

EMPATHY AND MORALITY IN BEHAVIOUR READERS

Susana Monsó*

[This paper is forthcoming in *Biology & Philosophy*. The final publication is available at Springer via <http://dx.doi.org/10.1007/s10539-015-9495-x>]

Abstract: It is tempting to assume that being a moral creature requires the capacity to attribute mental states to others, because a creature cannot be moral unless she is capable of comprehending how her actions can have an impact on the well-being of those around her. If this assumption were true, then mere behaviour readers could never qualify as moral, for they are incapable of conceptualising mental states and attributing them to others. In this paper, I argue against such an assumption by discussing the specific case of empathy. I present a characterisation of empathy that would not require an ability to attribute mental states to others, but would nevertheless allow the creature who possessed it to qualify as a moral being. Provided certain conditions are met, a behaviour reader could be motivated to act by this form of empathy, and this means that behaviour readers could be moral. The case for animal morality, I shall argue, is therefore independent of the case for animal mindreading.

Keywords: empathy; morality; moral emotions; nonhuman animals; behaviour reading; mindreading/theory of mind

* Department of Logic, History and Philosophy of Science, Universidad Nacional de Educación a Distancia (UNED), Spain.

1. Introduction

Could a creature engage in moral behaviour if she were incapable of attributing mental states to others? Could she be moved to help or care for others if she lacked an understanding of their mental lives? It is quite tempting to assert that behaving morally requires being able to grasp the fact that other beings, beside oneself, can feel pain and suffer. It also seems *prima facie* plausible that being a moral creature requires having the ability to comprehend how one's actions can affect others' well-being, a feat that is arguably unattainable without the ability to attribute mental states to others. I will argue that, despite appearances, acting for moral reasons does not require such sophisticated intellectual abilities.

My aim in this paper will be to argue that a creature who lacked mindreading capabilities could nevertheless act on the basis of empathy (understood as a moral emotion), and that this would allow her to qualify as a moral subject. The term 'moral subject' will be used, following Rowlands (2012), to refer to a being who behaves on the basis of moral motivations, even if she cannot be held responsible for her behaviour. I will argue that lacking mindreading capabilities is not an impediment to acquiring the status of moral subject, because a being incapable of mindreading could still be motivated to act by empathy as a moral emotion.

The motivation behind the case I want to make is twofold. Firstly, there is a growing amount of evidence that points to the presence of moral behaviours in quite a large number of nonhuman species (for a review, see Bekoff & Pierce, 2010). This contrasts with the current state of the animal mindreading debate, where there is a limited amount of (problematic) evidence that is largely restricted to great apes and corvids (for a review, see Lurz, 2011). Arguing for the independence of both these capacities seems crucial. Secondly, the category of moral subjecthood that was introduced by Rowlands (2012) is, I believe, a rather important one that has been commonly neglected. Separating moral motivation from moral responsibility allows us, amongst other things, to make sense of many of our intuitions that cannot be properly understood from an over-intellectualised vision of morality. For instance, we may feel inclined to describe the behaviour of a child who is a bully, or a mentally ill person who commits a

crime, as *morally bad*, but making the morality of one's motivations or behaviour dependent upon the possibility of being held morally responsible makes this problematic. The category of moral subjecthood carves out a space where behaviour can be moral in the absence of responsibility. One of the central aims of Rowlands' (2012) project was to show that moral subjecthood can obtain while in the absence of metacognitive capacities. My aim is to develop his idea, and show that mindreading is not required for moral subjecthood, either.

In order to make my case, I will assume that nonhuman animals (hereafter 'animals') are incapable of thinking about others' beliefs, desires, perceptions, sensations, and emotions, and that their social cognition is limited to detecting and reading behavioural cues. In making this assumption, I am following the strongest reading of Povinelli and colleagues' work on animal mindreading. I will show how, even if this assumption were true, the case for moral subjecthood in animals would still be open. The term 'mindreading' will thus be used in this paper in the broad sense in which Povinelli and colleagues use it,² that is:

to refer to *any* cognitive system, whether theory-like or not, that predicts or explains the behaviour of another agent by postulating that unobservable inner states particular to the cognitive perspective of that agent causally modulate the agent's behaviour. (Penn & Povinelli, 2007, p. 732)

The ability to mindread, therefore, constitutes the ability to conceptualise mental states and attribute them to others in order to explain or predict their behaviour. Two qualifications must be made at this point. First, although this definition of mindreading only makes reference to the attribution of mental states to others, later in that very paper,³ as well as in other texts (Gallagher & Povinelli, 2012, p. 164), it is made clear that Povinelli and colleagues believe that mindreading is also involved in the attribution of mental states to oneself. A being is thus thought unable to reflect upon her own mental states if she lacks mindreading capabilities. I will

² I will use the term 'mindreading', instead of 'theory of mind', or 'ToM', as Povinelli *et al.* usually call it, in order to avoid the theory theory connotations of the latter term.

³ '[A] subject lacking an f_{ToM} [i.e. mindreading capabilities] would not have access to r-states [i.e. information] about his own internal cognitive states [...]' (Penn and Povinelli 2007, p. 738).

follow Povinelli *et al.* and assume that a behaviour reader cannot reflect upon her own mental states, while noting that this is an assumption that has been questioned (e.g. Gallagher & Povinelli, 2012, p. 163-4) and that my argument is in no way critically dependent upon it.

The second qualification that must be made is that, although Povinelli and colleagues tend to focus on the predictive role of mindreading (due, possibly, to the fact that they are experimental scientists and this aspect of mindreading is probably easier to test for), I am going to be focusing on its explanatory role.⁴ This is so because this is the aspect of mindreading that is more relevant to morality. Indeed, mindreading can be thought to be relevant for morality in two main ways and both of them are related to explanation and understanding, rather than prediction. Firstly (following our previous assumption), mindreading allows one to reflect on one's motivations, an ability that is often thought to be central to moral autonomy (e.g. Korsgaard, 2004, 2006).⁵ Secondly, morality may seem dependent upon an ability to understand others as having a welfare, that is, as possessing mental states that can be affected by our actions. Indeed, the idea of taking into account others' welfare and interests is often thought to be the basis of key moral principles. Understanding that others have a welfare is arguably impossible without the ability to attribute mental states to others.

In order to argue for the possibility of moral subjecthood in behaviour readers, I will, following Rowlands (2012), show that neither reflecting on one's motivations nor understanding others as having a welfare is central to acting for moral reasons. In order to make my case, I will focus on the specific case of empathy as a moral emotion. I will show that, provided certain conditions are met, a behaviour reader could be motivated to act by empathy, where this notion is understood as a moral emotion, and that this means that a behaviour reader could be a moral subject.

2. Three Key Notions: Moral Subjecthood, Behaviour Reading and Empathy

⁴ The importance of mindreading for the prediction of behaviour has been questioned by Andrews (2012). Povinelli *et al.* also acknowledge that, in humans, the attribution of mental states is used more often for explanation than for prediction (Gallagher & Povinelli, 2012, p. 151).

⁵ However, see Andrews (2013) for an account of autonomy that does not require mindreading.

The three key terms I will employ in my argument ('moral subject', 'empathy', and 'behaviour reading') may be subject to confusions, and so this section will be devoted to clarifying what I will (and will not) mean by them. Following that, I will present a conception of empathy that will allow behaviour readers to potentially qualify as moral subjects.

2.1. Rowlands' Third Moral Category: Moral Subjecthood

Rowlands (2012) puts forward the idea of moral subjecthood to account for the evidence that points to seemingly moral behaviour in animals (see Bekoff & Pierce, 2010). This is a moral category that is different from the traditional concepts of moral patienthood and moral agency. To develop this new category, Rowlands begins by noting that the concepts of moral evaluation and moral motivation are distinct, so that we can make sense of the idea of moral motivation without moral evaluation. This means that we should be able to conceptualise a new moral category that pertains only to moral motivation – the category of moral subjecthood (Rowlands 2012, chap. 3).

A moral subject is a being who is motivated to act by moral reasons, but a mere moral subject cannot be praised or blamed for her behaviour. Moral subjecthood is thus distinguished from moral patienthood in that a moral subject is not merely a legitimate object of concern, but also a subject of moral motivations. At the same time, it is distinguished from moral agency in that a mere moral subject lacks a sufficient understanding of her motivations and subsequent behaviour to be held responsible for whatever she does.⁶ Rowlands gives the following definition of a moral subject:

X is a *moral subject* if X possesses (1) a sensitivity to the good- or bad-making features of situations, where (2) this sensitivity can be normatively assessed, and (3) is grounded in the operations of a reliable mechanism (a "moral module"). (Rowlands, 2012, p. 230)

⁶ In Rowlands' scheme, moral agency is gained through understanding and not through control. See section 4.

These are sufficient conditions for a being to qualify as a moral subject, but they are not necessary conditions. There may be other ways of conceiving moral subjecthood.

This definition requires us to understand as an objective moral fact that there are certain features of situations that make those situations good or bad ones, but we don't need the moral subject to be able to *judge that* they are good or bad, nor even that she be able to judge them *as* good or bad. We just need her to be able to detect some of these features and react in a certain sort of way towards them. Her reaction should be grounded in the operations of a certain mechanism, in order to qualify as reliable. Rowlands calls this a 'moral module', although with the qualification that this is just a label and that he is not committing himself to any realistic conception of the word 'module'. The purpose of this idea is to add the condition that the being's reaction to the good- or bad-making features of situations she is sensitive to should not be merely accidental or contingent, but that there should be some sort of mechanism behind it that guarantees that it always occurs, provided certain circumstances obtain (Rowlands 2012, chap. 5).

Within the moral reasons that can motivate a moral subject's behaviour, we can find morally laden emotions. There may be other moral reasons, such as perhaps considerations of a rational and reflexive character, but the type of moral reasons that Rowlands is interested in is that which may be applicable to animals. If there is a class of moral reasons that they may be motivated by, the best candidate are morally laden emotions (Rowlands 2012, p. 231).⁷

Rowlands defines morally laden emotions in the following way:

An emotion, *E*, is *morally laden* if and only if (1) it is an emotion in the intentional, content-involving, sense, (2) there exists a proposition, *p*, which expresses a moral claim, and (3) if *E* is not misguided, then *p* is true. (Rowlands, 2012, p. 69)

Let us illustrate this with an example. Indignation, in Rowlands' framework, counts as a moral emotion. This is so because, firstly, it involves factual content. It does not make sense to

⁷ Within Rowlands' scheme, morally laden emotions qualify as reasons, and are not mere causes, because they are emotions that involve content (see section 3.2).

say that one is indignant about nothing in particular. To be an actual case of indignation, one must be indignant about something. Borrowing Rowlands' example, we can think of the case of Smith, who is indignant *that* Jones snubbed him. For Smith's indignation to count as a moral emotion, it must not only involve factual content, but also morally evaluative content. Such content would be the proposition "It was wrong for Jones to snub me." Crucially, however, this second proposition need not be entertained by Smith. It is sufficient for his emotion to *track* such proposition.

In order to understand what it means for Smith's emotion to track the proposition "It was wrong for Jones to snub me," it is useful to look at the errors that can be implicated in Smith's indignation. These can be of two types, depending on whether they affect the factual content of his indignation, or the evaluative content. Smith's indignation is, firstly, based on the factual assertion "Jones snubbed me." That is, Smith believes that Jones snubbed him. If Jones did not, in fact, snub Smith, but merely failed to notice his presence as they crossed paths in the busy lobby, then Smith's indignation would be *misplaced*, for it would be based on a false belief. The error may instead affect the evaluative content of Smith's emotion, in those cases in which it was not, in fact, wrong for Jones to snub him. Say Jones had overheard Smith criticising him behind his back. Smith would then have no right to be indignant that Jones snubbed him. His indignation would be *misguided*, because it would be based on the false evaluative claim "It was wrong for Jones to snub me." What it means for Smith's emotion to track this evaluative proposition is, therefore, that, regardless of whether Smith actually entertains it, there is a truth-preserving relation between Smith's emotion and the proposition itself, such that the non-misguided status of Smith's indignation is what guarantees its truth. Indeed, if his indignation is not misguided, then this proposition *must* be true.

Rowlands (2012) only makes reference to the tracking of moral propositions on behalf of emotions. We can, however, also make sense of the idea of a behavioural response being motivated by an emotion and, by extension, tracking the same proposition that the emotion tracks. So, for instance, suppose that Smith's indignation causes him not to invite Jones to his birthday party. Such a behavioural response would also be misguided if Smith had originally

deserved Jones' snubbing. The truth of the proposition "It was wrong for Jones to snub me" thus also seems to be guaranteed by the non-misguided status of Smith's behavioural reaction. This idea will be relevant in upcoming sections. Apart from this small qualification, I shall adopt the conceptual framework offered by Rowlands to make my case.

As a side note before moving on, it is worth mentioning that Rowlands seems to be committed to a view that some find controversial – moral realism (see Rowlands, 2012, pp. 222–5). Although I am myself inclined towards this metaethical position, attempting a full-blown defence of moral realism is clearly beyond the scope of this paper, so I will limit myself to the following clarifications. On the one hand, it is important to note that, for the arguments to be developed in this paper, all that is required is the assumption that certain features of the world contribute to the goodness or badness of a situation, but not the commitment to any particular theory in normative ethics. The model of minimal empathy I will develop, in turn, assumes that it is *not unreasonable* to suppose that certain mental states such as pain and suffering are, *ceteris paribus*, bad-making features of situations. This does not seem an extravagant assumption to make, if we take into account both the phenomenological characteristics of such mental states (which are, at least arguably, inextricably bound up with aversion or unpleasantness) and the general consensus, both within moral philosophy and in our folk intuitions, with regards to moral claims such as "Other things being equal, suffering is bad." On the other hand, however, Rowlands' account of morally laden emotions may well be compatible with less stringent metaethical positions than moral realism. All that his account needs is a way of making sense of the truth of certain moral claims, so as to speak of moral emotions tracking them. Some anti-realist stances, such as Blackburn's quasi-realism (Blackburn, 1993), allow us to talk about the truth of moral claims, while remaining skeptical about the ontological status of moral properties. During the remaining paper, I shall bracket this issue and assume the plausibility of the realist stance, but there are no reasons to think that my arguments cannot be reinterpreted to fit the quasi-realist standpoint. Let us now turn to the notion of behaviour reading.

2.2. Behaviour Reading versus Mindreading

Research on animal mindreading aims to determine whether animals are capable of attributing mental states to others in order to predict or explain their behaviour. It is usually understood that mental states are unobservable entities with causal power over one's behaviour and they are thought to include sensations, perceptions, emotions, intentions, beliefs, and desires. There is a relatively widespread consensus with respect to the idea that *some* animals are capable of attributing *some* mental states to others: in particular, perceptual states such as 'seeing' and 'hearing', intentions and states of ignorance/knowledge. It is generally thought to be the case that animals cannot, however, attribute more complex states, such as false beliefs (for a review of this research, see Call & Tomasello, 2008; Lurz, 2011, chap. 1).

Some researchers, however, remain altogether sceptical with respect to the case for animal mindreading. Among these, the most notable exception is that of Povinelli and colleagues (Povinelli and Vonk 2003; Povinelli and Vonk 2004; Penn and Povinelli 2007; Penn, Holyoak, and Povinelli 2008; Gallagher and Povinelli 2012; Penn and Povinelli 2013). They, together with other sceptics such as Heyes (1998) and Lurz (2009; 2011), formulated the idea of a behaviour reader: a being who can detect behavioural cues, conceptualise them employing abstract terms, and use them to predict (or explain) the behaviour of others.

The idea of a behaviour reader has been used on behalf of these sceptics to argue that current experimental paradigms in animal mindreading research face an insurmountable 'logical problem'. The argument goes as follows:

P1: Other agents' mental states cannot be subjected to direct observation.

P2: Mental state attribution is always based on behavioural or environmental cues.

P3: Current experimental paradigms cannot, even in principle, distinguish a behaviour reading response from its mindreading counterpart, since a task that would only be achievable by means of a mental state attribution is never required of the experimental subjects.

P4: Current experimental results can be more parsimoniously explained by positing that animals are behaviour readers, since this is less cognitively demanding and nothing is gained from positing that they are mindreaders.

C1: There is currently no evidence whatsoever that animals can mindread.

(C2: Animals cannot mindread.)

Conclusion 2 is sometimes asserted by Povinelli and colleagues (Penn *et al.*, 2008, p. 129; Penn & Povinelli, 2013, p. 10), even though it does not strictly follow from their argument. Nevertheless, in what remains, I will assume that Povinelli and colleagues are right in their critique of animal mindreading research and, furthermore, that conclusion 2 is true and animals cannot attribute mental states to others. I will argue that, even assuming that Povinelli *et al.* are right,⁸ the case for empathy as a moral emotion in animals (and the broader case for moral subjecthood in them) still stands. In order to make my case, I am going to employ their notion of a behaviour reader. It is important, then, for my purposes, to clarify exactly what is meant by this term. Povinelli and Vonk (2004) define it as:

a psychological system dedicated to social cognition, but one which forms and uses concepts about only 'behaviors' which can, in principle, be observed. (Povinelli and Vonk 2004, p. 4)

A behaviour reader thus engages in social cognition, which means that it predicts and explains the behaviour of others. The crucial idea is that, in doing so, it can only make use of behavioural concepts formed by interpreting observable behavioural cues. As an example of the formation and utilisation of a behavioural concept, we can find the following:

chimpanzees, confronted with the particular behavior of another chimp (e.g., pursing lips, bristling hair, etc.) are able to represent this behavior in terms of a more abstract interpretation (e.g. threat [...]). Seeing a 'threat display', the chimpanzee likely has a sense of the kinds of

⁸ In making this assumption, I am bracketing several important critiques of Povinelli and colleagues' approach made on both theoretical and empirical grounds, for instance, Andrews, 2005; Buckner, 2014.

behavior that will follow ('charging', 'being hit', etc.). We have referred to this as the 'behavioral abstraction hypothesis'. It holds that chimpanzees: (a) construct abstract categories of behavior, (b) predict future behaviors following from past behaviors, and (c) adjust their own behavior accordingly. (Gallagher & Povinelli, 2012, p. 150)

Behaviour readers are thus not 'behaviouristic' beings, but rather:

fully "cognitive creatures" endowed with mental representations and inferential abilities similar to those of humans but [their] representational capabilities might not encompass *all* the same semantic possibilities as human subjects. (Penn & Povinelli, 2013, p. 74)

What is meant by this last claim is, precisely, that what a behaviour reader lacks is the ability to conceptualise and reason about unobservable entities, and these include mental states. The question of whether a behaviour reader can be a moral subject thus amounts to the question of whether a moral subject needs to be able to understand and reason about others' mental states. As stated in the introduction, I am going to answer this question by looking at the particular case of empathy. If we can make a case for a behaviour reader being motivated by empathy as a moral emotion, then we can state that there are at least some grounds for arguing that moral subjecthood does not require mindreading capabilities. Before going on to presenting the conception of empathy that can achieve this, it is important to first look at how this term is used in the literature.

2.3. Empathy and its Relation to Mindreading

Most papers that aim to deal with a problem related to empathy usually begin by giving a definition of the term, for there is no consensus on what this term denotes. This has resulted in a situation in which '[t]here are probably as many definitions of empathy as people working on the topic' (de Vignemont & Singer, 2006, p. 435). This disagreement stems, at least in part, from the fact that the notion of empathy plays an important role in two significantly different research strands (see Batson, 2009). The first of these research strands studies social cognition, and here

empathy is thought to have something to do with how people come to understand what others think and feel. The second of these research strands studies moral psychology or moral development, and here empathy is often thought to be a key ability whose absence can result in social dysfunctions such as psychopathy. Within both of these different research strands, very disparate notions of empathy can be found, and each may bear a different relationship to mindreading.

Within the field of social cognition, three distinct relations between empathy and mindreading can be found. Firstly, empathy is at times used as a synonym for mindreading. This use of the term can be found, for instance, in Povinelli (1998). Following Michael (2014), we can distinguish two further relations between the notions of empathy and mindreading in this research strand. On the one hand, empathy is at times conceived as a mechanism, process or ability that *enables* the attribution of mental states to others. On the stronger versions of the simulation theory of mind, for instance, one cannot attribute mental states to others if one lacks the ability to empathise, understood as a process of mental simulation whereby one puts oneself in the other's 'mental shoes' (e.g. Goldman, 2006). On the other hand, empathy is at times conceived as a complex ability that serves the main purpose of understanding others and that *presupposes* the ability to mindread. Within this last type of framework, if one lacks the capacity to attribute mental states to others, one cannot achieve an understanding of others by empathising with them (e.g. de Vignemont & Jacob, 2012).

The notions of empathy that are relevant for our present purposes are, however, those that can be found within the second research strand. It is here where we find the notions of empathy, mindreading and morality forming three interrelated concepts. It is common, within this debate, to conceive of empathy as a multi-level ability with different levels of increasing cognitive complexity. Such a conception of empathy can be found, for instance, in de Waal's work (de Waal, 2006, 2008; Preston & de Waal, 2002).

Within this framework, the most basic level of empathy is often thought to be emotional contagion:⁹ an involuntary emotional resonance triggered by the attended perception of another's affective state. Such a form of empathy is automatic, often unconscious, and does not require any understanding of what initiated it. Therefore, no mindreading is involved here. Bare emotional contagion, however, may give way to a self-centred reaction, often a mood with no intentional content. This is what is usually termed 'personal distress', and is not oriented towards the welfare of the other, so there is no reason to suppose it has a moral character.

If emotional contagion results, not in a self-centred reaction oriented towards alleviating one's own distress, but in a reaction that intends to alleviate the other's distress, this by itself is usually not enough for the emotional contagion to be considered a moral motivation (Batson, 2009; de Waal, 2006, 2008; Hauser, 2001). Helping behaviour triggered by mere emotional contagion is often thought to be a case of what de Waal terms 'avoidance of aversive vicarious arousal' and which is presumed to lack a moral character:

Perhaps the most compelling evidence for emotional contagion came from Wechkin, et al. (1964) and Masserman et al. (1964), who found that monkeys refuse to pull a chain that delivers food to them if doing so delivers an electric shock to and triggers pain reactions in a companion. Whether their sacrifice reflects concern for the other [...] remains unclear, however, as it might also be explained as avoidance of aversive vicarious arousal (de Waal 2008, p. 283).

Aversive vicarious arousal, then, means that the distress of the target is perceived by the subject as a negative stimulus. It is argued that the resulting behavioural response (avoidance of aversive vicarious arousal) may not qualify as a moral one that reflects 'concern for the other', even if it is oriented towards alleviating the target's distress. This is so because it is thought to serve the selfish purpose of alleviating the subject's own distress by way of the extinction of its cause (i.e. the distress of the target). This idea can be found in other texts such as Hauser (2001, p. 276).

⁹ Nothing in my argument turns on whether emotional contagion is indeed a form of empathy.

For an instance of emotional contagion to be considered a moral motivation, it is common to establish that some cognitive overlays are required. Among these cognitive overlays is an explicit understanding of the fact that one's affective state was caused by the affective state of the other, which implies mindreading (de Waal 2006; Batson 2009; de Vignemont and Jacob 2012). It is also common to find the explicit claim that empathy as a moral motivation requires the ability to attribute mental states to others. For instance, Batson (2009) claims the following:

Feeling for another person who is suffering [...] is the form of empathy most often invoked to explain what leads one person to respond with sensitive care to the suffering of another. [...] To feel for another, one must think one knows the other's internal state [...] because feeling *for* is based on a perception of the other's welfare [...]. (Batson 2009, pp. 9-10)

A similar idea can be found in de Vignemont and Jacob (2012, p. 310). Tschudin (2006, p. 626) further suggests a causal link between the development of mindreading and the appearance of empathy. O'Connell (1995) goes as far as presenting a collection of examples of apparently empathically-motivated prosocial behaviour in chimpanzees as intended proof of the presence of mindreading capabilities in that species. The consensus among scholars, therefore, seems to be that empathy as a moral motivation requires mindreading, and thus that empathy is off-limits for a behaviour reader. In addition, other cognitive requirements are usually included. For instance, de Waal (2006; 2008) further requires that there be a certain contextual appraisal of what caused the other's affective state. For Dixon (2008), there needs to be a moral judgement with regards to the other's affective state; an explicit consideration of her suffering as a morally bad thing. Both these requirements also presuppose the ability to mindread.

It thus seems that the most that a behaviour reader could aspire to would be aversive vicarious arousal: an involuntary emotional resonance caused by the attended perception of the target's distress that results in an attempt to alleviate one's own distress by means of the alleviation of the target's. This, as we have seen, is thought to be a response lacking in moral character (Hauser 2001; de Waal 2006, 2008; Batson 2009). In what follows, however, I will

argue that such a conclusion is hasty and that aversive vicarious arousal may be considered a moral reaction, so long as certain conditions are met. Among these, crucially, mindreading will not be included.

3. Minimal Moral Empathy

In this section, I will present a minimal conception of empathy that will qualify as a moral emotion, in Rowlands' (2012) terms. Following that, I will analyse the cognitive requirements that a subject of this form of empathy must meet.

3.1. Empathy as a (Minimal) Moral Emotion

If we follow Rowlands' framework (see section 2.1.), we can arrive at a minimal definition of empathy as a moral motivation:

Creature *C* possesses *minimal moral empathy* (MME) if: it has (1) an ability to detect distress behaviour that, (2) due to the action of a reliable mechanism, results in an emotion that is directed towards the distress behaviour, and built into which is (3) an urge to change the situation that, together with the emotional reaction, (4) tracks a relevant moral proposition.

There are thus four conditions whose fulfillment would allow *C* to count as a subject of MME.¹⁰ Condition (1) entails that *C* should be able to detect distress behaviour and discriminate it from non-distress behaviour. Such detection does not have to rely solely on visual cues, but can also depend upon auditory or olfactory cues. It does not have to be infallible, either. Some instances of distress behaviour may remain undetected by *C*. But she should be fairly reliable in recognising at least some forms of distress behaviour in others. What is crucial for the purposes

¹⁰ These four conditions are sufficient for it to be an instance of empathy as a moral emotion, but they are not necessary. Other forms of empathy as a moral emotion may be conceived, some of which may be sufficient for a being to count as a moral agent, and not merely as a moral subject.

of our discussion, is that we do not need C to be able to postulate any unobservable mental states with causal power over the behaviour of the distressed individual. Thus, it is sufficient for C to be a behaviour reader.

Condition (2) is what preserves what is considered by many scholars¹¹ as the constitutive characteristic of empathy, that is, an isomorphism between the emotional state of the subject and that of the target that is caused by the attended perception of the target's emotional state. MME has emotional contagion at its core, but it is not the non-intentional version of emotional contagion that yields merely personal distress. It is an intentional form of emotional contagion, which means that the emotion it results in is not a mere mood, but an emotion that possesses a specific content:¹² one related to the other's welfare. This means that the emotion that is the result of the emotional contagion is directed towards the distress behaviour that C is witnessing. C should be able to understand that she is witnessing distress behaviour, and, as a result of her emotional contagion, she should experience this behaviour she is witnessing as unpleasant. In the same way that Smith is indignant *that* Jones snubbed him, creature C is distressed *that* the target is displaying distress behaviour.

Condition (3) establishes that the emotional contagion has a motivational aspect to it. This means that there will be a behavioural response that is motivated by it. And this response, as well as the emotional reaction, have to track a moral proposition, as is determined by condition (4). This last condition is what ensures that MME is not a regular emotion, but a moral emotion, by allowing it to involve morally evaluative content. This content will be the proposition "This creature's distress is bad." As in the Smith-Jones example, it is not required that C be able to entertain this proposition. All that is required is that her emotional reaction and subsequent

¹¹ But not all. See, for instance Zahavi & Overgaard (2012).

¹² The word 'specific' is important. At times, 'mere' emotional contagion gives way, not to a non-intentional form of personal distress, but one that has content. For instance, when a bird is startled and flies away, and immediately the rest of the flock follows it, the latter may be due to an emotional contagion which has yielded a state of fear directed towards the possibility of there being a predator around. Although this would be an emotion involving content, it would still be a form of personal distress, for the content involved has no relation whatsoever to the welfare of the conspecific who triggered the emotion.

behavioural response track this proposition, so that, if neither C's distress nor her helping reaction are misguided, then the proposition "This creature's distress is bad" must be true.¹³

Let us put forward an example to illustrate MME. Jane is sitting on a couch and crying. Her dog, Higgins, as dogs often do in these situations, nervously comes up to her and engages in various affiliative behaviours such as resting his head on her lap, or licking her arm. If we assume that Higgins is a behaviour reader, then he cannot have understood that Jane is sad, because sadness is a mental state. But he has understood that she is displaying certain behavioural cues (tears falling from her eyes, a certain bodily posture, a characteristic odour, etc.) that pertain to the more abstract category of distress behaviour. If Higgins, due to a certain reliable mechanism in him, has indeed become distressed at the sight (or smell) of Jane's crying, and is experiencing her behaviour as unpleasant, and if this, in turn, is what moves him to come up to her and engage in affiliative behaviour, then we can conclude that he is being a subject of MME, for his emotional reaction and subsequent behaviour track the moral proposition "Jane's distress is bad."

If Jane were not really crying, but only faking it to tease Higgins, then his emotion and subsequent behaviour would be misplaced, for Higgins' belief that Jane is displaying distress behaviour would be false. If, on the other hand, Jane were really sad and crying, but merely due to the fact that she is watching the ending of Titanic on TV, then Higgins' reaction would be misguided, for it would be based on the false evaluative assertion "Jane's distress is bad." The misguided character of Higgins' reaction would obtain independently of his capacity to understand the fact that Jane's distress is bad. And it is precisely because the non-misguided status of his reaction would guarantee the truth of the proposition "Jane's distress is bad," that we can assert that his reaction *tracks* this moral proposition.¹⁴

¹³ MME involves both factual content ("This creature is displaying distress behaviour") and evaluative content ("This creature's distress is bad"). C does not need to be capable of explicitly entertaining the evaluative content of her emotion, but I am assuming that she *is* capable of entertaining the factual content, and thus, of believing something along the lines of, but not necessarily equivalent to, "This creature is displaying distress behaviour." A case could perhaps be made for MME only requiring that the factual content be *tracked*, but I shall not follow this road, since Povinelli and colleagues' work strongly suggests that they would grant me this assumption (see sections 2.2. and 4).

¹⁴ Note that the attribution of MME to Higgins not only relies on the badness of Jane's distress, but also, more importantly, on the experiential form that Higgins' emotional contagion takes. In being distressed

This framework allows us to avoid the objection made by de Waal, Hauser and Batson with respect to ‘aversive vicarious arousal’. MME might indeed be a form of aversive vicarious arousal, because the other’s distress behaviour is perceived as a negative stimulus, and this results in an urge to get rid of it somehow. Indeed, in this last example, Higgins is reacting to Jane’s crying as an unpleasant stimulus that he wants to eliminate. But the traditional dichotomy between aversive vicarious arousal and moral behaviour is a false one. Experiencing Jane’s distress behaviour as an aversive stimulus does not preclude Higgins from the possibility of having a moral reaction towards it. The fact that Higgins’ emotional contagion does not result in a self-centred emotion, but in a form of distress that has Jane’s distress behaviour as its intentional object is significant. The fact that Higgins reacts to Jane’s distress behaviour by approaching her and interacting with her in a caring manner instead of merely running away from the aversive stimulus is also significant. The unpleasantness that Higgins experiences when witnessing Jane’s behaviour does not have to necessarily result in a selfish reaction, but rather, may instead constitute the *experiential form* his concern for her takes. In experiencing Jane’s distress behaviour as unpleasant, Higgins is demonstrating a sensitivity to at least one bad-making feature of situations – her underlying distress. Because Higgins is inclined to have this sort of reliable and relevant emotional reaction to the morally salient feature of situations that is someone’s suffering (even if he doesn’t understand it as such), he counts as a moral subject.¹⁵

3.2. The Cognitive Requirements of *Minimal Moral Empathy*

Although I have described the definition of empathy I have given as ‘minimal’, there are a few cognitive requirements that must be met by C in order to be a subject of MME. First and foremost, C must be capable of experiencing emotions. The definition I have given of empathy

that she is displaying distress behavior, he is experiencing Jane’s distress (behaviour) as something distressful, i.e., as something bad.

¹⁵ Moral subjecthood is not restricted to positive moral emotions. If Higgins were to possess some reliable mechanism that resulted in him *rejoicing* in the display of distress behaviour in others, and if this motivated him to contribute to their suffering, we would also have to conclude that he is a moral subject. In this case, the moral emotion Higgins would be a subject of might be labeled cruelty or *schadenfreude*. Note, however, that his lack of understanding would prevent Higgins from being held responsible for his behaviour.

makes reference only to emotions of an unpleasant or negative type (broadly referred to as 'distress'). I have done this for the sake of simplicity, and also following a general trend in the literature (see, e.g. Wispé 1986; Michael 2014), but we could arguably conceive of a form of MME with a pleasant emotion at its core. Suppose Wanda reliably undergoes a form of emotional contagion whenever she is in the presence of someone displaying joyful behaviour. If this results in Wanda being happy *that* the target is displaying joyful behaviour, and built into such happiness is an urge to preserve the situation in a way that tracks a relevant moral proposition ("This creature's happiness is good"), then this will fulfill the requirements for Wanda to be a subject of MME. I will henceforth ignore this complication and refer only to MME with unpleasant emotions at its core, the form of MME that is, probably, more likely to be found in animals, for it has quite a clear evolutionary value (see, e.g. de Waal 2008; Rowlands 2012, p. 17). The following analysis, however, should be taken to be perfectly applicable to the case of MME with pleasant emotions.

C must be able to experience at least some of the negative emotions that fall under the label 'distress': anxiety, fear, aversion, sadness, suffering, anguish, etc. The case for MME in animals is thus dependent upon the evidence for emotions in them. The empirical evidence for the presence of negative emotions in animals is quite strong, at least in the case of mammals (for a review, see Panksepp, 2004; Bekoff, 2008). There are further physiological and evolutionary considerations that quite strongly suggest that some species of animals are capable of possessing negative emotions (DeGrazia, 1996, chap. 5; Rowlands, 2002, chap. 1). For the purpose of this paper, it is not necessary to delve into all this evidence, but merely to point out that the case for MME in animals will indeed be dependent upon it.

Not only is it required that C be capable of possessing emotions, it must also be the case that C possesses intentional states, so that she is capable of possessing not only moods, but also emotions with content. We need this in order to avoid the claim that the C is acting out of mere personal distress. The urge to change the situation must be oriented towards the other's welfare, and not only towards an alleviation of C's personal distress (although this can still be acquired as a by-product of the being's moral behaviour, see Rowlands 2012, pp. 226-7). The content

towards which C's emotion is directed, i.e. the distress behaviour she is witnessing, is what transforms C's emotion from a mere cause, to a reason for her behaviour. It is a reason because it now provides a justification for C's emotion. The content turns her emotional contagion, and subsequent urge to change the situation, into what she *should* feel, given the circumstances (Rowlands 2012, p. 192).

The case for empathy will thus also be dependent upon the evidence for intentional states in animals. A full-blown defence of the presence of MME (or any other morally laden emotion) in a certain animal species will have to provide evidence for their possession of intentional states, or provide a counter-argument to the philosophical arguments against intentional states in animals. These have most famously been developed by Davidson (2001) and Stich (1979). For the purposes of this paper, however, I do not need to provide a counter-argument to their ideas,¹⁶ for Povinelli and colleagues would undoubtedly grant me the assumption that animals can believe and have other propositional attitudes (see section 2.2).

In order to avoid the claim that C is 'getting things right' (i.e. acting in the morally correct way) by accident, we also need some sort of reliable mechanism to be at the basis of her emotional reaction and subsequent behaviour – what Rowlands (2012) labels a 'moral module'. We need a mechanism that will reliably link the perception of certain bad-making features of situations (i.e. distress behaviour in others) to an emotional reaction and urge to change the situation that track a relevant moral proposition. A possible candidate for such a 'moral module' could be a perception-action mechanism (PAM).

Preston and de Waal (2002) postulated the idea of a PAM that would be responsible for motor and emotional resonance in certain species (including, but not limited to, humans). A PAM, according to them, ensures that when one perceives certain actions and emotions in others, there will be a largely unconscious activation of representations of those same actions and emotions in oneself, and that this eventually leads to the same physiological or psychological reactions one would have if one were actually experiencing such action or emotion. The discovery of mirror neurons, which fire when one executes an action, as well as

¹⁶ For counter-arguments, see, e.g. DeGrazia, 1996; Regan, 2004; Rowlands, 2009, 2012.

when one observes that same action being executed by another, was taken by Preston and de Waal as evidence for the presence of a PAM in humans and other species. Mirror neurons were originally discovered in macaque monkeys (di Pellegrino *et al.*, 1992) and there is some evidence that suggests the involvement of an analogous mirror system in emotional contagion in humans (Carr *et al.*, 2003; Dapretto *et al.*, 2006; Jabbi, Swart, & Keysers, 2007; Rizzolatti & Craighero, 2004). My notion of MME is dependent upon the evidence for a mechanism, such as a PAM, that enables emotional contagion in (at least some species of) animals, but not, as we shall now see, upon the evidence for mindreading capacities in them.

4. An Empathic Behaviour Reader

In order to understand why MME is indeed minimal, we need to understand the overintellectualisation that is often built into the idea of a moral motivation. Many theories of moral motivation follow what Rowlands (2012) calls the ASCNM (access-scrutiny-control-normativity-morality) schema. According to this schema, a certain motivation cannot be *moral* if it exerts no *normative* grip on its subject, i.e. if it is not something that the subject can embrace or resist. This, in turn, can only happen when the subject in question has *control* over it and the actions that follow from it. This control is understood to be gained through critical *scrutiny* of the motivation and the actions it motivates. In order to be able to critically scrutinise her motivation, she must have *access* to the process that delivers it. Therefore, access, scrutiny and control on behalf of the subject are often thought necessary for her motivations to acquire the status of moral (Rowlands 2012, chap. 6; see also Korsgaard 2004, 2006).

The bulk of Rowlands' book focuses on arguing that the ASCNM schema breaks down at the scrutiny-control stage, for scrutiny over one's motivations in no way guarantees control over them. Animals and humans should both be thought to have as much or as little control over their motivations as the other, because the only form of control that could be thought applicable only to humans, that is, control gained through scrutiny, is spurious. Our unique scrutinising abilities do give us responsibility over our actions, not because they give us control over our motivations

and actions, but because they give us an understanding of them that animals lack (at least to our extent) (Rowlands 2012, chap. 9). In Rowlands' characterisation of moral subjecthood, access, control, and scrutiny lose their privileged position in granting moral status. My characterisation of MME, in turn, shows how empathy can be a moral motivation without the need for access, scrutiny and control. This is, indeed, what makes this notion of empathy minimal, while allowing it to be moral.

Firstly, it is not required that C be able to access the operations of her 'moral module'. Whichever operations give rise to emotional contagion can be completely unconscious, so long as the emotional contagion indeed takes place and generates a certain type of intentional emotion. Secondly, no scrutiny over C's motivations is needed. It is not necessary that she be able to reflect upon the emotional contagion she is experiencing and the urge to change the situation it is causing. That would probably be necessary for C to count as a moral agent, but not for moral subjecthood. Finally, no control is needed, either. The emotional contagion can be involuntary and automatic. The urge to change the situation that is built into it can be conditioned or deterministic. It will not matter so long as they both track a relevant moral proposition.

Because MME needs no access, no control, and no scrutiny, it counts as minimal. Indeed, it is as minimal as it can get while retaining its moral character. Anything less cognitively demanding would collapse into personal distress or mere emotional contagion, thus losing its moral character. A 'moral module' is needed to avoid the claim that C is 'getting things right' accidentally. The possession of emotions is a prerequisite for emotional contagion to take place, and the capacity to entertain content is a necessary condition if we want MME to qualify as a reason and not a mere cause. We cannot rid C of any of these three components if we do not want her to lose the status of moral subject. However, access to the operations of her 'moral module', scrutiny over its deliverances and control over such deliverances and the behaviour they motivate are not needed. And it is within the abilities that allow scrutiny over one's motivations where we would find the ability to mindread.

The term 'scrutiny', indeed, encompasses all those abilities that allow a being to reflect upon her motivations and the behaviour they stir up. It is a form of metacognition, for it is a reflection performed upon one's own mental states. Rowlands distinguishes three components to scrutiny:

First, there is *recognition*: the ability to recognize a motivation. Then there is *interrogation*: the ability to ask oneself whether this motivation is one that should be embraced or one that should be resisted. Third, there is *judgment*: the ability to assess the compatibility between a motivation and an antecedently adopted moral principle or proposition. (Rowlands, 2012, p. 180)

The recognition cluster already implies the ability to attribute mental states (motivations) to oneself, which can be understood as a form of mindreading (see section 1). But crucially, both the interrogation and the judgement cluster require the ability to attribute mental states to others. Indeed, in order to adequately determine whether a motivation should be embraced or resisted (interrogation), one must be capable of understanding how one's actions can affect others' well-being, which implies mindreading. Furthermore, one cannot entertain moral propositions and assess their compatibility with one's motivations (judgement) without understanding the notion of moral patienthood, which involves being able to attribute, at the very least, sentience to other beings. Understanding that other beings are sentient means grasping that they feel pain and suffering as unpleasant, and thus also implies mindreading.

Such reflections are not needed, however, for MME to take place. If we were to require that C *entertain* the moral proposition "This creature's suffering is bad", then we *would* need such sophisticated intellectual abilities. But, as shown in section 3.1, in order to make sense of the morality of C's reaction, we just need it to *track* this proposition. It is thus sufficient for her to be distressed at the fact that the target is displaying distress behaviour, meaning that she experiences it as unpleasant due to an emotional contagion, and undergoes a subsequent urge to change the situation. MME, therefore, does not require mindreading. Povinelli's behaviour reader can, indeed, be a moral subject, at least to the extent that it can be motivated to act by a minimal form of empathy as a moral motivation. A subject of MME does not need to construe

others' behaviour in terms of the mental states that underlie it. She does not need to be able to understand and attribute the phenomenology that accompanies displays of distress behaviour.

Recall that behaviour reading implies a certain amount of cognitive activity; an amount that is sufficient for MME to take place. Povinelli and colleagues, as we saw in section 2.2, insist upon the idea that a behaviour reader can construct abstract behavioural categories and use them to reason about others' behaviour. Therefore, our requirement that C be able to identify distress behaviour and represent it as such can be met by Povinelli's behaviour reader. Indeed, in the same way as Povinelli (Gallagher and Povinelli 2012) acknowledges a behaviour reader's capacity to interpret certain behavioural cues ('pursing lips, bristling hair, etc.') in terms of a more abstract behavioural category ('threat display'), it can be expected of a behaviour reader to be able to interpret certain behavioural cues that are, in fact, markers of distress (e.g. vocalising in a certain way, displaying certain bodily postures, etc.), as pertaining to the more abstract category of 'distress behaviour'. And, so long as there is a reliable mechanism, such as a PAM, that ensures an intentional emotional reaction, and the latter, in turn, (together with the behavioural response it motivates) tracks a relevant moral proposition, then the behaviour reader in question is a moral subject.

A certain objection can be anticipated at this point. If a behaviour reader can detect and reason about behavioural cues, but not understand behaviour in terms of the mental states that underlie it, how can we say of it that it possesses a sensitivity to (at least) one bad-making feature of situations? Is it not precisely the phenomenology underlying displays of distress behaviour, that is, what it feels like for the subject in distress, that which makes it an objective bad-making feature of situations? This is right, but not a problem for my argument. A subject of MME is sensitive, not just to distress behaviour, but to distress itself. Indeed, distress behaviour is a reliable marker of distress, and it is sufficient for our subject to be sensitive *to* distress in others (in the sense that, when faced with distress in others, she undergoes a certain form of emotional contagion that tracks a relevant moral proposition), but we do not need her to be sensitive *that* it is, in fact, a case of distress as an unpleasant mental state. That is, the fact that she reliably experiences a feeling of (intentional) unpleasantness when in the presence of

another creature in distress is enough to say that she is sensitive to the bad-making feature of situations that is the suffering of others. The fact that she doesn't *represent* distress in others as an unpleasant mental state is irrelevant, for it doesn't preclude her from having such sensitivity.

Before concluding, it is important to emphasise two points. Firstly, I have not tried to defend that empathy is in any way *necessary* for morality. I have merely argued that behaving on the basis of MME would *suffice* to count as a *moral subject* (i.e. as the subject of at least one type of moral motivation). A being who lacked the requirements for MME could still be a moral subject if her behaviour were sometimes motivated by, say, other moral emotions, or moral considerations of a reflexive character. My framework is thus entirely compatible with the position of philosophers such as Prinz (2011) and Maibom (2009), who argue that empathy is not necessary for morality. Secondly, the thesis I have advanced is different from the point made by Kennett (2002), who argued in favour of the moral capacities of autistic individuals. These people's mindreading skills are deficient in many respects and nowhere near as sophisticated as those of a 'normal' adult human. But they still have *some* understanding of others' mental lives,¹⁷ and can certainly engage in moral reasoning. My point is a different and more fundamental one, namely, that beings who lack *all* capacity to mindread and form moral judgements *of any sort* may nevertheless pertain to the realm of moral subjects. Perhaps, however, MME never does take place throughout the animal kingdom, and all the cases of seemingly empathically motivated behaviour we see in animals have some sort of mindreading component to them. That is something that the empirical researchers must determine. My aim in this paper is simply to point out that there is no necessary connection between empathy understood as a moral motivation and mindreading. The case for empathy as a moral motivation in animals, and the broader case for moral subjecthood in them, should, therefore, not be considered dependent upon the evidence for animal mindreading.

5. Conclusion

¹⁷ At least the individuals that Kennett is interested in (see Kennett, 2002, p. 345).

The aim of this paper has been to show that, even if Povinelli and colleagues are right in their critique of animal mindreading research, this does not affect the possibility of moral subjecthood in animals. In order to argue in favour of this view, a theoretical construct has been presented – what was termed *minimal moral empathy* (MME). By way of this construct, we have shown that it is possible to conceive of a form of empathy that would a) not require mindreading capabilities, b) preserve (what is usually considered) the constitutive characteristic of empathy, and c) qualify as a moral emotion. While this is a theoretical construct, it should be noted that the possibility of its existence is relatively backed up by empirical research, which presents examples of helping behaviour in animals and points to the existence of the cognitive requirements of MME in at least some species. We can therefore conclude that there are (some) grounds for considering that (some) animals may be motivated to act by moral reasons. Povinelli's behaviour reader, while being incapable of reasoning about mental states, may nevertheless turn out to be moral.

Acknowledgements

This work was supported by the Spanish Ministry of Economy and Competitiveness, under Grants BES-2012-052504 and FFI2011-23267. I would also like to thank Kristin Andrews, Álex Díaz, José A. Gascón, Javier González de Prado, Mark Rowlands, Cristian Saborido, Kim Sterelny, and two anonymous reviewers for their comments and suggestions. Additionally, this material benefitted from discussion at the 2013 SEFA and SIFA meetings, as well as the IV PBCS workshop.

References

- Andrews, K. (2005). Chimpanzee theory of mind: Looking in all the wrong places? *Mind & Language*, 20(5), 521–536. <http://doi.org/10.1111/j.0268-1064.2005.00298.x>
- Andrews, K. (2012). *Do apes read minds?: Toward a new folk psychology*. The MIT Press.
- Andrews, K. (2013). Great ape mindreading: What's at stake? In A. Lanjouw & R. Corbey (Eds.), *Humans and other animals: Rethinking the species interface* (pp. 115–125). Cambridge University Press.
- Batson, C. D. (2009). These things called empathy: Eight related but distinct phenomena. In J. Decety & W. Ickes (Eds.), *The Social Neuroscience of Empathy* (pp. 3–15). Cambridge, MA, US: MIT Press.
- Bekoff, M. (2008). *The Emotional Lives of Animals: A Leading Scientist Explores Animal Joy, Sorrow, and Empathy - and Why They Matter* (First Trade Paper Edition). New World Library.
- Bekoff, M., & Pierce, J. (2010). *Wild Justice: The Moral Lives of Animals*. University Of Chicago Press.
- Blackburn, S. (1993). *Essays in Quasi-Realism*. New York: Oxford University Press.
- Buckner, C. (2014). The semantic problem(s) with research on animal mindreading. *Mind and Language*, 29(5), 566–589.
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12(5), 187–192. <http://doi.org/10.1016/j.tics.2008.02.010>
- Carr, L., Iacoboni, M., Dubeau, M.-C., Mazziotta, J. C., & Lenzi, G. L. (2003). Neural mechanisms of empathy in humans: A relay from neural systems for imitation to limbic areas. *Proceedings of the National Academy of Sciences*, 100(9), 5497–5502. <http://doi.org/10.1073/pnas.0935845100>
- Dapretto, M., Davies, M. S., Pfeifer, J. H., Scott, A. A., Sigman, M., Bookheimer, S. Y., & Iacoboni, M. (2006). Understanding emotions in others: Mirror neuron dysfunction in children with autism spectrum disorders. *Nature Neuroscience*, 9(1), 28–30. <http://doi.org/10.1038/nn1611>

- Davidson, D. (2001). Rational Animals. In *Subjective, Intersubjective, Objective*. Oxford University Press.
- DeGrazia, D. (1996). *Taking Animals Seriously: Mental Life and Moral Status*. Cambridge University Press.
- De Vignemont, F., & Jacob, P. (2012). What is it like to feel another's pain? *Philosophy of Science*, 79(2), 295–316. <http://doi.org/10.1086/664742>
- De Vignemont, F., & Singer, T. (2006). The empathic brain: How, when and why? *Trends in Cognitive Sciences*, 10(10), 435–441. <http://doi.org/10.1016/j.tics.2006.08.008>
- De Waal, F. B. M. (2006). *Primates and Philosophers: How Morality Evolved*. (J. Ober & S. Macedo, Eds.). Princeton, N.J.: Princeton University Press.
- De Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, 59, 279–300. <http://doi.org/10.1146/annurev.psych.59.103006.093625>
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, 91(1), 176–180.
- Dixon, B. A. (2008). *Animals, emotion & morality: Marking the boundary*. Prometheus Books.
- Gallagher, S., & Povinelli, D. J. (2012). Enactive and behavioral abstraction accounts of social understanding in chimpanzees, infants, and adults. *Review of Philosophy and Psychology*, 3(1), 145–169.
- Goldman, A. I. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading* (1st ed.). Oxford University Press, USA.
- Hauser, M. (2001). *Wild Minds: What Animals Really Think* (1st ed.). Holt Paperbacks.
- Heyes, C. (1998). Theory of mind in nonhuman primates. *The Behavioral and Brain Sciences*, 21(1), 101–114; discussion 115–148.
- Jabbi, M., Swart, M., & Keysers, C. (2007). Empathy for positive and negative emotions in the gustatory cortex. *NeuroImage*, 34(4), 1744–1753. <http://doi.org/10.1016/j.neuroimage.2006.10.032>

- Kennett, J. (2002). Autism, Empathy and Moral Agency. *The Philosophical Quarterly*, 52(208), 340–357. <http://doi.org/10.1111/1467-9213.00272>
- Korsgaard, C. M. (2004). Fellow creatures: Kantian ethics and our duties to animals. In G. Peterson (Ed.), *Tanner Lectures on Human Values* (Vol. 25/26). Salt Lake City: University of Utah Press.
- Korsgaard, C. M. (2006). Morality and the distinctiveness of human action. In S. Macedo & J. Ober (Eds.), *Primates and philosophers: How morality evolved*. Princeton University Press.
- Lurz, R. W. (2009). If chimpanzees are mindreaders, could behavioral science tell? Toward a solution of the logical problem. *Philosophical Psychology*, 22(3), 305–328. <http://doi.org/10.1080/09515080902970673>
- Lurz, R. W. (2011). *Mindreading Animals: The Debate over what Animals Know about Other Minds*. Cambridge, Mass.: MIT Press.
- Maibom, H. L. (2009). Feeling for Others: Empathy, Sympathy, and Morality. *Inquiry*, 52(5), 483–499.
- Masserman, J., Wechkin, S., & Terris, W. (1964). “Altruistic” behaviour in rhesus monkeys. *American Journal of Psychiatry*, 121(6), 584–585.
- Michael, J. (2014). Towards a consensus about the role of empathy in interpersonal understanding. *Topoi*, 33(1), 157–172. <http://doi.org/10.1007/s11245-013-9204-9>
- O’Connell, S. M. (1995). Empathy in chimpanzees: Evidence for theory of mind? *Primates*, 36(3), 397–410. <http://doi.org/10.1007/BF02382862>
- Panksepp, J. (2004). *Affective Neuroscience: The Foundations of Human and Animal Emotions* (1st ed.). Oxford University Press, USA.
- Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin’s mistake: Explaining the discontinuity between human and nonhuman minds. *The Behavioral and Brain Sciences*, 31(2), 109–130; discussion 130–178. <http://doi.org/10.1017/S0140525X08003543>

- Penn, D. C., & Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a “theory of mind.” *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 362(1480), 731–744.
<http://doi.org/10.1098/rstb.2006.2023>
- Penn, D. C., & Povinelli, D. J. (2013). The comparative delusion: The “behavioristic”/ “mentalistic” dichotomy in comparative theory of mind research. In H. A. Terrace & J. Metcalfe (Eds.), *Agency and joint attention* (pp. 62–81). Oxford University Press, USA.
- Povinelli, D. J. (1998). Can animals empathize? Maybe not. *Scientific American*, 9, 67–75.
- Povinelli, D. J., & Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences*, 7(4), 157–160.
- Povinelli, D. J., & Vonk, J. (2004). We don’t need a microscope to explore the chimpanzee’s mind. *Mind and Language*, 19(1), 1–28.
- Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, 25(1), 1–20; discussion 20–71.
- Prinz, J. (2011). Is empathy necessary for morality? In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and Psychological Perspectives* (pp. 211–229). Oxford University Press.
- Regan, T. (2004). *The Case for Animal Rights*. University of California Press.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192. <http://doi.org/10.1146/annurev.neuro.27.070203.144230>
- Rowlands, M. (2002). *Animals Like Us*. Verso.
- Rowlands, M. (2009). *Animal Rights: Moral Theory and Practice* (Second Edition). Palgrave Macmillan.
- Rowlands, M. (2012). *Can Animals Be Moral?*. Oxford University Press, USA.
- Stich, S. P. (1979). Do animals have beliefs? *Australasian Journal of Philosophy*, 57(March), 15–28.
- Tschudin, A. (2006). Dumb animals, deaf humans? In S. Hurley & M. Nudds (Eds.), *Rational Animals?* (pp. 413–438). Oxford: Oxford University Press.

Wechkin, S., Masserman, J., & Terris, W. (1964). Shock to a conspecific as an aversive stimulus. *Psychonomic Science*, *1*, 17–18.

Zahavi, D., & Overgaard, S. (2012). Empathy without isomorphism: A phenomenological account. In J. Decety (Ed.), *Empathy: From bench to bedside* (pp. 3–20). Cambridge, MA,: MIT Press.