoming invasion of robots in the activities that make up our daily lives (Fry, 2019). I am not referring to small vacuum cleaners that can be operated from a remote dedicated app as is the case with air conditioners, heating systems, washing machines, safety devices and so on. Ultimately we can interpret these mechanical systems as prostheses of our will. From this point of view they are not very interesting, if not for the fact that they constitute a further chapter of our never concluded hybridization with machines and artifacts, a real human universal present in all cultures where homo sapiens is present. What is most interesting from my point of view are those cases where machines are able to make decisions on their own (Danaher, 2016).

It is necessary to specify what "autonomously" means here. As is known, the concept of autonomy is the basis of our conception of the individual as a personal entity capable of making plans and making decisions that can be traced back to the

individual himself as his focal center. The autonomous subject would therefore be the one who traces in himself the source of his decisions. This concept, which underlies both the Cartesian conception of substance and the obsession of modern and contemporary philosophy with the subject, is highly controversial (Dworkin, 1988; Rosslenbroich, 2014; Schneewind, 1997). If we understand autonomy as a form of spontaneity, then it would seem that we are forced to enter into inevitable paralogisms. Since we are neither self-created entities nor self-programmed entities, it is difficult to believe that our decisions are self-created and self-programmed. Yet the concept of autonomy has to do with self-programming, since it means precisely the ability to give rules of conduct to oneself. So, when we talk about our ability to make decisions, we mean something very complicated and not the simple idea that action is the visible manifestation of a will that has no other foundation than in itself.

I will not go into the merits of the question of freedom of will, which inevitably arises every time autonomy is discussed. It is not my intention to affirm or deny this freedom. What it is enough for me to establish is that our autonomy is conditioned by many factors ranging from our genetic code, to the environment in which we have been brought up and educated, to the intertwining of our individual biology and conditions. socio-environmental that we find built around us and that we have not formed in any way. It is within this format, in fact, that our ability to choose is exercised. We process the information that comes to us from a selection of environmental stimuli and internal stimuli, we select those that we believe to be relevant, we draw from our past experience what can be useful for tracing persistent patterns, we adopt, in short, together with the previous information all a series of conditional strategies such as "if ... then". It is with these premises that our action is exercised, where autonomy must be considered a label that identifies a class of actions that we easily refer to ourselves as their contextual starting point (Sapolsky, 2018).

Actions are often exercised on the basis of our past experience and this past experience has value for us above all because we search for it and often find persistent patterns. Some of these persistent patterns we call "concepts". Concepts are indispensable tools for our orientation in the world. They make us understand what to do ("don't get too close to what only vaguely looks like a slightly oversized cat"), they make us understand what expectations it is rational to have ("it is not the same thing to take a psychotropic substance and an asparagus soup "), Allow us to make plans for the future (for example by making us understand that a vacation is something different from a move). All these examples are based on categorizations of our or others' past experiences and fall into the category of empirical or post-hoc concepts,

which need experience to be formed and used (Löhr, G., 2019; Vasilecas, O., Bugaite, D. , Trinkunas, J., 2014).

We attribute the intentional attitude with some ease to explain many activities related above all to the actions of living organisms. We make inflationary use of it indeed. According to Davidson, this inflationary use should be one of the elements that should advise us to be cautious and suspicious. Perhaps we are handling concepts without being well aware of their implications and perhaps we are assuming that thought, intentionality, rationality are involved in some holistic whole that unites at least a large part of living organisms. But how do we normally recognize those characteristics that we think are rational, including the attribution of intentionality and the ability to form patterns and recognize patterns (what we normally call producing concepts)? We recognize in others - we believe we recognize in others - an attitude that is in us, which Davidson in his influential article Rational Animals calls a "propositional attitude" (Davidson, 1985), that is, the ability to have things like a belief, a desire , an intention, the ability to be ashamed, to deceive, to distinguish true from false.

The intentional attitude, the shame, the lie, the distinction between true and false can all be grouped under the concept of belief. However, how can you be sure that some organism or some system with a cognitive organization of a certain complexity has a belief? For Davidson there is a sort of test to pass the examination of the ascription of a belief: one must be able to form the concept of belief. This is for Davidson a central intuition and of the utmost importance. No animal can have a belief if it does not also have the concept of belief. For Davidson this is not possible and for a direct reason. Nobody can have a belief unless they also understand the possibility of being deceived or being wrong. This means having the necessary mental qualities to distinguish true from false, that is, the ability to discriminate between true beliefs and false beliefs. I don't know if it can be said that Davidson has presented an argument to prove this intuition. It seems to me not, but I don't even think this is entirely relevant. I believe that Davidson has produced something different in this position, that is, he has introduced a sophisticated network of concepts and problems, to show that we do not have sufficiently clear ideas on the whole issue, even if one of his strongest conclusions is that to have a belief it is necessary to have the concept of belief and this leads him to argue that it is not at all clear in what sense an animal can have a belief. This conclusion qualifies, at least prima facie, his position as highly skeptical about the ascription of conceptual abilities in animals.

Think of this example: I see a dog chasing a cat running across a square that has only one large tree, a majestic magnolia, let's say. This majestic magnolia is also the

oldest tree in the whole neighborhood. The cat goes around the magnolia and takes refuge on the tree. The dog barks furiously at the tree. In what sense does the dog have the belief that the cat is on the magnolia? Was the concept of cat formed, was the concept of tree formed? None of us would think that the concept of the oldest tree in the whole neighborhood has been formed. He is certainly not able to substitute "magnolia" for "oldest tree in the whole neighborhood" in the relevant contexts salvo veritate. But is it really necessary for something like this to happen (Toribio, 2010)? Forming a concept can be considered an operation that involves different degrees of complexity. As far as we know, one of the major complexities is allowed by the formation of concepts within human language. In this context, it makes sense to request that it be possible to replace "magnolia" with "oldest tree in the whole neighborhood" salva veritate, in order to verify a correct understanding of the belief that magnolia is the oldest tree in the whole neighborhood . If a normal speaker is unable to make the usual substitutions in the relevant salva veritate contexts, then we conclude that he has not formed the concepts correctly. However, what is correct for human speakers, since it cannot be asked of the dog, must it lead us to conclude that the dog did not form the concept of cat or the concept of tree? Somehow if the magnolia were surrounded by a flower bed the dog should know that the cat could not have climbed the flowers (Colin, 1999). He must have somehow thought that since the cat is an animal that climbs easily on surfaces that tend to be vertical, then he must have climbed the tree and not the flowers. Even if it does not have the concept of belief in the sense requested by Davidson, the dog must still have the belief that the cat is on the tree. He must have well-founded reasons in support of this belief, because otherwise we would find it difficult to explain why the dog continues to bark at the tree. This idea somehow takes for granted the idea that thought, although in homo sapiens closely intertwined with language and although it is impossible for us to clearly establish the boundaries between thought and language (as well as between belief and propositional attitude), nevertheless is also present in organisms that do not possess verbal language, while they possess other communication systems, including interspecific ones. There seems to be no cheaper and better way to explain dog behavior than through its network of beliefs which also involves the previous formation of conceptual networks (Carruthers, 2013). What is true for the dog obviously also applies to the cat that when the dog chases him "he knows that the dog knows" that he must not allow him to reach the majestic magnolia in the center of the square. Is it true, however, that this is an example of how very easily we are induced to inflationary interpretations in the attribution of conceptual networks, beliefs, intentionality? I do not think, while if I did not know with a certain amount of significant detail what a missile is, I could think that the missile that

hit a target manifested its intention to hit that target after having identified it, discriminated against it from the rest of the environment. , having formed conceptual networks that allowed him to perform the task he had chosen to do. Obviously, this would be a completely wrong explanation, and anyone who points out that behind the missile hitting the target there are human operators who make decisions would not simply provide an alternative explanation to mine, but would have a better explanation. Its explanatory network is superior to mine because it describes reality in a better way than mine, even though we both possess a concept of belief that distinguishes us from the dog and cat of the previous example.

What is common to our different and elaborate belief systems is that they can be described as true or false. This does not seem to be possible for Davidson for other systems such as those implemented by animals or by infants, or by Alzheimer's patients. Having a concept means having the belief that the mental scheme that defines it is related to the ability to distinguish some things from others in a significant number of qualified contexts. Is this an excessive request? Of course we have only indirect indices of the ability of other minds to form concepts in this sense and eventually to have the concept of belief. But what does it mean to have a belief? I think having a belief can be described as thinking that something in the world is or has been or will be so and so. To ascribe a belief to someone is to believe that they have a strategy of behavior in the world based on these beliefs (Alechina and Logan, 2010). To have a belief in this sense, is it also necessary to know how to specify that this is a belief? I think no. It is certainly necessary to have what may be called belief related to belief if we do some reflexive thinking operations (Burgos, 2007). It is obviously true that there are biological systems that possess beliefs and concepts in the sense of Davidson, for example homo sapiens is such a system. It also seems clear to me that the ascription of intentionality is as much a crucial operation of exhaustive description of an action as a natural strategy.

The problem is therefore not intentionality for us or the ability to have a belief and to believe we believe, but the fact that this is not an excessively demanding condition for interpreting other systems of thought. Animals that are unable to make claims of reflexivity on their own thought systems are not even able to have a strong belief? Probably this is the case, and if it is correct to suggest that it is this lack that should make us think that one thing is the ability to discriminate and another thing, that is a qualitatively different operation, is having an intentional attitude, it should not lead us to conclude that it is It is necessary to have an intentionality in a reflexive sense (precisely the sense that Davidson suggests) to be able to say that one has a belief, differentiated from the simple discriminatory capacity that is typical of cats, dogs and other animals. If we then compare living beings and artifi-

cial systems, one of the first distinctions that could be made is that living systems force us, when we interpret their behavior, to make our own an assumption of rationality that cannot be present in artificial systems except as an attribution by the designer.

What Davidson says, in my opinion, is not precisely the denial of intentionality to systems less complex than the adult human with average cognitive performance, but rather our difficulty in marking precisely where the inability to be intentional and to have a cupboard. It is very likely that neither the dog nor the cat that is chased by the dog possesses the concept of truth, since the true and the false are not in things, but in our descriptions of states of the world and as far as we know we cannot say that the dog and the cat have propositional descriptions of the world. However, that doesn't mean they don't have descriptions of the world as it appears to them. This was the position of Stitch, for example, who thought that if animals have beliefs, then the best explanation of their behavior must be in terms of an explanatory theory similar to the one we adopt for human behavior (Stitch, 1979). Davidson would probably take this position as an undue assumption of the point that it is about to prove, but the problem in my opinion is precisely whether a demonstration of something like this is possible. Since we have to rely on behavioral indices, it will never be possible, if not through these indirect indices, to have true beliefs justified about the beliefs that certain systems have or could have. Belief could be thought of as a complex system that has degrees and is placed along a continuum. Perhaps thought is also a continuum. This is not at all in contradiction with the fact that some cognitive performances were only possible starting from the appearance of language and the mechanisms of anticipation, prediction, representation, feedback that language allows and thanks to which it constituted an extraordinary evolutionary advantage for homo sapiens.

The condition of believing to believe in order to have the certainty that one has intentionality is, however, too demanding a condition. More: I think that even the condition of being able to have a propositional attitude as a necessary condition for having the ability to form concepts is too demanding. In fact, I think it is too demanding even with respect to the condition of truth. Of course neither the dog nor the cat are able to form propositions, but if we accept that the dog knows that it is chasing a cat and we think that it can discriminate between a cat and a remote-controlled mechanical object, then it will also know that the remote-controlled object does not it is a cat, although at first it was deceived and led to chase it. He had a belief about the remote controlled object that proved inadequate to his expectations, if we don't want to use the word "false", and this inadequate belief forced him to reset his behaviors with respect to the remote controlled object and with respect to

many other things that they are in its immediate environment. It can then be said that the dog was able to distinguish between an inadequate and an adequate belief. After all, aren't the facts reporting this to us? If we are willing to admit it, shouldn't we also be willing to concede that true and false are properties of beliefs even before they are properties of utterances? After all, wouldn't it be an evolutionary advantage to be able to discriminate true (adequate) beliefs from false (inadequate) beliefs?

Before all this, the ability to form concepts must be located. What might be some of the characteristics you need to exhibit to pass a concept forming test? One test might be the ability to improve your performance against past errors. Evidence that many animal species are capable of such learning abounds (Safina, 2016; Godfrey-Smith, 2017). I believe that this evidence simply justifies the idea that animals form concepts.

I think that the ability to form concepts precisely in the sense of forming patterns that can be used for the purposes of discrimination in the environment and for the purposes of self-learning is a necessary condition of thought, but I also believe that it is absolutely not a sufficient condition. In fact, self-learning and the ability to discriminate have also been characteristics of artificial expert systems for some time and we do not believe that these systems think. Think of the progress that Google Translate has shown in recent years, moving from a generally crude system to a system capable in many cases of remarkable linguistic translation performance. All this happened thanks to the enormous computing power that by accessing the big data archives can compare the occurrences of a word or an expression in milliseconds and choose the most appropriate context for that expression. In the security sector, facial recognition techniques have shown impressive ability to discriminate faces in contexts characterized by the presence of large crowds, such as in stadiums. The expert systems that finalize machine learning, deep learning and machine learning, no longer live in a limited deterministic world which could be considered the chessboard on which Deep Blue defeated Garry Kasparov in 1997, but in a world where there are machines. to gain experience (Vespignani, 2019). In our world we are subjects who intuitively know many more things than we can ever communicate. For example, how to get out of a car (a task that is still difficult for a robot to perform) or recognize a face. Expert systems elaborate automated schemes from the analogue of experience and this analogue is big data.

By using statistical functions that classify repetitions, the data are interpreted as the probability of events and the probability of an event is a classification with predictive capacity. Thus a machine is able to recognize with a given probability whether a human voice belongs to a man or a woman, whether it is an adult, elderly

person or a child. The algorithm that recognizes images operates on the basis of statistical occurrences relating to the position of an object, its lighting, its angle with respect to the plane and the background. All these are stylizations, it may be said, but for the purposes of learning in numerous sciences, such as anatomy and ornithology, the drawings that select the traits that must be considered relevant in that knowledge are extremely effective. The occurrence of the single bird or single organ under our eyes is recognized in deep learning by stratifications of abstractions - this is the meaning of "deep", which indicates not so much a depth behind reality, but rather a reticular system of multi-level abstractions (Accoto, 2019) -. Thus a face is recognized starting from the analysis of the individual pixels up to being able to outline what appears with high probability as a face and so on. Similarly, an autonomous driving system discriminates against other vehicles from pedestrians who are crossing the road or a diagnostic algorithm surrounds the cancer cells in a diagnostic image. The basic point, which is the foundation of ever more extensive learning (the so-called "deep neural network"), is always the identification of recurring patterns (which also allow the recognition of statistical anomalies, such as fraudulent transactions with credit cards, for example). In this sense it seems indubitable to me to argue that machines form concepts, just as animals form concepts too.

Another thing can then be said. The Bayesian interpretation of probabilities interprets frequencies as degrees of confidence and approximation to truth. Is it possible to give a similar interpretation of deep learning, as an increase in the degree of approximation to the best result for the algorithm that is close to the truth? If so, then it seems difficult not to establish some further kinship between us, the non-human animals and them, the machines we have created. In addition to forming concepts, machines also come close to truth. It doesn't make much sense, I think, to say that we have developed the initial algorithm and therefore the machines that implement the performances that are allowed by these algorithms do not really form concepts. The same could be said of our biological ancestry, but this obviously makes no sense. However, there is one thing in common between deep learning and machine learning systems and our biological ancestry: we often don't know how the outputs are produced. A whole line of research has originated from this theme that underlines the hopes and potential risks of AI (Domingos, 2017; Bostrom, 2016). The idea very often underlying these fears is that we can only rely on machines if we are able to understand their operation in detail.

A concept can be considered an algorithm, that is a procedure to produce a result (a schematization, a knowledge, a set grouping) in a finite number of moves. The trend we are experiencing is where an impressive amount of decisions are made by algorithms not so much in activities of little interest because they are highly repeti-

tive, but in decision-making processes that have to do with the practices that we associate with democracy, practices ranging from bureaucratic dimension to the legislative one to the legal one. There are algorithms that select the people to whom to check baggage at the airport, algorithms that predict the tax evasion rate of certain categories, algorithms that recommend potential partners to contact on dating sites, algorithms for predicting our electoral preferences designed in based on our online activities.

By now, numerous scholars are beginning to talk about the dangers of algocracy (Danaher, 2016), a term that indicates not so much a world governed by computers, that world touted by countless science fiction films, as our world and the tendency to rely on machines for make decisions of public importance. Two different problems are condensed into this concern, both relating to a question of opacity. On the one hand, there is the problem of opacity in the use of our data, those data that we continuously produce. It is almost never clear how this data is collected and how it is used. Obviously they are used for commercial profiling purposes, but they can also be used for political propaganda aimed at small groups of voters. This data has been compared to the new oil and in fact there has been someone who has argued that the big companies that manipulate and manage these enormous masses of information should pay to obtain them (but there is a problem: we have voluntarily delivered , but not in exchange for anything, but in exchange for sociality). Those who have access to this huge mass of data, we do not know how they will want to use it, therefore. But this is only the first concern. There is another relating not to the opacity of decisions, but to what we could call epistemic opacity. The computing power of these systems, their ability to self-learn and self-correct, is such that cases begin to appear in which we do not understand how the final result was reached. Some fear that this will soon produce generations of algorithms that will far surpass the cognitive limits of human beings relative not to producing the result, as is already the case, but rather to understanding the result itself. This raises questions, which I will not enter into now, which concern the extent of political authority and the legitimacy of decisions, that is, our ability to participate in an informed manner in political decisions. The challenge of the systemic opacity of deep learning and machine learning is probably only at the beginning, because there is no lack of selective discrimination dangers, which could be considered biases implemented by the machine (Bacchini and Lorusso, 2019). And what is a prejudice if not a nonrepresentative or only partially representative conceptual scheme of empirical reality? Machines are capable of producing concepts even if they are not yet able to think and therefore certainly not capable of producing a belief about their own conceptual production.

How should Davidson's indications on the problem of ascribing conceptual skills be taken? Are they to be regarded as a resolutely skeptical position on the ability of animals to have concepts, as some have interpreted them, encouraged in this by Davidson himself? Or should they be considered as an indication of our difficulty in ascribing beliefs? Or they must also be taken for the holistic affirmation of a certain conception of thought that recalls that of Leibniz for which all the spiritual units, which he called monads and which exhaust the domain of reality, have a mental activity, differing in degrees. , from the dark to the conscious perception, which he called apperception (Leibniz, 2001) and which in Kant will become the focal center of the thought of the transcendental subject? Among these three options, I believe that the first must be excluded, while the second and third must be accepted. These are two options that are profoundly in solidarity with each other, because they hypothesize a continuum both in belief and in thought. In other words, belief is not at all a quality that is either present or absent, but resembles a color that can be more or less intense. The more the cognitive systems become complex on the evolutionary scale, the more one has a more "saturated" belief. But it is very difficult to say where precisely the belief appears and where it disappears. So my idea is that the ability to form concepts is a completely necessary condition for there to be thought, but it is absolutely not a sufficient condition. Many of the doubts that are raised about the ability of machines to produce and use concepts, now also autonomously, as happens with facial recognition software, derive from problems concerning the opacity of learning processes. This opacity generates concern when it becomes opacity in decision-making processes, as has been said with regard to algocracy, because it expropriates decision-making opportunities. Whether this is always bad is controversial. Just as it is of course controversial whether the machines we build will ever be able to think one day. This was not the subject of these pages which instead concerned the ability of machines to produce concepts. My thesis was that if this ability is recognized to certain animals, then there is no reason not to recognize it to certain machines. This is not enough for there to be thought and belief to the extent that thought and belief exist in the human being. But the foundations were laid by evolution in one case and our creativity in the other.

References

Accoto, C. (2019), *Il mondo ex machina. Cinque brevi lezioni di filosofia dell'automazione*, Roma, Egea.

Alechina, N. and Logan, B. (2010), *Belief ascription under bounded resources*, "Synthese", pp. 179-197.

Andrews, K. (2015), *The Animal Mind: The Philosophy of Animal Cognition*, London, Routledge.

Bacchini, F., Lorusso, L. (2019), *Race, again. How face recognition technology reinforces racial discrimination*, "Journal of Information, Communication and Ethics in Society", pp. 321-335.

Bostrom, N. (2016), *Superintelligence. Paths, Dangers, Strategies*, Oxford, Oxford University Press.

Burgos, J.E, (2007), *About Aboutness: Thoughts on Intentional Behaviorism*, *"*Behavior and Philosophy*"*, pp. 65-76.

Carruthers, P. (2013), *Animal Minds Are Real, (Distinctively) Human Mind Are Not*, *"*American Philosophical Quarterly*"*, pp. 233-248.

Cingolani, R. (2019), *L'altra specie. Otto domande su noi e loro*, Bologna, Il Mulino.

Colin, A. (1999), *Animal Concepts Revisited: The Use of Self-Monitoring as an Empirical Approach*, "Erkenntnis", pp. 33-40.

Danaher, J. (2016), *The Threat of Algocracy: Reality, Resistance and Accommodation*, "Philosophy & Technology", pp. 245-268.

Davidson, D., (1980), *Essays on Events and Actions*, Oxford, Clarendon Press.

Davidson, D., (1985), *Rational Animals*, in LePore E. and B. McLaughlin (eds.), *Actions and Events,* Blackwell, Oxford, pp. 473-480.

Dennett, D. (1989), *The Intentional Stance*, New York, Bradford.

Domingos, P. (2017), *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World*, London, Penguin.

Dworkin, G. (1988), *The Theory and Practice of Autonomy*, Cambridge, Cambridge University Press.

Funkhouser, E. (2017), *Beliefs as signals: A new function for belief*, "Philosophical Psychology", pp. 809-831.

Fry, H. (2019), *Hello World: Being Human in the Age of Algorithms*, New York, Norton & Company.

Godfrey-Smith, P. (2017), *Other Minds: The Octopus, the Sea, and the Deep Origins of Consciousness*, New York, Farrar Straus & Giroux.

Glock, H. J. (2000), *Animals, thoughts, and concepts*, "Synthese", 123: 35–64.

Glock, H. J. (2009), *Can animals act for reasons?* "Inquiry", pp. 232-254.

Jago, M. (2014), *The Problem of Rational Knowledge*, "Erkenntnis", pp. 1151-1168.

Leibniz, G. W. (2001), *Monadologia*, Milano, Bompiani.

Löhr, G. (2019), *Embodied Cognition and Abstract Concepts: Do Concept Empiricists Leave Anything Out?* , "Philosophical Psychology", pp. 161-185.

MacIwan, I. (2019), *Macchine come me e persone come voi*, Torino, Einaudi.

Oliver, K. (2009), *Duplicity Makes the Man Or, Can Animals Lie?* in Martin, C., *The Philosophy of Deception*, pp. 104-117.

Oslo, K. (2010), *Rationality in the Domesticated Dog and Other Non-Human Animals*, "Teorema", pp. 135-145.

Roelofs, L. (2017), *Rational Agency without Self-Knowledge: Could 'We' Replace 'I'?*, "Dialectica", pp. 3-33.

Rosslenbroich, B. (2014), *On the Origin of Autonomy*, New York, Springer.

Schneewind, J. (1997), *The Invention of Autonomy: A History of Modern Moral Philosophy*, Cambridge, Cambridge University Press.

Safina, C, (2016), *Beyond Words: What Animals Think and Feel*, London, Picador.

Sapolsky, R. (2018), *Behave: The Biology of Humans at Our Best and Worst*, New York, Vintage.

Stich, S. (1979), *Do Animals Have Beliefs?,* "Australasian Journal of Philosophy", pp. 15-28.

Toribio, J. (2010), *The Animal Concepts Debate: A Metaphilosophical Take*, "Teorema", pp. 11-24.

Vasilecas, O., Bugaite, D., Trinkunas, J. (2014), *Knowledge Expressed By Ontology Transformation into Conceptual Model*, "Communication and Cognition", pp. 13-24.

Vespignani, A. (2019), *L'algoritmo e l'oracolo. Come la scienza predice il futuro*, Milano, Il Saggiatore.