Les ateliers de l'éthique
The Ethics Forum

Moral Animals and Moral Responsibility

Albert W. Musschenga

Volume 10, numéro 2, été 2015

URI : https://id.erudit.org/iderudit/1035326ar
DOI : https://doi.org/10.7202/1035326ar

Aller au sommaire du numéro

Éditeur(s)

Centre de recherche en éthique (CRÉ)

ISSN

1718-9977 (numérique)

Découvrir la revue

Citer cet article

Musschenga, A. W. (2015). Moral Animals and Moral Responsibility. *Les ateliers de l'éthique / The Ethics Forum*, *10*(2), 38–59. https://doi.org/10.7202/1035326ar

Résumé de l'article

La question centrale dans cet article est celle de savoir si les animaux sont moralement responsables de ce qu'ils font. Répondre à cette question nécessite une argumentation minutieuse et progressive. Dans les sections 1 et 2, j'explique ce qu'est la moralité, et qu'être doté de moralité signifie de se conformer à des règles ou à des normes morales. Dans les sections 3 et 4, je pose que certains animaux ne se contentent pas de montrer des régularités dans leur comportement social, mais qu'ils suivent aussi véritablement des normes sociales. Toutefois, les normes qu'ils suivent sont-elles *morales*? Dans la section 5, je prétends, en me référant aux travaux de Shaun Nichols, que les compétences ou capacités morales de base sont déjà présentes chez les primates non humains. Respecter des règles ou des normes morales est bien plus que d'agir conformément à ces normes; cela requiert d'être motivé par des règles morales. Dans la section 6, me référant à Mark Rowlands, j'explique que d'être capable de motivation morale ne nécessite pas d'être un agent moral; le fait d'être un sujet moral suffit. Contrairement aux agents moraux, les sujets moraux ne sont pas responsables de leur comportement. Affirmant qu'il y a de grandes similitudes entre le comportement moral d'un animal et le comportement inconscient, automatique et habituel d'un humain, je me penche, dans la section 7, sur la question de savoir si les humains sont responsables de leur comportement moral habituel, et si tel est le cas, sur quelle base on peut alors nier le fait que les animaux sont eux aussi responsables de leur comportement. La réponse à cette question est que les humains sont responsables de leur comportement habituel s'ils disposent d'une capacité d'intervention réfléchie, délibérée. Bien que les animaux soient capables d'intervenir dans leurs comportements habituels, ils ne peuvent le faire de manière réfléchie, délibérée.

MORAL ANIMALS AND MORAL RESPONSIBILITY

ALBERT W. MUSSCHENGA
PROFESSOR OF PHILOSOPHICAL ETHICS
DEPARTMENT OF PHILOSOPHY, VRIJE UNIVERSITEIT

ABSTRACT:

The central question of this article is, Are animals morally responsible for what they do? Answering this question requires a careful, step-by-step argument. In sections 1 and 2, I explain what morality is, and that having a morality means following moral rules or norms. In sections 3 and 4, I argue that some animals show not just regularities in their social behaviour, but can be rightly said to follow social norms. But are the norms they follow also *moral* norms? In section 5, I contend, referring to the work of Shaun Nichols, that the basic moral competences or capacities are already present in nonhuman primates. Following moral rules or norms is more than just acting in accordance to these norms; it requires being motivated by moral rules. I explain, in section 6, referring to Mark Rowlands, that being capable of moral motivation does not require agency; being a moral subject is sufficient. Contrary to moral agents, moral subjects are not responsible for their behaviour. Stating that there are important similarities between animal moral behaviour and human, unconscious, automatic, habitual behaviour, I examine in section 7 whether humans are responsible for their habitual moral behaviour, and if they are, what then the grounds are for denying that moral animals are responsible for their behaviour. The answer is that humans are responsible for their habitual behaviour if they have the capacity for deliberate intervention. Although animals are capable of intervention in their habitual behaviour, they are not capable of deliberate intervention.

RÉSUMÉ :

La question centrale dans cet article est celle de savoir si les animaux sont moralement responsables de ce qu'ils font. Répondre à cette question nécessite une argumentation minutieuse et progressive. Dans les sections 1 et 2, j'explique ce qu'est la moralité, et qu'être doté de moralité signifie de se conformer à des règles ou à des normes morales. Dans les sections 3 et 4, je pose que certains animaux ne se contentent pas de montrer des régularités dans leur comportement social, mais qu'ils suivent aussi véritablement des normes sociales. Toutefois, les normes qu'ils suivent sont-elles *morales*? Dans la section 5, je prétends, en me référant aux travaux de Shaun Nichols, que les compétences ou capacités morales de base sont déjà présentes chez les primates non humains. Respecter des règles ou des normes morales est bien plus que d'agir conformément à ces normes; cela requiert d'être motivé par des règles morales. Dans la section 6, me référant à Mark Rowlands, j'explique que d'être capable de motivation morale ne nécessite pas d'être un agent moral; le fait d'être un sujet moral suffit. Contrairement aux agents moraux, les sujets moraux ne sont pas responsables de leur comportement. Affirmant qu'il y a de grandes similitudes entre le comportement moral d'un animal et le comportement inconscient, automatique et habituel d'un humain, je me penche, dans la section 7, sur la question de savoir si les humains sont responsables de leur comportement moral habituel, et si tel est le cas, sur quelle base on peut alors nier le fait que les animaux sont eux aussi responsables de leur comportement. La réponse à cette question est que les humains sont responsables de leur comportement habituel s'ils disposent d'une capacité d'intervention réfléchie, délibérée. Bien que les animaux soient capables d'intervenir dans leurs comportements habituels, ils ne peuvent le faire de manière réfléchie, délibérée.

INTRODUCTION

According to a widely accepted view, only humans are considered to be morally responsible for their behaviour. Only humans can be praised or blamed for what they do, and might in certain cases even be said to deserve punishment. Only humans have moral duties that do not stop at the border of the human community. Their moral community comprehends—at least—all sentient beings. In that view, human and nonhuman sentient beings have a different moral status. Duty-bearing humans are moral agents, while nonhuman animals as the objects of the duties of agents are moral patients. Humans as moral agents have duties *to* other humans and duties *regarding* all sentient beings. This became the received view after Peter Singer and others successfully attacked speciesism.[1] The distinction between humans as moral agents and animals as moral patients is now disputed, since, according to some scholars in animal behaviour, there is evidence that some animal species—e.g., chimpanzees and bonobos—have their social behaviour regulated by a morality, comparable to human morality, and not just by a functionally equivalent regulative system. A morality is not just a pattern of behavioural regularities. It is a system in which the regularities in social behaviour result from following social norms or rules. The recognition that at least some animals are moral—belong to a species that has a morality—might require us to revise the received view. Are moral animals morally responsible for what they do?

I cannot, of course, simply assume that some animals have a morality. I have to provide arguments. In sections 1 and 2, I explain what morality is, and that having a morality means following moral rules. In sections 3 and 4, I argue that some animals do not just show regularities in their social behaviour, but can be rightly said to follow social norms. But are the norms they follow also *moral* norms? In section 5, I contend, referring to the work of Shaun Nichols, that the basic moral competences or capacities are already present in nonhuman primates. Following moral rules or norms is more than just acting in accordance with these norms. It requires moral agency. However, can animals be moral agents? Since agency is generally seen as condition for moral responsibility, by showing that animals have agency, I would also have answered the question about the moral responsibility of animals. In section 6, I follow Mark Rowlands who distinguishes between moral subjects and moral agents. Moral subjects have the capacity for moral motivation. However, they are not moral agents, in the sense of what some authors call full agency. Animals can be motivated by moral reasons that are internal, but not available to their conscious, rational scrutiny. These reasons are, according to Rowlands, embodied in their non-conscious processing operations that are cognitively impenetrable. In the last section, section 7, I make a last attempt to prove that animals can have moral responsibility. I contend there that animal moral behaviour shows important similarities to habitual behaviour of humans. I then ask if the reasons why we hold humans responsible for their habitual behaviour also justify attributing responsibility to moral animals. The answer is that we cannot hold animals responsible while they lack the capacity for deliberate intervention in their behaviour.

1. A FUNCTIONAL DEFINITION OF MORALITY

Before going into the question of whether animals can have a morality, I have to clarify what I mean by morality. In my view, the best definition of *human* morality is a functional one, formulated by G.J. Warnock: "(t)he 'general object' of morality, appreciation of which may enable us to *understand* the basis of moral evaluation, is to contribute to betterment—or non-deterioration—of the human predicament, primarily and essentially by seeking to countervail 'limited sympathies' and their potentially most damaging effects." According to Warnock, becoming a moral person implies learning to resist and control one's always-present self-regarding tendencies. Morality's biggest enemy may be the pure egoist. But pure egoism is as rare as pure altruism. The average person has sympathy and concern, but only for a limited number of people—usually his or her family and friends. Therefore the proper business of morality is, in Warnock's view, "to expand our sympathies, or, better, to reduce the liability inherent in their natural tendency to be narrowly restricted."[2] Next to self-interestedness then, favouritism and partiality are on this view the most widespread moral problems.

Warnock speaks of expanding our sympathies. In his view, universal intent is a formal characteristic of morality. Moral rules are meant to guide and protect everybody. This is why his definition doesn't cover the moralities of human societies in which the moral community coincides with the social group. It is evident that within such a definition no system for the regulation of the social behaviour of a non-human species qualifies as a morality. A similar, but less restrictive functional definition is found in Jessica Flack and Frans de Waal. In their view human morality needs to take human nature into account either by fortifying certain natural tendencies—such as sympathy, reciprocity, loyalty to the group and family, and so on—or by countering other tendencies—such as within-group violence and cheating.[3] Flack and de Waal's definition can be broadened to cover animal morality, simply by skipping the adverb "human" in "human morality" and by substituting "human nature" for "animal nature." They themselves avoid speaking of animal morality. In their view, non-human primates have a *proto*-morality. Human moral systems, they say, rely on basic mental capacities and social tendencies that humans share with other co-operative primates, such as chimpanzees. That is why they regard it as justified to conclude that these other primates have a protomorality. Morality, however, also requires capacities that are present only in humans—such as a greater degree of rule internalization, a greater capacity to adopt the perspective of others, and the unique capacity to debate issues among themselves and transmit them verbally.

The broadest definition of the function of morality is given by Dale Peterson: "The function of morality, or the moral organ, is to negotiate the inherent conflict between self and others." This definition, he says, includes the possibility that at least mammals have moral systems homologous to ours.[4] Marc Bekoff and Jessica Pierce define morality as "a suite of interrelated other-regarding behaviors that cultivate and regulate complex interactions within social groups. These behaviors relate to well-being and harm, and norms of right and wrong attach to many of them."[5] They rightly distinguish between prosociality and altruism on the one hand, and morality on the other. To have a morality, they say, a given species must meet certain threshold requirements. These thresholds are the following: a level of complexity in social organization, including established norms of behaviour to which attach strong emotional and cognitive cues about right and wrong; a certain neural complexity that serves as a foundation for moral emotions and for decision-making based on perceptions about the past and the future; relatively advanced cognitive capacities (such as a good memory); and a high level of behavioural flexibility. All moralities consist of well-developed systems of other-regarding prohibitions and proscriptions.[6] The set of actions that constitutes moral behaviours varies among species. So does the degree of moral complexity. Morality can be thought of as nested levels of increasing complexity and specificity. Bekoff and Pierce don't enumerate the animal species that meet the threshold requirements. What they do say is that animals with a highly developed moral capacity may include chimpanzees, wolves, elephants, and humans.[7] This is not an exhaustive list. The distinction between human morality and animal morality is for them quantitative rather than qualitative. Humans appear to have evolved an unusually high level of moral complexity.[8]

The definition I suggest combines elements of the definitions by Warnock, Flack & de Waal, Petersen, and Bekoff & Pierce:

> Morality cultivates and regulates social life within a group or community by providing rules (norms) that fortify natural tendencies that bind the members together—such as sympathy, (indirect) reciprocity, loyalty to the group and family, and so on—and which counter natural tendencies that frustrate and undermine cooperation—such as selfishness, within-group violence, and cheating.

This definition leaves open the question of whether animals can have a morality. What it does say is that the mechanism for the moral regulation of social behaviour is rules. A species can be said to have a morality only if its supposedly moral behaviour is rule-governed. With humans not all rules are moral rules. If we want to find out whether a rule that people follow is a moral rule we ask them— e.g., how they justify it and what their motives are for following the rule, how other people react when they violate the rule, and what kind of feelings they themselves have on such occasions. Scholars in animal behaviour must start with observing the behaviour, body language, and facial expressions of animals, and the sounds they produce. If they observe certain regularities in their social behaviour, the next thing they do is to examine whether the regularity is caused by following a rule. Even if they can prove that animals follow a rule, additional evidence is needed to establish that the rule is a moral one and that the animals who follow it generally have moral motives. However, direct proof is impossible. The usual approach of animal behavioural scientists is more indirect. They try to find out if an animal species possesses the capacities that are needed for acting morally. If the answer is affirmative, they still have to show that an explanation of a certain behaviour in moral terms is the best one available. If we want to find out about the morality of another society, we usually start by looking for behaviour and practices that are similar to the behaviour and practices that fall within the scope of morality in our society. Scholars in animal behaviour do the same.

2. BEHAVIOURAL REGULARITIES, NORMS, AND MOTIVATION

Rules become visible in behavioural regularities, but not every behavioural regularity indicates the existence of a social rule. Habits are also behavioural regularities. For a group to have a habit, it is enough that the behaviour of most of its members on certain occasions in fact converges. This can be determined by observers from an external point of view, without recurring to the beliefs and attitudes of the group. Rules can guide behaviour in two ways: externally and internally. Rules govern behaviour externally when subjects conform to rules out of fear for sanctions. Rules govern behaviour internally when subjects have accepted and internalized rules. How can it be established that behavioural regularities result from internalized rules? According to the highly influential philosopher of law H. L. A. Hart, a common behavioural regularity must be explained by a rule (norm) if (1) deviation of the regularity elicits criticism, (2) the deviation is generally accepted as a good reason for criticism, and (3) the norm is seen as binding and obligatory. The third criterion requires that a subject accepts—and not just observes—a rule. This is what Hart means by taking "the internal point of view."[9]

I venture to say that the first and the third criteria are also useful for distinguishing norm-based behavioural regularities of animals from mere behavioural regularities such as habits. The second criterion is not relevant for that goal, since reasons for criticism can only be expressed in language. Criticism itself, however, can also be expressed in non-verbal form. Even humans use non-linguistic means to show their disapproval of a certain behaviour. They express it by gestures, facial expressions, and sounds, which are means of communication also available to animals. While the first criterion points to reactions of group members to norm transgression, the third criterion refers to the attitude of the agent toward the norm. When an agent only follows rules out of fear for sanctions, they are not binding and obligatory for him or her.

Hart's third criterion for distinguishing a statistical behavioural regularity from a rule-based one is the presence of the internal point of view. The internal point of view with regard to norms or rules is the point of view taken by someone who has internalized the norm, or, in more technical terms, who has the practical attitude of norm-acceptance. Someone who has internalized a norm is motivated by a felt obligation.

In the next section I discuss whether some animal species can be said to follow social rules or norms (Hart's modified first criterion). Section 4 deals with norm internalization (Hart's third criterion).

3. SOCIAL NORMS IN ANIMALS

In their article "Evolutionary precursors of social norms in chimpanzees," Claudia Rudolf von Rohr, Judith Burkart, and Carel van Schaik develop a theoretical framework for recognizing different functional levels of social norms and distinguishing them from mere statistical regularities.[10] They define social norms as behavioural regularities that are normative to a varying degree and generate social expectations. These expectations do not have to be experienced consciously. Their satisfaction or violation might, according to von Rohr et al., produce distinct reactions observable from the outside. Since meeting expectations is the normal situation, no reactions have to follow. But when a certain behaviour violates expectations, nearly always negative reactions ensue. Most important are the negative reactions by uninvolved bystanders.[11] Von Rohr et al. distinguish three types of negative reactions from bystanders on the violation of three different types of norms: 1) *Quasi social norms*. The negative reactions might simply be caused by specific cues. For example, when an infant is attacked and screams, bystanders flow to the scene and harass the perpetrator. This type of bystander reaction does not reflect violated social expectations, and most likely does not involve emotions such as indignation towards the perpetrator. Bystanders in this category probably do not possess any specific inference about how the distress of an infant and the behaviour of the perpetrator are linked together and thus are not able to perceive harming an infant as norm violation per se. Their reaction is a response to the stimulus of hearing the child scream.[12] 2) *Protosocial norms.* If bystander reactions cannot be explained by simple stimulus-response mechanisms, it might be that they are responding to norm violation as such. In this case, bystander reactions might also involve emotions comparable to indignation in humans. Bystander reactions to norm violation per se require the capacity to exhibit some empathetic competence because only this would enable bystanders to understand to some extent the distress of the mistreated infant , and also its cause. Von Rohr et al. assume that apes but not monkeys have empathetic competence, because monkeys seem to lack the capacity to attribute mental states to others.[13] 3) *Collective social norm.* Humans are endowed with sophisticated empathetic and cognitive abilities, which enable them to grasp the full extent and far-reaching consequences of mistreating children. Moreover, they are able to reason that infants are completely defenceless and therefore highly vulnerable creatures.[14] An important difference between the reactions of chimpanzees and those of humans to norm violation is that chimpanzees might experience indignation in a fairly individualistic way, while humans are able to share their feelings of indignation. Referring to Tomasello and Carpenter,[15] von Rohr et al. state that, by analogy with shared intentionality, shared indignation goes beyond simultaneous experience by different individuals and includes the awareness of a collective experience,that may lead to collective protest against, and condemnation of, the violator. This exemplifies the collective nature of a social norm.

Negative reactions by non-involved bystanders to the deviation from a socially expected behavioural regularity, accompanied by feelings comparable to human indignation, indicate that a social norm lies at the base of the behavioural regularity. According to von Rohr et al., this kind of negative reaction requires the capacity to exhibit some empathetic competence, a capacity which is, according to them, present in apes such as chimpanzees, and not in monkeys and other species that lack the capacity to attribute mental states to others.[16] They conclude that norms might play a role in guiding the behaviour of chimpanzees, but that these norms are not collective social norms but protosocial norms. Since, in their view, moral norms are collective social norms, the conclusion must be that only humans have a morality.

Although I am inclined to accept that only humans are capable of shared indignation, I am not convinced that shared indignation marks the violation of the kind of social norms we call moral norms. An important step in their argument is the distinction that von Rohr et al. make between personal norms and social norms. A personal norm refers to a personal expectation about how an individual wants to be treated. Personal norms are precursors of social norms because it seems implausible that one would form expectations about how others should be treated before forming expectations about how one wants to be treated oneself. Moral behaviour, they say, starts when personal expectations are generalized and extended to others. It seems that they call norms 'personal' if violation of a norm elicits a negative reaction only from the individual that is negatively affected. I find this concept of a personal norm implausible. If I punish my neighbour when he does not bow to me, because I personally expect that younger persons should bow to older ones, this clearly is a personal norm—provided that I myself also bow to older persons. Nobody else punishes youngsters who do not bow to older persons. The norm is personal because it is not shared by others. Suppose I am talking to some neighbours at the back of my house, when I see a stranger climbing over my fence. I get angry at that person and shout that he has no right to enter my garden without my consent. Although I am the only one who starts shouting, my neighbors—tacitly—approve of my reaction and would do the same if someone climbed over their fence. In the view of von Rohr et al., the norm that one should not climb over a fence protecting another person's house is a personal norm. I disagree. I consider it to be a shared, social norm because all individuals in my group would react negatively when the violation of a norm directly affects them—that is, when someone climbs over their fence. According to von Rohr et al., a norm can be called a social, collective norm only if the indignation about violating that norm in a group is not just simultaneously, but collectively experienced. Thus, the norm forbidding climbing over the fence can only be said to be social if it can be established that my neighbours collectively experience disapproval. How can that be established? Does it suffice that they declare that they share my indignation? Shared indignation among animals cannot be determined by asking them. It can be established only if it results in collective protest against, and condemnation of, the violator. However, von Rohr et al. state that collective experience *may* lead to collective protest. In the absence of clear behavioural expressions it is impossible to prove the presence of collective experience. When von Rohr et al. say, "It is this collectivity upon which the viability and the enforceability of a social norm ultimately rests on and which on current evidence appears to be absent in chimpanzees," it remains unclear what kind of evidence they have in mind.[17] All in all, serious doubts about their conception of a collective social norm are possible. Their argument does not force them to conclude that only humans can have a morality.

4. INTERNALIZATION OF NORMS

Hart's third criterion for distinguishing a statistical behavioural regularity from a rule-based one, as we have seen, is the presence of the internal point of view, the point of view taken by someone who has internalized the norm. Humans are capable of internal guidance by norms. Are animals also capable of internalizing these rules or is their following of rules based only on fear of sanctions? In his paper "Normative Guidance," Peter Railton explores central features of normative guidance, the mental states that underlie it, and its relation to our reasons for feeling and acting, using fictive examples describing everyday activities involving all sorts of norms.[18] He develops in several steps what he calls "a partial largely functional characterization of conditions a piece of behavior must meet to be norm-guided." This characterization applies to all norm-guided behaviour, not only to behaviour guided by moral norms. I jump over these steps, and go right to the last formulation he gives, which I adapt—in his spirit—because I am here only interested in moral norms:

> Agent *A*'s conduct *C* is guided by norm *N* only if 1) *C* is the manifestation of *A*'s disposition to act in a way conducive to compliance with *N*, so that 2) *N* plays a regulative role in *A*'s *C-ing,* where this involves some disposition on *A'*s part 3) to notice failures to comply with *N*, 4) to feel shame or guilt[19] when this occurs, and 5) to exert effort to comply with *N* even when the departure from *N* is unsanctioned and non-consequential.

Condition 1—the disposition to act in a way conducive to compliance with *N*—expresses that "To be norm-guided is a matter of how one is disposed to think, act, and feel, not simply of how one sees oneself, or would like to." Condition 2—*N* plays a regulative role in *A*'s *C-ing*—says that reference to *N* must be a necessary part of the explanation of A's behaviour. Condition 3—the disposition to notice failures to comply with *N*—refers to the fact that A must monitor his or her behaviour because compliance with *N* matters to him or her. That it matters to him explains why he or she takes pains to comply with the norm even if non-compliance doesn't cause a disadvantage to him or her and goes unnoticed by other people (condition 5). The sanctions are internal: feelings of shame and guilt (condition 4).

Railton is not satisfied with a functional characterization of conditions that a certain piece of behaviour must meet to be norm-guided, and goes on to to explore the distinctive role of norm-guidance in an agent's psychology. He wants to know what mental acts or states of mind give a norm this sort of role in one's life. He reviews several candidates that are discussed in recent philosophical literature: acceptance of norms, endorsement of norms, and identifications with norms. None of these attitudes accounts for the role of norms in shaping our lived world and contributing to the reasons for which we act:

> Humble *internalization* of norms without the self's permission, approval, or identification, like humble acquisition of beliefs without the benefit of judgement or reflection, provides much of our substance as agents. And the critical assessment and revision of norms that saves us from mere conformity and inertia, like the critical assessment and revision of what we believe, proceeds more often by trial-and-error feedback and unselfconscious readjustment over the course of experience than by spontaneous higher-order acts of endorsement or self-definition.[20]

To this he adds that these higher-order acts do play a crucial role in making us candidates for moral agency and moral accomplishment. The distinction between humble internalization of norms and higher-order acts of endorsement or self-definition is important for our subject. "Humble internalization" might be the right term to describe the way that animals can be said to possess norms that guide their behaviour. According to Flack & De Waal, one of the distinctive characteristics of humans is that they have a greater degree of norm internalization than non-human primates.[21] They think that this capacity is required for having a morality. They do not explain what this greater degree of norm internalization consists in. Following Railton, we can now interpret them as saying that humans, in contradistinction to animals, are capable of endorsing norms and of self-identifying with norms. By speaking of humans having a greater degree of norm identification, Flack & De Waal suggest that some degree of norm identification is required for having a protomorality. Railton would qualify that as a humble internalization of norms.

In the previous sections, I argued that some animal species follow rules (norms), and that these rules are not external, but, as Railton calls them, humbly internalized. I didn't discuss whether these species are following *moral* rules. In section 5, I examine if animals have the capacities for moral behaviour. Section 6 discusses whether animals can be morally motivated.

5. EMPATHY, CONCERN FOR OTHERS, AND HELPING BEHAVIOUR

Many specialists in animal behaviour suggest that the basic moral competences or capacities are already present in nonhuman primates. One of these competences or capacities is empathy. Bekoff and Pierce even call empathy the cornerstone of what in human society is called morality.[22] As is well-known from the literature in developmental psychology, empathy is not a single behaviour.[23] There is a whole class of behavioural patterns with varying degrees of complexity.[24] Empathy occurs in nested levels, with the inner core as a necessary foundation for the other layers. The simplest forms of empathy are body mimicry and emotional contagion—largely automatic physiological responses. The next layer consists of somewhat more complex behaviours such as emotional empathy and targeted helping. Empathy of the two lowest levels can be found, for example, in mice. More complex is cognitive empathy, the capacity to feel another's emotion and to understand the reasons for it. Cognitive empathy appears to emerge developmentally and phylogenetically with other markers of mind, including perspective taking (PT), mirror self-recognition (MSR), deception, and tool use.[25] According to Preston and de Waal, cognitive empathy may be found in a wider range of species, in the hominoid primates and perhaps elephants, social carnivores, and cetaceans (whales, dolphins, and porpoises). Most complex is the capacity of attribution, in which an individual can take the other's perspective, which requires the use of imagination. According to Koski and Sterck, chimpanzees' capacities to understand others' emotional states operate at the level of what Hoffman calls 'quasi-egocentric empathy'—a complete separation between one's own distress and that of the other has not yet been established.[26] Chimpanzees would also be able to show initial other-regard. There is some evidence, for instance, that chimpanzees can attribute goals.[27] Research also suggests that non-human primates are sensitive to a conspecific's distress signals.[28]

More insight in the role of empathy and concern for the distress of others in human morality is provided by Shaun Nichols.[29] He examined the moral capacities of very young children. Nichols builds on the distinction made by Turiel and his colleagues between conventional and moral rules.[30] Turiel and colleagues contend that moral persons distinguish themselves by regarding the violation of moral rules as special along what they call the dimensions of seriousness, wide applicability, authority, independence, and justification. Violation of moral rules is above all serious when it causes harm to other people. Although the domain of morality is probably wider than that of harm-based violations, Nichols assumes that rules whose violation brings about harm constitute the core of morality. The capacity to see harm-based violations as very serious, generalizable, authority-independent, and wrong, because of well-being considerations, appears, according to Nichols, early in children's ontogenetic development—before their third year—and seems to be cross-culturally universal. Nichols calls this capacity the capacity for Core Moral Judgment (CMJ). CMJ depends on two mechanisms: a normative theory prohibiting harming others and a basic altruistic motivation that is activated by representing suffering in others. In referring to the studies of psychologist Robert Blair,[31] Nichols contends that psychopaths, known to be deficient in affective response to the distress of others, do have a normative theory prohibiting harming others. A striking feature of psychopaths is that they provide conventional-type justifications for why violating moral rules is wrong, rather than offering justifications in terms of harm suffered by the victim. This leads Nichols to the conclusion that the normative theory is at least dissociable from the affective system. As far as I understand, a normative theory is for Nichols simply a system of norms.

Nichols wants to know the cognitive and affective mechanisms underlying altruistic motivation. He argues that altruistic motivation depends on the minimal mind-reading (or empathic) capacity for an enduring representation of pain or some other negative affective or hedonic states in others. Thus, according to Nichols, altruistic motivation does not depend on sophisticated mind-reading capacities. How can attributing distress to others lead to altruistic motivation? Nichols assumes that the altruistic motivation is mediated by an affective response. He gives two accounts of this affect. The available evidence does not weigh in favour of either of these two accounts. The first account is that there is a distinctive basic emotion of sympathy. The other is that distress attribution might produce a kind of second-order contagious distress in the subject. Representing the sorrow of another person may lead one to feel sorrow. This would produce an empathic response—to help, for example. Nichols suggests that perhaps both affective mechanisms are operative. He introduces an overarching term for these two affective mechanisms: Concern Mechanism. Neither reactive distress nor concern require, according to Nichols, sophisticated mind-reading abilities.

Can we extend conclusions from Nichols's findings on the moral capacities of very young children to the moral capacities of animals? Although Nichols doesn't discuss the moral capacities of animals, he thinks it possible that at least some nonhuman animals have the minimal mind-reading capacity to attribute distress to another.[32] He notes that it is unclear from the available data which mechanism is operative in nonhuman primates—whether it is a form of concern or reactive distress.[33] As we have seen above, eminent scholars of animal behaviour think that at least some animals are able to attribute distress to others. However, they disagree whether this mind-reading capacity is required for following norms. Kristen Andrews, for example, argues that animals such as chimpanzees are capable of following norms and punishing violations *without* mind-reading.[34] She thinks that norms can exist prior to understanding others' beliefs and pro-attitudes. Andrews doesn't distinguish between moral norms and other norms. She just assumes that moral norms are among the norms that can be understood without a theory of mind.[35] Although Andrews may be right that norms in animal behaviour can be understood without a theory of mind, I doubt that the same applies to moral norms. In Nichols's theory, the capacity to attribute distress and to be motivated by the perception of distress is central to the moral capacities of very young children and, possibly, also to those of nonhuman primates. Maybe mind reading is not required for following norms, but it might be required for following *moral* norms.

6. MORAL SUBJECTHOOD AND MORAL MOTIVATION

In the previous sections, I made an attempt to clarify that the behaviour of at least some animal species is guided by a morality. In the previous section, I argued that it is possible to assume that at least some nonhuman animals possess basic moral competences and capacities. In section 4, I examined how norms can be present in the minds of animals and how they can guide animal behaviour. This section discusses whether animals can have agency.

In section 4, I referred to Railton's account of the distinctive role of norm guidance in an agent's psychology. However, his statements on guidance by norms and agency seem to be contradictory. First, he says that humble internalization of norms provides much of our substance as agents. Later on, he states that higher-order acts of endorsement of norms and self-definition make us candidates for agency and moral accomplishment. It is not clear whether or not Railton would call norm-following animals "agents." For many of those who keep animals, work with them, or study them, it is quite obvious that animals of many species have agency. To quote the historian Jason Hribal:

> Faking ignorance, rejection of commands, the slowdown, foot-dragging, no work without adequate food, refusal to work in the heat of the day, taking breaks without permission, rejection of overtime, vocal complaints, open pilfering, secret pilfering, rebuffing new tasks, false compliance, breaking equipment, escape, and direct confrontation, these are all actions of what the anthropologist James C. Scott has termed 'weapons of the weak'.[36] Hence, while rarely organized in their conception or performance, these actions were nevertheless quite active in their confrontation and occasionally successful in their desired effects.[37]

Is it really justified to interpret the behaviour of these animals, like Hribal does, as "*acts* of resistance"? Donald Davidson would probably admit that we often succeed in explaining, and sometimes predicting, the behaviour of non-linguistic animals by attributing beliefs and desires and intentions to them.[38] This method, Davidson says, works for dogs and frogs, much as it does for people. Moreover, we have no practical alternative framework for explaining animal behaviour. Davidson thinks that we are justified in applying this method, provided that we acknowledge that we are applying a pattern of explanation that is "far stronger than the observed behavior requires, and to which the observed behavior is not subtle enough to give point."[39] Contrary to humans, dogs and frogs are not rational, and they have no intentional agency. In his view, intentional agency is connected to the capacity to have propositional attitudes—beliefs, intentions, and desires—and to attribute them to others. And this capacity requires language. This is a highly distinctive concept of agency, but is it the only one? Referring to studies of developmental psychologists, Helen Steward states that there is much evidence supporting the view that a basic conception of purposive agency is in place, prior to the emergence of full-scale propositional attitude psychology. In her view, propositional-attitude psychology is a rather sophisticated outgrowth of the basic concept of agency—an outgrowth that is particularly suited to enable us to deal with our human conspecifics. The concept of agent is a more general and less demanding notion.[40]

According to Nichols, very young children are moral before they are capable of having propositional attitudes and ascribing them to others—that is, prior to their having full agency. Being moral doesn't require full agency, only agency in

Steward's less demanding sense. This is also the view of Bekoff and Pierce.[41] They argue that humans are not the only moral beings; at least some animal species also have a morality. [42] Bekoff and Pierce accept what they regard as the philosophical implications of their position: one cannot argue that animals have a morality while denying that they have agency. To that they add that nonhuman animals are not moral agents *in the same sense* in which most adult humans are. Moral agency is species-specific and context-specific; animals are moral agents *within the limited context of their own community*.[43] Unfortunately, Bekoff and Pierce do not tell us what their conception of animal agency is. To say that moral animals are not moral agents in the same sense as human adults is not very informative. Neither do they discuss whether animal agency implies that moral animals are responsible for their behaviour. Adherents of the Davidsonian conception might take the stance that responsibility requires what Steward terms *full agency*: the ability to have and attribute propositional attitudes. Animals clearly lack that ability. But how should we conceive of moral animals if they are not agents? What is the mechanism that makes them act morally when this cannot be explained by referring to the 'mechanism' of reflective capacities?

Mark Rowlands argues in his book *Can Animals Be Moral?* that animals can be morally motivated although they lack moral agency. Moral animals are moral *subjects*:

> X is a moral *subject* if and only if X is, at least sometimes, motivated to act by moral considerations.[44]

The notion of a moral subject has, according to Rowlands, almost invariably conflated with that of a moral *agent*:

> X is a moral *agent* if and only if X is (a) morally responsible for, and so can be (b) morally evaluated (praised or blamed, broadly understood) for, its motives and actions.[45]

In Rowlands's view, moral agency and moral subjecthood should be as conceptually distinct as the concept of motivation is distinct from the concept of evaluation. The main issue that troubles him in his book is that there are persuasive reasons to think that the distinction between motivation and evaluation is not applicable in the moral case.[46] The standard view, Rowlands says, is that an individual's action can only be morally motivated if he or she is conscious of the motivating reason and has control over it.[47] In this view, an individual can only be said to act morally if he or she is not only doing the right thing, but also for the right reason. If the standard view is correct, the distinction between a moral subject and a moral agent collapses. Moral agency is then a condition for moral subjecthood. To have moral normativity, a reason must be under the control of the acting individual. Rowlands sets himself to the task of showing that this concept of control is empty. He builds his argument by introducing the figure of an individual whom he calls "Myshkin"—after the prince in Dostoyevsky's *The Idiot*:

> Prima facie, Myshkin has the soul of a prince. Throughout his life he performs many acts that seem, to the impartial bystander, to be kind or compassionate. Moreover, he performs these acts because he is the subject of sentiments that—again, at least prima facie—seem to be kind or compassionate ones. When he sees another suffering, he feels sad and compelled to act to end or ameliorate that suffering. When he sees another happy, he feels happy because of what he sees. If he can help an individual get what he or she wants without hurting anyone else, he will help because he finds that he enjoys doing it. In short, Myshkin deplores the suffering of others and rejoices in their happiness. His actions reflect, and are caused by, these sentiments. Thus, Myshkin is, or at least seems to be, motivated by sentiments where these are understood as states individuated by their content. …What Myshkin does not do, however, is subject his sentiments and actions to critical moral scrutiny. Thus, he does not ever think to himself things like: "Is what I am feeling the right feeling in the current situation—that is, is what I am feeling what I *should* be feeling?" Nor does he think to himself things like: "Is what I propose to do in this circumstance the (morally) correct thing to do (all things considered)?."[48]

Rowlands does not think that Myshkin is incapable of reflection, but he supposes that his dealings with others operate on a more visceral level. This is the picture that Rowland gives us of Myshkin:

> (1) Myshkin performs actions that seem to be morally good, (2) Myshkin's motivation for performing these actions consists in feelings or sentiments that seem to be morally good, (3) Myshkin is able to subject neither the actions nor the sentiments to critical scrutiny.[49]

Since Myshkin does not reflect on his motivations, and is thus unable to articulate his reasons for action, we cannot know whether he is doing the right thing for the right reason. According to the standard view, we should thus conclude that Myshkin does not act morally. As a part of his attempt to avoid this conclusion, Rowlands introduces the figure of Marlow.[50] Marlow is a moral agent who understands his actions and their consequences. He has both knowledge that a given course of actions is wrong, and also why it is wrong. Marlow has access to the operations of his moral module, and is capable of critically evaluating his motivations. He is an ideal spectator and adjudicator of moral matters. Being a moral agent, Marlow is responsible for what he does and is a legitimate target of moral praise and blame. Suppose, says Rowlands, that Marlow arrives at the conclusion that the sentiments that drive Myshkin's actions are the morally correct ones, and that his actions are the morally right ones. The implication is that we don't have to continue saying that Myshkin *seems* to act morally good and *seems* to have the morally correct sentiments. Would we then still say that Myshkin does not act morally? It still could be that Myshkin's action is accidently morally right and morally correctly motivated. Rowlands's answer is that Myshkin is motivated by moral reasons that are internal, but not available to his conscious, rational scrutiny. They are embodied in his nonconscious, subpersonal processing operations that are cognitively impenetrable, which means that the operations of the moral module cannot be penetrated by, and so are not available to, subsequent belief- and concept-forming operations.[50] In the traditional view, motivating moral reasons must be not only internal, but also under the subject's control. Rowlands's solution—which I will not reconstruct in detail for reasons of space—is to provide us with an alternative account of agency that relies not on the concept of control, but on that of understanding. Marlow, who has knowledge that a given course of actions is wrong, and also knowledge of why it is wrong, and who is capable of critically evaluating his motivations and the principles underlying his actions, has a level of understanding of actions that moral animals are lacking: "What demarcates moral subjects from moral agents, it seems, is a kind of level of understanding."[52]

Rowlands's view that (animal) moral behaviour can be motivated by reasons resulting from unconscious, automatic processes finds support in present-day cognitive science and social psychology. Most psychologists nowadays agree that there are two types of cognitive processes or reasoning systems. Roughly, one system is associative and its computations reflect similarity and temporal structure; the other system is symbolic, and its computations reflect a rule structure.[53] Stanovich and West labeled these systems or types of processes "System I" and "System II."[54] There is now considerable agreement on the characteristics that distinguish the two systems. The operations of System I are fast, automatic, effortless, associative, and difficult to control or to modify. System I is cognitively impenetrable. The operations of System II are slower, serial, effortful, and deliberately controlled; they are also relatively flexible and potentially rule governed. The perceptual system and the intuitive operations of System I generate *impressions* of the attribute of objects of perception and thought. System II is uniquely human. Recent studies show that most of human judgments are not simply the outcome of conscious—System II—reasoning. To a large extent, they are intuitive and automatic—System I—responses to challenges, elicited without awareness of underlying mental processes.[55]

Automaticity is responsible for a large part of our judgments as well as for that part of our behaviour that we characterize as habitual. Habitual moral behaviour engages only System I processes. So does animal morality. It seems plausible to assume that habitual human morality, making use only of System I processes, guides human behaviour in the same manner as animal morality guides animal behaviour.[56] If we think that moral animals, not being agents, cannot held responsible for their automatic and unconscious morally (in)correct behaviour, what does that imply for the responsibility humans can be said to have for their automatic and unconscious moral behaviour? Reversely, if humans are held responsible for their habitual moral behaviour, why not moral animals?

7. RESPONSIBILITY AND HABITUAL BEHAVIOUR

This leads us to the question of what the conditions are for attributing responsibility to an individual for his or here behaviour. One of the most influential theories of responsibility is that of John Martin Fischer and Mark Ravizza.[57] Fischer & Ravizza explain responsibility in terms of (moderate) reasons-responsiveness. For them, responsibility involves a kind of control (guidance control) and this control does not require alternative possibilities (regulative control). Someone "exhibits guidance control of an action insofar as the mechanism that actually issues in the action is his own, reasons-responsive mechanism."[58] Guidance control requires that the actual mechanism be such that it would respond differently in the presence of different reasons.

What do Fischer & Ravizza mean by reasons responsiveness? The distinction between reason tracking and reason responding might be helpful here.[59] Creatures, says Karen Jones, might track reasons and respond to reasons. Reason trackers are capable of registering reasons and behaving in accordance with them. They need not possess the concept of a reason or a self-concept. Nonhuman animals may be seen as reason trackers. When a bird flees after hearing the warning signal of a member of its species, it registers a reason to flee and behaves accordingly. Jones assigns the function of tracking reasons to emotions and the affective systems. It is the intuitive system that tracks reasons. Contrary to reason trackers, reason responders are capable of deliberative reasoning. They can guide their actions via reasons understood as reasons. According to Jones, persons are both reason trackers and reason responders.

Fischer & Ravizza state that *moral* responsibility ought to be characterized not merely as a responsiveness to reasons, but rather as a responsiveness to a range of reasons that include *moral* reasons.[60] Young children act often on processes of thought that are reasons responsive, insofar as their ability to reason practically would have led them to do otherwise in response to some other sufficient reason to do otherwise (e.g., a threat of punishment). Still, we usually do not hold children *morally* accountable because they lack the ability to grasp and respond to specifically *moral* reasons.[61]

(Moral) reasons-responsiveness requires conscious, practical deliberation. If being able to engage in practical deliberation is a condition for attributing responsibility, animals cannot be held responsible for their behaviour since this abilityis clearly absent in all non-linguistic animals. However, Fischer & Ravizza distinguish between two reasons-responsive mechanisms: practical reason and non-deliberative habit. Non-deliberative, habitual actions are also reasons responsive. Reflection and deliberation are, they say, not the only reasons-responsive mechanisms.[62] If they are right, we still have guidance control over habitual actions. How should we conceive this guidance control?

Before going into this question, I need to clarify the concept of habitual action. According to psychologists, habits are represented in memory as direct context-response associations that develop from repeated co-activation of the context and response. These responses are not mediated by representations of goals.[63] However, goals can guide habits by providing the initial outcome-oriented impetus for response repetition. These habits are often a vestige of past goal pursuit. Goals might direct habit learning when people repeatedly implement goals to respond to a particular context cue (e.g. skill learning) as well as when they repeatedly implement goals to respond that do not specify contexts (e.g., implicit learning).[64] When we deliberate, says Bill Pollard, we have direct control over our actions. The kind of control we have over habitual actions is indirect. We have the capacity to intervene in our habitual behaviour:

> Since there was a time when we didn't do such things, it will normally still be possible for us to refrain from doing them in particular cases (though perhaps not in general). We intervene by doing something else, or nothing at all, either during the behavior, or by anticipating before we begin it. In this way habitual behaviors contrast with other automatic, repeated behaviors such as reflexes, the digestion, and even some addictions and phobias in which we cannot always intervene, though we may have very good reason to do so.[65]

Pollard calls this kind of control "intervention control." In Pollard's view, having intervention control over a piece of behaviour is sufficient for someone to be responsible for that behaviour. Humans intervene in their habitual behaviour if they have reasons to do so. These reasons may be external or internal. An external reason is, for example, a prohibition by a superior entity, or a request by someone who is annoyed by that habit. An internal reason is, for example, the insight that the reason underlying the habit is no longer valid. Imagine John, who, having discovered a quicker route to his place of work, invariably goes that way. This route is so ingrained in his habits that he sometimes even takes it when, on his way to some other destination, he is deep in thought. At a party at his children's school, he meets Gerard, who works in an office in the vicinity of his work. Gerard knows a still faster route. If John decides to adapt his usual route, he shows intervention control. The talk with Gerard made him reflect on his habitual route. The reason behind taking the old route—that it is the fastest one—is no longer valid.

In the language of psychologists, Pollard says that new information might lead an individual to re-represent the goals that initially provided the outcome-oriented impetus for the repetitive behaviour. However, habits tend to persist, even when they no longer align with the initial goals. Suppose that Myshkin, compassionate as he is, has a habit of giving generous alms to beggars. He learns that beggar Pjotr has gained a small fortune by means of his begging activities. Pjotr repeatedly succeeds in getting alms from the same people by disguising himself in different guises. No recognizable human being, even someone whose actions are usually automatic and habitual, would in the long run continue giving alms to a beggar once becoming aware of the fact that the beggar is far from poor. He will, perhaps slowly, realize that the reason underlying his alms-giving habit doesn't apply to this beggar. However, Rowlands's unreflective Myshkin would go on giving Pjotr alms because the image of a begging Pjotr still provides the cues that trigger his habit of giving alms. Unreflective Myshkin lacks the capacities for deliberate intervening.

Animals are capable of inhibiting habitual behaviour, which depends on having the ability of self-control.[66] However, the kind of intervention that Pollard has in mind is deliberate, reasons-based intervention. Self-control may also be a necessary condition for deliberate intervention, but certainly not a sufficient one. Having the ability to reflect is a necessary condition for *deliberate* intervention in habitual behaviour. Lacking that ability, moral animals are unable to intervene deliberately in their habitual behaviour. Therefore, the conclusion that even moral animals are not morally responsible for their behaviour is unavoidable.

CONCLUSION

The aim of this article was to examine if we can attribute moral responsibility to animals. The intuitive answer of most people will be that we cannot do so. However, many people also find that no animal species can have a morality. Given that this belief is highly contested, there is also reason to take a fresh look at the issue of animal responsibility. This is what I attempted to do in this article. Although the final answer is still the same—we cannot hold animals responsible for what they do—it became clear on our way to that answer that many still widely shared beliefs about animals have to be revised: some animal species do have a morality, and these species' members are capable of being motivated by moral reasons. I am convinced that these revised beliefs also affect our view on the moral status of moral animals, but I have to leave this issue for another occasion.

NOTES

[1] Singer, Peter, *Animal Liberation,* New York, Random House, 1975.
[2] Warnock, Geoffrey J., *The object of morality.* London, Methuen, 1971, p. 26.
[3] Flack, Jessica C. and Frans de Waal, "'Any Animal Whatever'. Darwinian Building Blocks of Morality in Monkeys and Apes," *Journal of Consciousness Studies*, vol. 7, no. 1-2, 2002, p. 23.
[4] Peterson, Dale, *The Moral Lives of Animals*, New York, Bloomsbury Press, 2011, p. 58.
[5] Bekoff, Mark and Jessica Pierce, *Wild Justice. The Moral Lives of Animals,* Chicago/London, The University of Chicago Press, 2009, p. 7.
[6] *Ibid.*, p. 13.
[7] *Ibid.*, p. 20.
[8] *Ibid.*, p. 139.
[9] Hart, Herbert L. A., *The Concept of Law,* Oxford, Oxford University Press, 1961, p. 153.
[10] Von Rohr, Claudia R., Judith M. Burkart and Carel P. Van Schaik, "Evolutionary Precursors of Social Norms in Chimpanzees: A New Approach," *Biology and Philosophy*, vol. 26, no. 1, 2011, pp. 1-30.
[11] *Ibid.*, p. 3ff.
[12] *Ibid.*, p. 16.
[13] *Ibid.*, p. 16.
[14] *Ibid.*, p. 17.
[15] Tomasello, Michael and Malinda Carpenter, "Shared Intentionality," *Development Science*, vol. 10, no. 1, 2007, pp. 121-125.
[16] Tomasello, Michael and Josep Call, *Primate Cognition,* New York, Oxford University Press, 1997.
[17] Von Rohr et al., "Evolutionary Precursors of Social Norms in Chimpanzees: A New Approach," *op. cit.*, p. 18.
[18] Railton, Peter, "Normative Guidance," *in* Russ Shafer-Landau (ed.), *Oxford Studies in Metaethics*, Volume 1, Oxford, Oxford University Press, 2006, pp. 3-35.
[19] Railton speaks here of discomfort, not of shame and guilt, which are more specific moral feelings.
[20] Railton, "Normative Guidance," *op. cit.*, p. 31f.
[21] Flack and De Waal, "'Any Animal Whatever'. Darwinian Building Blocks of Morality in Monkeys and Apes," *op. cit.*.
[22] Bekoff and Pierce, *Wild Justice, op. cit.*, p. 87.
[23] See Eisenberg, Nancy, "Emotion, Regulation, and Moral Development," *Annual Review of Psychology*, vol. 51, no. 1, 2000, pp. 665-697; Hoffman, Martin L., *Empathy and Moral Development: Implications for Caring and Justice,* Cambridge, Cambridge University Press, 2000.
[24] See Preston, Stephanie D. and Frans De Waal, "Empathy. Its Ultimate and Proximate Bases," *Behavioural & Brain Sciences*, vol. 25, no. 1, 2002, pp. 1-20; Bekoff and Pierce, *Wild Justice, op. cit.*.
[25] Preston and De Waal, "Empathy. Its Ultimate and Proximate Bases" *op. cit.*.
[26] Koski Sonja Elena and Elizabeth H. M. Sterck, "Empathetic Chimpanzees. A Proposal of the Levels of Emotional and Cognitive Processing in Chimpanzee Empathy," *European Journal for Development Psychology*, vol. 7, no. 1, 2009, pp. 38-66.
[27] See Premack, David and Guy Woodruff, "Does the Chimpanzee Have a Theory of Mind?," *Behavioral and Brain Sciences*, vol. 1, no. 4, 1978, pp. 516-526; Call, Josep and Michael Tomasello, "Distinguishing Intentional from Accidental Actions in Orangutans, Chimpanzees, and Human Children," *Journal of Comparative Psychology*, vol. 112, no. 2, 1998, pp. 192-206.

[28] E.g., Miller, Robert, James Banks and Nobuya Ogawa, “Role of Facial Expression in ‘Cooperative-Avoidance Conditioning’ in Monkeys,” *Journal of Abnormal and Social Psychology*, vol. 67, no. 1, 1963, pp. 24-30.

[29] Nichols, Shaun, *The Sentimental Rules,* Oxford, Oxford University Press, 2004.

[30] Turiel, Elliot, *The Development of Social Knowledge: Morality and Convention,* Cambridge, Cambridge University Press, 1983; Turiel, Elliot, Melanie Killen and Charles Helwig, “Morality: Its Structure, Functions, and Vagaries,” *in* Jerome Kagan and Sharon Lamb (eds), *The Emergence of Morality in Young Children,* Chicago, University of Chicago Press, 1987, pp. 155-244.

[31] Blair, R. James R., “Cognitive Developmental Approach to Morality: Investigating the Psychopath,” *Cognition*, vol. 57, no. 1, 1995, pp. 1-39; Blair, R. James R., “Moral Reasoning and the Child with Psychopathic Tendencies,” *Personality and Individual Differences,* vol. 22, no. 5, 1997, pp. 731-739.

[32] Nichols, *The Sentimental Rules, op. cit.*, p. 60.

[33] *Ibid.*, p. 61.

[34] Andrews, Kristin, “Understanding Norms without a Theory of Mind,” *Inquiry*, vol. 52, no. 5, 2009, pp. 433-448.

[35] *Ibid.*, p. 446.

[36] Scott, James C., *Weapons of the Weak. Everyday Forms of Peasant Resistance,* New Haven, Yale University Press, 1987.

[37] Hribal, Jason C., “Animals, Agency, and Class: Writing the History of Animals from Below,” *Human Ecology Forum*, vol. 14, no. 1, 2007, p. 103.

[38] Davidson, Donald, “Rational Animals,” *Dialectica*, vol. 36, no. 4, 1982, pp. 317-327.

[39] *Ibid.*, p. 32.

[40] Steward, Helen, “Animal Agency”, *Inquiry*, vol. 52, no. 3, 2009, pp. 217-231.
This is Steward’s concept of agency:
(i) an agent can move the whole, or at least some parts, of something we are inclined to think of as *its* body;
(ii) an agent is a centre of some form of subjectivity;
(iii) an agent is something to which at least some rudimentary types of intentional state (e.g., trying, wanting, perceiving) may be properly attributed;
(iv) An agent is a *settler* of matters concerning certain of the movements of its own body i.e. the actions by means of which those movements are effected are considered to be non-necessitated events, attributed always first and foremost to the *agent,* and only secondarily to environmental impacts or triggers of any sort. (Steward, “Animal Agency”, p. 226).

[41] Bekoff and Pierce, *Wild Justice*, *op. cit.*.

[42] Bekoff and Pierce argue that to have a morality, a given species must meet certain threshold requirements. These thresholds are the following: a level of complexity in social organization, including established norms of behaviour to which attach strong emotional and cognitive cues about right and wrong; a certain neural complexity that serves as a foundation for moral emotions and for decision making based on perceptions about the past and the future; relatively advanced cognitive capacities (such as a good memory); and a high level of behavioural flexibility (Bekoff and Pierce, *Wild justice*, *op. cit.*, p. 13). Bekoff and Pierce do not provide us with a list of animal species that meet the threshold requirements. What they do say is that animals with a highly developed moral capacity may include chimpanzees, wolves, elephants, and humans (Bekoff and Pierce, *Wild justice*, *op. cit.*, p. 20).

[43] Bekoff and Pierce, *Wild justice*, *op. cit.*, p. 144.

[44] Rowlands, Mark, *Can Animals Act Morally?*, Oxford, Oxford University Press, 2011, p. 89.

[45] *Ibid.*, p. 75.

[46] *Ibid.*, p. 93.

[47] *Ibid.*

[48] *Ibid.*, p. 124.

[49] *Ibid.*, p. 125

[50] Marlow is "christened after the skilled scrutinizer of motives who narrates some of Joseph Conrad's novels" (*Ibid.*, p. 128).

[51] *Ibid.*, p. 146f.

[52] *Ibid.*, p. 239f.

[53] See Sloman, Steven A., "The Empirical Case for Two Systems of Reasoning," *Psychological Bulletin*, vol. 119, no. 1, 1996, pp. 3-22.

[54] Stanovich, Keith E. and Richard F. West, "Individual Differences in Reasoning: Implications for the Rationality Debate?," *Behavioural and Brain Sciences*, vol. 23, no. 5, 2000, pp. 645-665.

[55] See Bargh, John A., Mark Chen and Lara Burrows, "Automaticity in Social Psychology: Direct Effects of Trait Construct and Stereotype Activation on Action," *Journal of Personality and Social Psychology*, vol. 71, no. 2, 1996, pp. 230-244; Bargh, John A. and Tanya L. Chartrand, "The unbearable automaticity of being," *American Psychologist*, vol. 54, no. 7, 1999, pp. 462-479.

[56] In contradistinction to System II cognitive processes, which are rule-based and computational, System I processes are said to be associative and/or heuristic based. This image is, according to Peter Carruthers, wrong. See Carruthers, Peter, "The fragmentation of Reasoning," *in* Pablo Quintanilla, Carla Mantilla and Paola Cépeda (eds), *Cognición social y lenguaje. La intersubjetividad en la evolución de la especie y en el desarrollo del niño*, Lima, Pontificia Universidad Católica del Perú, 2014. Carruthers refers to research by Gallistel and colleagues on conditioning of animals that shows that the behaviour of animals involved in conditioning experiments is best explained in rule-governed, computational terms, rather than in terms of associative strengths (Gallistel, C. R. and John Gibbon, "Time, Rate and Conditioning," *Psychological Review*, vol. 107, no. 2, 2000, pp. 289-344; Gallistel, C.R. and Adam P. King, *Memory and the Computational Brain*, Oxford, Wiley-Blackwell, 2009). Carruthers concludes that nonhuman animals engage in unreflective processes that can be both flexible and rule-governed. Otherwise learning by animals could not be explained

[57] Fischer, John M. and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility,* Cambridge, Cambridge University Press, 1998.

[58] Fischer and Ravizza, *Responsibility and Control*, *op. cit.*, p. 39.

[59] I derive this distinction from Karen Jones. Jones says that even brain-damaged persons who are unable to form long-term memories can have functioning fear systems that enable affective learning that 'tracks' their practical reasons without generating higher-level understanding of that tracking (Jones, Karen, "Emotion, weakness of will and the normative conception of agency," *in* Anthony Hatzimosyis (ed.), *Philosophy and the Emotions,* Cambridge: Cambridge University Press, 2003, p. 185). Jones refers to Joseph LeDoux (who reports the case of a woman who, though unable to recognize her doctors from one meeting to the next, was able to learn not to shake hands with a doctor who had previously pricked her with a tack concealed in his palm (Ledoux, Joseph, *The emotional brain,* New York, Simon & Schuster, 1996).

[60] Fischer and Ravizza, *Responsibility and Control*, *op. cit.*, p. 81.

[61] See also, Wallace, R. Jay, *Responsibility and the Moral Sentiments*, Cambridge, Harvard University Press, 1994, p. 189.

[62] Fischer and Ravizza, *Responsibility and Control*, *op. cit.*, p. 232ff. It is not clear how Fischer and Ravizza's reasons responsiveness by habitual actions relates to Jones's concept of tracking reasons.

[63] Wood, Wendy and David T. Neal, "A New Look at the Habit–Goal Interference," *Psychological Review*, vol. 114, no. 4, 2007, p. 845.

[64] Wood and Neal, "A New Look at the Habit–Goal Interference," *op. cit.*, p. 851.

[65] Pollard, Bill, "Can virtuous actions be both habitual and rational?," *Ethical Theory and Moral Practice*, vol. 6, no. 4, 2003, pp. 415f.

[66] According to Baumeister et al., the human capacity for self-regulation appears to be much more extensive than what is found in other animals, which may suggest that the evolutionary pressures that guided the selection of traits that make up human nature, such as participation in cultural groups, found self-regulation to be especially adaptive and powerful (Baumeister, Roy F., Matthew C. Gailliot, C. Nathan DeWall and Megan Oaten, "Self-Regulation and Personality: How Interventions Increase Regulatory Success, and How Depletion Moderates the Effects of Traits on Behavior," *Journal of* Personality, vol. 74, no. 6, 2006, pp. 1773-1801).