

## Letting animals off the hook

Nicolas Delon, New College of Florida

[ndelon@ncf.edu](mailto:ndelon@ncf.edu)

**Abstract.** A growing literature argues that animals can act for moral reasons without being responsible. I argue that the literature often fails to maintain a clear distinction between moral behavior and moral agency, and I formulate a dilemma: either animals are *less moral* or they are *more responsible* than the literature suggests. If animals can respond to moral reasons, they are responsible according to an influential view of moral responsibility—Quality of Will. But if they are responsible, as some argue, costly implications must be acknowledged. If, however, they should not be considered responsible, then we may have to reassess the meaning of animal morality. I discuss ways to eschew responsibility or to tailor it to animals and argue that each requires a revised conception of animal morality.

**Word count:** 10,800

**Acknowledgments:** Thanks to Diane Michelfelder and her students and colleagues at Macalester College, Louise Daoust and her students at Eckerd College, and the audience and organizers of the Normative Animals conference at York University. Susana Monsó and Alicia Bednar provided valuable feedback on early drafts. I'm also indebted to several anonymous referees' comments and encouragements. Ideas related to this paper were discussed at the Rocky Mountain Ethics Congress at the University of Colorado, Boulder.

### 1. Introduction

What kinds of moral agents are there? Computers and cars are not moral agents. Now imagine you're walking past a playground. The children are agents, but you're reluctant to hold them morally responsible. Likewise, maybe the dogs playing at the park are agents of some sort but not *moral* agents.<sup>1</sup> They're playing by some tacit rules, but they're not moral. The realm of agents is larger than that of moral agents. Only the latter are open to attributions of moral responsibility and reactive attitudes. Parents are morally responsible and can be blameworthy for what happens to their children and their dogs. But if children or dogs fight at the park, they may be reprimanded, not held morally responsible. There are important differences between our responses to children and dogs.

---

<sup>1</sup> Agency does not entail moral agency. For recent work on animal agency, see Arruda and Povinelli (2018); Delon (2018); Jamieson (2018); Sebo (2017); Thomas (2016); Wilcox (2020).

For instance, children will normally become moral agents; dogs won't. Children need and dogs need not be brought into scaffolding practices where we hold each other accountable and raise budding agents. The standards we apply to children are sensitive not just to what they are, but to what they are starting to become and the contexts in which they grow up.

Consider cases of nonhuman animals (henceforth "animals") engaging in prosocial helping. These are anecdotes, but they are numerous enough to warrant consideration, and they illustrate growing evidence collected in laboratory and field settings in various species. On a busy highway in Chile, a dog has been hit by a vehicle and lies unconscious in the middle of the road. Another dog weaves in and out of the traffic and manages to drag the dog to safety (Rowlands 2012: 6). A monkey was filmed rescuing a friend stuck and unconscious on railway tracks at a train station in India.

Chimpanzees will sometimes help conspecifics without any direct benefit to themselves. In a remarkable video shot in Uganda at a busy road crossing, dominant male chimpanzees aid females and youth cross safely (Hockings et al., 2006; discussed in Andrews and Gruen 2014).<sup>2</sup> Animals such as apes, elephants, cetaceans, and corvids, seem to engage in mourning behavior, expressing curiosity, distress, perhaps grief around the corpses of conspecifics (Gruen 2014; King 2013; Monsó and Osuna Mascaró 2020). African and Asian elephants are known to manifest concern with distressed or deceased individuals, assisting the ailing, and showing a special interest in dead bodies of their kind. Elephants have demonstrated a capacity for empathic understanding through coalition formation, the offering of protection and comfort to others, retrieving and "babysitting" calves, aiding individuals that would otherwise have difficulty in moving, and removing foreign objects attached to others. The research also suggests that helping and empathetic behavior are not

---

<sup>2</sup> Links to the videos: <https://www.youtube.com/watch?v=-HJTG6RRN4E> (dogs); <http://www.theguardian.com/world/video/2014/dec/22/monkey-saves-dyingfriend-train-station-india-video> (monkey); <https://www.youtube.com/watch?v=f5Q0pWSeeZc> (chimpanzees).

restricted to closely related kin (Byrne et al. 2008; Douglas-Hamilton et al. 2006; Plotnik and de Waal 2014). A female elephant, Grace, was observed trying to help the dying matriarch of another family and distressed when unable to do so effectively (Douglas-Hamilton et al. 2006; Rowlands 2012)

A growing literature documents animal “proto-morality”, spearheaded early on by primatologist Frans de Waal and colleagues (de Waal 2006; 2009). Many primates exhibit “building blocks of morality”: empathy, consolation, conflict resolution, cooperation, fairness (i.e., inequity aversion) (Flack and de Waal 2000). While animals lack full-blown morality, they manifest behavior that is to some degree moral—genuinely prosocial and other-regarding. Ethologist Marc Bekoff and philosopher Jessica Pierce have argued that many species can follow moral norms (of empathy, fairness, cooperation, and mutual help), but that such norms are species-specific: there is human morality, wolf morality, rat morality, and so on. “[A]nimals are moral agents within the limited context of their own communities.” (2009: 144) Alongside the empirical literature, philosophical work on animal morality has blossomed (Andrews and Gruen 2014; Back 2018; Behdadi 2020; Clement 2013; Monsó and Andrews, forthcoming; Ferrin 2019; Fitzpatrick 2017; Monsó 2015; 2017; Monsó et al. 2018; Monsó and Wrage 2021; Musschenga 2015; Rowlands 2012; Shapiro 2006; Shupe 2021; Vincent et al. 2019; for earlier arguments, see Clark 1982; DeGrazia 1996; Pluhar 1995; Sapontzis 1987).

This paper seeks to refine our conceptual understanding of the animal morality debate. What would it take for animals to be moral agents, for their conduct to have moral worth? Can animals, as Rowlands argues, act for moral reasons? If so, what do we appraise morally—the act, the motivations, the character? I will force a dilemma on the view that animals can act for moral reasons. If they can, resisting their moral responsibility requires more work if we want to preserve

an intermediate category of moral subjects, who act for moral reasons but are not moral agents. We'll need fine-grained conceptual distinctions, which in turn may lead to a deflated sense of 'acting for moral reasons', undermining the category of moral subjects. Thus, animals are either *less moral* or *more responsible* than many argue in the animal morality debate. I proceed as follows. Section 2 reconstructs Rowlands's influential theory of animal morality. The reconstruction leads to a dilemma putting pressure on the demarcation between moral subjects and moral agents (Section 3). I draw on Quality of Will theories of responsibility for a few reasons. Whereas it originally ruled out animals, some theories of animal morality have explicitly appealed to it (Burgis 2018; Ferrin 2019; Behdadi 2020) and it bears striking similarities to Rowlands (2012). I argue that even theories of animal morality purporting to eschew claims of responsibility face pressure from Quality of Will. Section 4 considers two ways of defusing the dilemma and accommodating moral subjects—that responsibility has different degrees and faces, respectively. I conclude with some optimism about the liberal horn and recommend some revisions to make it more palatable.

## 2. Animal morality

The inference from prosocial behavior to responsibility is typically blocked by a missing necessary condition: a capacity for deliberation or reflective assessment of motivations, or an understanding of moral concepts (Korsgaard 2006; Dixon 2008). Even arguments that animals could be virtuous (Clark 1982; Sapontzis 1987) stopped short of asserting responsibility. Commonly accepted grounds of responsibility include an agent's actions originating in a reasons-responsive mechanism or being the product of self-government or conscious deliberation. What matters is that *some* property demarcates candidates for responsibility from others, even if borderline cases exist such as children and psychopaths.

However, the demarcation only holds if moral responsibility does hinge on such features. If there is continuity between animal and human behavior and responsibility does not require conscious deliberation, then what, if anything, blocks the inference? Much of human behavior is automatic, habitual, affective, and opaque, and nonetheless open to moral appraisal (Arpaly 2003; Cova 2013; Ferrin 2017; Musschenga 2015; Railton 2014). It is then tempting to conclude that animals are open to similar forms of moral appraisal. Thus, work on animal morality suggests that animal behavior may be open to appraisal relative *at least* to group-specific norms—rules delimiting appropriate behavior within the social group, according to which individuals sometimes evaluate and sanction each other. Even when they don't, we can perform the evaluation.

On the other hand, we could be concerned about the collapse of the demarcation. It is beyond the scope of this paper to argue that seeing animals as morally responsible imposes unfair burdens on them, but I will sketch a rationale. Intuitively, claims about animal morality are not degrading or disrespectful, they do not objectify animals or reinforce prejudices about their inferiority. Quite the opposite. So whence the worry? Elsewhere, I argue that the inductive risk associated with mistakenly attributing morality to other animals is not negligible (Delon, ms). Recent work on methodology in animal cognition focuses on the risks associated with failing to ascribe certain cognitive capacities to other animals (e.g., Andrews and Huss 2014; Birch 2017; Mikhalevich 2015). Traditionally, the risk of over-attribution has been seen as worse than that of under-attribution. This recent work rightly argues that prioritizing false negatives over false positives is misguided, for reasons both scientific and ethical. Both are errors, and the former can have high ethical costs. But there are also risks to attributing capacities that animals lack (Birch 2018). Some studies suggest that attributing morally laden capacities to animals can affect our attitudes. Jared Piazza, Justin Landy, and Geoffrey Goodwin (2014) have found that perception of harmfulness (having a harmful as

opposed to benevolent disposition, relative to human welfare) has a negative effect on attributions of moral standing, independently of their sentience or intelligence. Because “perceiving an animal as having a benevolent disposition enhances people’s moral consideration for that animal, which is likely to promote better treatment of it” (121), seeing animals are moral or immoral could have unforeseen consequences. Thus, moral subjects are not just moral patients: they deserve distinctive protections and respect (Rowlands 2012: 248-54); being able to exercise one’s moral abilities is constitutive of flourishing (Monsó et al. 2018).

How work on animal morality can affect our treatment of animals is an open question—can it warrant punishment or third-party intervention? Morality has two sides, and not all moral animals play nice—predation, aggression, and callousness are pervasive. Our perception of predators could change if we saw them as moral agents. We might see chimpanzees, dolphins, and orcas as sometimes immoral. Our attitudes to coyotes and even wolves, already considered a nuisance by farmers and the US Fish and Wildlife Service, could further deteriorate. Such moral costs must be part of inductive risk assessments. I will thus work on the assumption that the costs of overattributing moral characteristics deserve serious scrutiny.

The abovementioned cases, for Rowlands, “form parts of a large and growing body of evidence for the claim that some animals can exhibit moral behavior.” (2018: 469) Most scientists and philosophers deny that possibility by setting stronger conditions on moral behavior: *X can act morally if and only if X can be morally responsible*, and responsibility requires metacognitive abilities that animals lack. While endorsing a standard, reflective picture of moral responsibility, Rowlands argues, *pace* Korsgaard (2006) and Dixon (2008), that animals *can* act for reasons, even without metacognition. Animals are “motivated to act by moral reasons, not merely causes ... where these reasons take the

form of emotions with identifiable moral content.” (2012: 35) We can reconstruct Rowlands’s reasoning as follows (I depart slightly from his four-part “unpacked” argument, pp. 33-35):

- i. To be a moral subject is to be motivated to act by moral considerations, which provide reasons for those actions.
- ii. Moral considerations can take the form of *morally laden emotions*.
- iii. An emotion is morally laden if it tracks a moral evaluation or judgment as part of its content.
- iv. Some emotions in some animals have evaluative content under some plausible description.
- v. Therefore, some animals are capable of morally laden emotions (from iii and iv).
- vi. Therefore, morally laden emotions provide motivating reasons for animals to act (from i, ii, and v).
- vii. Therefore, animals can be moral subjects.

A crucial premise (ii) is that animals are capable of *morally laden emotions*, intentional states with identifiable moral content such as “This creature’s distress is bad”. Such emotions consist of two components: cognitive (a representation of a state of affairs) and evaluative (an affective valence) (iii). Rowlands uses an intricate “tracking” strategy for ascribing content. It consists in using sentences as “*de dicto* ascriptions of content to ourselves to explain the behavior of animals.” (2012: 57) A similar strategy applies to evaluative content. “Emotions, if they are legitimate, track true evaluative propositions, but they do not require that the subject of an emotion entertain, or even be capable of entertaining, such a proposition.” (67) Animals can experience moral emotions but not form moral judgments.

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. (2012: 69)

The claim that E tracks  $p$  means that there is a truth-preserving relation between E and  $p$  such that (3) (see Monsó 2017 for an application of the strategy to care and empathy). Thus, tracking allows us to assess emotions for correctness.<sup>3</sup>

Suppose Rowlands is correct that emotions involve intentional content such that they can (in)correctly represent. Emotions do not just represent, they can also motivate. A controversial aspect of Rowlands's view, granted for the sake of the argument, is his *externalism* about moral motivation. If emotions are responsive to reasons, they can be motivations that track moral reasons, *even if* the subject does not or cannot entertain such reasons. Responsiveness to reasons is responsiveness to morally relevant objective features of the world, such as suffering or distress. In sum, emotions constitute morally evaluable motivations if they represent features of the world that happen to be reasons for the animal's conduct and if they are efficacious.

Rowlands argues that some animals can be moral *subjects* even though none, except for human beings, are moral *agents*. A "minimal moral subject" meets the following sufficient conditions:

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"). ... Moral subjects are ... sensitive

---

<sup>3</sup> One subtle difference: in cognitive tracking, the animal does have a belief; in evaluative tracking, the animal's emotion simply *tracks* the evaluative proposition. Thanks to [ ].



to the good- and bad-making features of situations in the sense that they entertain intentional content emotionally. (2012: 230-1)

In contrast, “the extent to which one is an agent is the extent to which one understands what one is doing, the likely consequences of what one is doing, and how to evaluate those consequences” (240). Moral agency and moral subjecthood are “logically independent”. Moral agents possess further capacities to understand *that* certain motives and actions are right or wrong and *why* (243). Rowlands concedes that moral agency, being a function of understanding, may come in degrees. However, animals can be moral agents “to such a small extent that, if we were to think of agency as a categorical matter ... then we would almost certainly say [they are] not an agent at all.” (ibid.)

Rowlands’s key move is to dissociate moral evaluation and responsibility, making moral subjecthood a “desirable” category (2017: 471), and several authors concur (Burgis 2018; Monsó 2015; 2017; Monsó et al. 2018). If certain facts or properties can be evaluated morally without presupposing a responsible agent, then animals lacking capacities for moral agency may still be open to moral evaluation of their behavior or motivations if they are reliably responsive to moral reasons. Remember that reasons need not play a conscious or deliberative role in the animal’s mental life.<sup>4</sup> The reasons, however, are implicit in the phenomenology of their emotions. Another’s distress is experienced as unpleasant and motivates one to engage in affiliative behavior. The badness of distress, and its motivational pull, color the subject’s experience of the target’s distress (Monsó 2017: 351). This makes the presumed “moral module” efficacious.

---

<sup>4</sup> For Monsó (2015: 676), Rowlands endorses realism about moral facts. To me, despite “a reasonably robust sense of ethical objectivity” (Rowlands 2017: 472), he is only committed to *reasons-externalism*. The good- or bad-making features are independent of “the subjective states of the *agent*”. This is different from saying they are mind-independent.

Consider the Chilean dog again.

His companion lies unconscious on a busy road. This is, let us suppose, a bad-making feature of the situation. The first requirement is that the would-be rescuer is sensitive to this bad-making feature. Such sensitivity does not require that the dog is able to think thoughts such as “This is bad!” The appropriate sensitivity can, in fact, be purchased by other means [e.g., empathetic capacities; response to distress] ... Nowhere in this general picture is there any suggestion that the dog has control over his sentiments, still less that he is able to critically scrutinize them. (2018: 473)

The last bit is crucial. Much of Rowlands’ argument consists in burden-shifting, aimed at the SCNM (Scrutiny-Control-Normativity-Motivation) schema (2012: Ch. 6-7) or “nexus” (2017). The initial appeal of the idea that morality depends on metacognitive abilities “rests on the fallacy of the miracle-of-the-meta.” (2012: 189). According to SCNM,

The ability to critically scrutinize one’s motivations gives one control over them. This control permits these motivations to make a normative claim on their subject, and so makes them the sort of motivations that might be moral. (2017: 470)

Rowlands argues at length that the appeal to control leads to regress and rests on confusion about its role in making motivations normative. His central thesis is that the moral value of an action is logically distinct from the blame- or praiseworthiness of the agent. While the latter requires control, and so perhaps metacognition, the former does not. Thus, an animal’s motivation can be moral without metacognition.

Moral motivations may come cheap, but Rowlands hasn't argued that responsibility requires metacognition. Nor has he shown that moral subjecthood isn't sufficient for responsibility. The dilemma arises from dismantling the SCNM nexus: weakening the conditions for having moral motivations weakens the conditions for responsibility; on the other hand, reinstating stringent conditions on the latter presupposes something like SCNM. It is unclear why a stringent view of responsibility would welcome an entirely separate ("logically independent") category of moral evaluation. I will return to this possibility in section 4.

Before presenting the dilemma, let us recap. There are moral agents and mere agents. Some mere agents are moral subjects, causally responsible for their actions, not morally responsible yet capable of acting for moral reasons. The challenge is to prevent sufficient conditions for responsibility from trickling down into our evaluations of moral subjects. As noted, being too liberal with our attributions is risky, so we should be reluctant to expand the scope of moral responsibility unless we have sufficient epistemic and practical reason to.

### **3. Animals on the hook**

#### *3.1. Rowlands's dilemma*

I argue that Rowlands's argument, when combined with certain views about moral responsibility, entails that some animals can be morally responsible. My argument does not generalize to all theories of responsibility, but its focus is not arbitrary. First, the view I focus on, Quality of Will, bears revealing parallels to Rowlands's picture of moral motivation. Furthermore, it is a prominent theory, as a quick glance at recent discussions of responsibility responses and reactive attitudes would show. Maybe the best theory of responsibility does not entail that all moral subjects are also moral agents. But since Rowlands does not offer or endorse a positive conception of moral

responsibility, this remains an open question. In any case, we can take the forthcoming argument to be conditional on the plausibility of the Quality of Will view.

Nomy Arpaly's (2003) influential account of "moral worth" will bring the problem into relief. On her view, blameworthiness and praiseworthiness are constitutive of moral responsibility and depend on responsiveness to moral reasons, which is manifested by a depth of concern for what *happens to be* moral rather than what an agent *takes to be* moral. I will consider each of these features shortly. Admittedly, Arpaly doesn't share Rowlands's view of animals' capacity to act for moral reasons. She writes,

creatures not acting for reasons at all cannot be either morally praiseworthy or morally blameworthy ... One cannot blame or praise a creature who cannot be expected to perceive the morally relevant features of situations any more than an elephant can be expected to perceive legal factors, aesthetic factors, or contexts in which a baseball player should not bunt. (2003: 131)

Rowlands would agree. But she also writes,

the dog's mind presumably cannot grasp—nor can it track, the way even unsophisticated people can—such things as increasing utility, respecting persons, or even friendship. ... Thus, even if this animal can act for reasons, to some extent, it cannot respond to *moral reasons*, even though it may occasionally come close. (146)

If animals were responsive to reasons, they could be open to moral praise and blame, but responsiveness to reasons requires a capacity for moral concern, which presupposes conceptual understanding. Importantly, animals are not blame- or praiseworthy according to Arpaly, not

because they lack “agent-autonomy”, or the capacity to reflect, deliberate, and determine their motives (she denies that responsibility presupposes autonomy), but because acting for moral reasons requires more demanding cognitive capacities than it does according to Rowlands.

Rowlands argues that animals are responsive to moral reasons. Animals lack “understanding”, but, on his conception of reasons-responsiveness and moral content, morally laden emotions are sufficient for moral motivation. If he is correct, some animals are capable of what Arpaly calls “moral concern”. But if this is genuine moral concern, then animals are morally responsible by the same token. If they are not, however, then they are not reasons responsive. Both pressures are real. As noted, the empirical evidence for animal proto-morality is growing and compelling. The same evidence suggests that, maybe, some animals could be moral agents. Rowlands has only shown that animals can be moral without being responsible *given some disputed theoretical demarcation*. This is not to say the demarcation is unacceptable, simply that the category of moral subject hinges on theoretical commitments. In sum, if we lower the standards for moral subjecthood, why not also lower the standards for moral agency?

It’s interesting that Rowlands’s qualms regarding even *human* responsibility surface throughout the book. While claiming that at least many humans but no animals can be moral agents, he does seem to think that the standard picture of agency is too demanding even for us. So, if we believe that human beings are morally responsible, then maybe we should reconsider our criteria for responsibility. But if we do, we risk collapsing the moral subjecthood/agency distinction. This is *Rowlands’s dilemma*:

- *Liberal horn*: accept moral subjecthood *and* moral agency for some animals.

- *Conservative horn*: deny moral agency for animals *but also* deflate the meaning of moral subjecthood.

The rest of the paper will motivate each horn and attempt to defuse the dilemma. The next section motivates the liberal horn, which proponents of animal morality should find the most attractive.

### 3.2. *Quality of will*

I focus on an influential family of views, henceforth “Quality of Will” (Arpaly 2003; McKenna 2012; Strawson 1974).

*Quality of Will*: (i) a person is morally responsible for an action when that action expresses her quality of will, that is, her goodwill, ill will, or indifference or lack of concern; (ii) goodwill consists in attitudes such as a desire for the right or the good or a concern for what is morally good or right.<sup>5</sup>

Importantly, the agent acting with goodwill is responsive to moral reasons *de re*, to what happen to be reasons for the action rather than the fact that it is good—the content of the agent’s attitude is not *de dicto* concern for morality (Arpaly 2003; Markovitz 2010). Mark Twain's Huckleberry Finn is praised for helping his slave friend Jim escape, even though Huckleberry views himself as flouting what he believes to be the right reasons (property rights, the law). His praiseworthiness derives from his being responsive to moral considerations *de re*. He does the right thing (=helping Jim escape) for the right reason (=Jim is a friend and a person), but without consciously entertaining

---

<sup>5</sup> David Shoemaker (2013) distinguishes between three interpretations of “quality of will”: *character*, *judgment*, or *regard*, yielding three “noncompeting conceptions of responsibility” and targets for distinct subsets of responsibility responses. Perhaps some animals exhibit quality of regard since they have affective and cognitive attitudes such as seeing a conspecific as in distress and to-be-helped or a companion as worthy of trust and reciprocity.

this being the right thing as his motivating reason. It's not that he is not deliberating. He is, in fact, torn. But his acting on the right reasons is not the product of his deliberative process. Had it been, Jim might have concluded he was doing the *wrong* thing!<sup>6</sup>

Rowlands denies that (ii) is a necessary condition for moral evaluation. Animals can have moral motivations even without any understanding of the concepts of right or wrong. But, while some Quality of Will views do require some such understanding for responsibility, this is not a core commitment. As noted, moral concern here is understood *de re* rather than *de dicto*. Insofar as an animal is motivated by what makes an action right, she has the required kind of concern. Moreover, Rowlands's tracking strategy seems specifically designed to allow for such attributions. The relevant moral proposition is implied by an animal's having the relevant moral emotion that does not misfire.

If goodwill does not require autonomy, can animals manifest it? Rowlands's tracking strategy enables the attribution of identifiable moral content to animals. Animals are responsive (*de re*) to features of the environment that our (*de dicto*) attributions identify as morally relevant: he's my buddy, she helped me last time, he's in distress, etc. Burgis (2018) and Ferrin (2019) argue that animals can manifest goodwill, thus taking the liberal horn of the dilemma. Burgis specifically argues that some animals can understand morally relevant considerations (*de re*, not *de dicto*) by Arpaly's lights. (2018: 132) Recall the example of Grace the elephant. She was acting for the right reasons in manifesting (*de re*) concern for the welfare of the matriarch, acting upon motivations

---

<sup>6</sup> In contrast, for Johnson King (2020), Finn's act lacks moral worth because he is *accidentally* doing the right thing; he has no idea that he is performing an act of the right type. Rather, he is motivated by the right-making features but does not understand the relationship between those features and the act's rightness. Moral worth requires *deliberately* doing the right thing. As a reviewer notes, this criticism, which would otherwise block the liberal horn of the dilemma, is not compatible with Rowlands's tracking account of moral motivation, for tracking is reliable. On the other hand, on Johnson King's more demanding view, animals' behavior cannot be moral.

whose content is responsive to moral considerations. She likely experienced empathy (distress by proxy) and sympathy (other-regarding concern) for a groupmate in distress. If such content is sufficient for goodwill, and if autonomous deliberation is not necessary, Grace is responsible according to Quality of Will. To express goodwill is to act for the right reasons, as Grace seemed to be doing.

Asia Ferrin (2019: 138-42) draws on the empirical literature to argue that empathetic capacities are sufficient (albeit not necessary) for the capacity to act for moral reasons. De Waal's (2009) Russian doll metaphor describes layers of empathy, from (i) state-matching (emotional contagion) at the core to (ii) sympathetic concern (consolation) to (iii) perspective-taking (targeted helping) on the outside. Many animals exhibit at least (i), including rodents (Bartal et al. 2011); many primates at least (ii). Ferrin defends two claims. First, empathy (affective and cognitive) is necessary and sufficient for moral action, especially responsiveness to others' states. Second, both affective and cognitive empathy are found to various degrees across species, including apes, cetaceans, and elephants. These animals meet the criteria for manifesting quality of will. Accordingly, their actions can have moral worth.

By the same token, Grace could have *failed* to show proper concern for the matriarch. Thus, if she is morally responsible, she could be blameworthy. Oddly, such an implication is rarely considered. Work on animal morality typically focuses on morally admirable behavior (but see Monsó 2021; Monsó and Wrage 2021; Shupe 2021). For Rowlands, "praise would be an inappropriate attitude to bear toward [moral subjects]." (2012: 252); Burgis suggests that moral animals are open to praise but not to blame (2018: 132). There seems to be a praise/blame asymmetry (for an empirical review, see Anderson et al. 2020). Rationales for it vary. Some argue that praise and blame have different control conditions--the ability to do otherwise is a condition of blame but not praise



(Nelkin 2008; Wolf 1980). One can be blameworthy only if one had alternative possibilities, while one can be praiseworthy even if one did not. Praise merely requires acting for the right reasons. The ability to do otherwise is a more stringent condition. On this view, some animals could meet the conditions for praiseworthiness but not blameworthiness because they lack control-relevant abilities but can still act for the right reasons. Rowlands would agree with the verdict but cannot avail himself of this justification for the asymmetry since he rejects control as a condition of moral evaluation. Moreover, theories of responsibility do not distinguish between moral subjects and moral agents, so it's at best unclear how these justifications mesh with his view.

A different but related assumption is that a higher bar of justification or evidence is required for blame than praise (Vilhauer 2015; Wolf 1980). Indeed, the risks of harm are lower in praising than blaming mistakenly; praise tends to benefit the target, blame tends to harm (Argetsinger 2022; Mackenzie 2021; McGeer 2012; Pickard 2011; 2017). The asymmetry is reinforced by the fact that our access to animals' motivations is opaque, so we should be charitable about their motivations. This echoes the caution favoring false positives over false negatives in animal research (Birch 2017). But these are epistemic and pragmatic considerations that don't bear on whether animals are *worthy* of blame or praise.

A natural thought is that elephants cannot express *ill* will when failing to help others in distress. Yet one could argue that orcas tormenting baby seals and chimpanzees brutally killing infant chimps are manifesting what seems like ill will, cruelty, or indifference toward suffering (Monsó 2021; Monsó and Wrage 2021). Still, most of us assume that failing to help or hurting others does not indicate such attitudes: animals are at least excused when their actions fail to express proper concern, either because they lack a crucial capacity or because of their circumstances (diet, scarcity). And so, we admire or praise the nice chimpanzees and let the nasty ones off the hook of our indignation or

blame. Regardless, our attitudes suggest that we treat them as responsible at least for the good things they do.

### 3.3. *Accountability*

Ferrin writes, “[t]hough animals are sometimes morally responsible, we may not be able to engage in practices of holding them responsible given the communication barrier and lack of overlapping social context.” (2019: 146) On the other hand, “[s]ome animals seem to experience reactive attitudes toward each other such as resentment, indignation, hurt feelings, anger, gratitude, reciprocal love, and forgiveness.” Ferrin suggests that animals are likely responsible to each other (“intraspecies accountability”) but not across species boundaries (“interspecies accountability”). We probably should not *hold* animals responsible, for reactive attitudes are only locally applicable by group members. If so, we may not be warranted in expressing reactive attitudes toward them even if we could recognize that their actions have moral worth. Thus, the recognition of moral subjecthood in other animals may entail *intra-* but not *inter-*species responsibility.

Dorna Behdadi (2020) takes a different route to the conclusion that some animals, who participate in “moral responsibility practices” (MRPs), are accountable to each other. Behdadi’s alternative to “capacity-focused approaches” sees moral agency as “the participation in certain social, inter-relational practices” (2) and argues specifically, from evidence on canine cognition and social play, that canids participate in MRPs, hence are moral agents. According to practice-focused approaches (which overlap with Quality of Will; see Strawson 1974; McGeer 2012; McKenna 2012; Vargas 2013), participants in MRPs “share a strong disposition to internalize norms and to participate in the attitudes, expressions, and practices that surround them.” (Behdadi 2020: 5) Canids are competent participants in canid normative “communicatory practices”, “a relevant analog to at least

some forms of moral exchange in terms of asking for reasons, explanations, or acknowledgment and responding by providing explanations, excuses, or acknowledging transgressions.” (11) Canids can thus be appropriate targets of blame when shared norm communication is possible—when canids are, as a Strawsonian could put it, potential moral interlocutors (Watson 2004: 235). Indeed, the dispositions and inclinations relevant to MRPs coincide with abilities enabling quality of will. Accordingly, they can adopt something akin to the Strawsonian “participant attitude” to each other.

Ferrin’s and Behdadi’s views have two implications. First, morality is species-specific, and the evaluation of moral subjects is relative to context (cf. Bekoff and Pierce 2009). Moral subjects are off the hook relative to us (but see Shupe 2021). By the same token, moral subjects internalize different norms, so that they act for different reasons than we do. We may be able to identify whether and when they act for moral reasons, but our spaces of moral reasons may not overlap much. So, *we* may lack standing to adopt the participant attitude toward *them*. Second, *if* interspecies communication and sufficient social overlap could be secured, interspecies accountability would make sense. Perhaps our “relations of mutual trust and affection” with companion animals provide such a context (Scanlon 2008: 166).

#### 3.4. *Protecting moral subjects*

In sum, we have philosophical and empirical reasons to extend Quality of Will to some animals. Yet the meaning and scope of their responsibility remain unclear. Granted, my argument is conditional on the plausibility of Quality of Will. Since Rowlands does not discuss it, I can only surmise what he would respond. Two cases he has offered to maintain the separation between moral motivation and moral responsibility will help.

First, consider the real-life case (“Evil children”), in England, of Robert Thompson and Jon Venables, two ten-year-old boys who, on February 12, 1993, abducted, tortured, and murdered three-year-old Jamie Bulger. Thompson and Venables became “the youngest convicted murderers in English history.” (2017: 471) As Rowlands notes, “Under questioning, they revealed that they had planned to abduct and murder a child that day”, so we presume they acted intentionally and were motivated to inflict suffering and kill. Even if one assumes that, because of their age, they fell below the threshold of responsibility, Rowlands expects the reader to agree that their motivations were morally evil.

Unfortunately, the only supporting claim is that denying that the boys had morally bad intentions, “if one is not in the grip of a peculiarly warped moral psychology, is as counterintuitive as a claim can get.” (ibid.) Let’s concede, then, that their motivations were evil. Does it not follow, according to Quality of Will, that they were somewhat morally responsible? They did the wrong thing for the wrong reasons. The question is then whether their motivations were morally laden. They clearly manifested a lack of moral concern. Should we expect the boys to manifest such concern? No less but also no more than what we expect of moral subjects. Thus, according to Rowlands, their motivations would be evil even if the boys were mentally ill or under the influence of factors beyond their control. They are moral subjects, open to moral evaluation, but not moral agents. Why not hold them accountable? The condition of their exemption is they are children, though it’s worth noting that, on a Strawsonian view, extreme evil serves as its own exempting condition by placing wrongdoers outside the bounds of the moral community. Gary Watson drew attention to this

ambivalence between antipathy and sympathy, blame and exemption, with regard to extreme evil. It could be clouding our intuition regarding the boys.<sup>7</sup>

A few things cast doubt on the moral status of the boys' motivations, though. First, they were held legally responsible and convicted, presumably partly on account of their motivations. According to Strawson, "[c]hildren are gradually becoming moral agents" (Watson 2004: 229) even if they lack full moral understanding. This suggests that the subject/agent distinction is porous. If, however, the boys were not responsible, this is because the moral psychology we deploy to explain their behavior discounts the moral status of their motivations. They may be malicious or vicious, and we may justifiably harbor antipathy toward them, but not evidently in a *moral* sense. Whether such psychology is "warped", as Rowlands says, requires argument. The claim that their motivations are obviously immoral, rather than pathological, made in support of the subject/agent distinction, lacks support.

A final point concerns Rowlands's appeal to parity (2017: 473). Rowlands takes this sort of case to confirm the logical independence of moral motivation and responsibility. However, why should our attitudes to children carry over to other species? There may be pragmatic reasons to appraise the boys' motivations that won't apply to animals, such as the need for social order, plaintiffs' legal claims, and scaffolding practices of moral education (McGeer 2012; Vargas 2013). We may separate

---

<sup>7</sup> In "Responsibility and the limits of evil: Variations on a Strawsonian Theme" [1987], Watson (2004: 219-259) has shown that Strawson's theory implies that evil counts as its own exemption condition. "If holding responsible requires the intelligibility of moral address, and if a condition of such address is that the other be seen as a potential moral interlocutor, then the paradox results that extreme evil disqualifies one for blame." (235) On the one hand, we have standing to blame evil wrongdoers; on the other hand, heartless murderers like Robert Harris do not seem capable of heeding our demands, and so cannot be morally addressed; we lack a shared framework of values. The alternative is to deny (*pace* Strawson) that responsibility requires membership in the moral community. According to McKenna (1998), while Harris *is* not a member of the moral community, he has the *capacity* to participate in it, which explains his responsibility. It's not that he does not understand our values; he repudiates them.

moral evaluation from the fact of holding responsible and yet think that the former is functionally justified by responsibility practices. We turn children into members of the moral community by evaluating their motivations before they can even be held responsible. None of those facts apply to other animals. Hence, even if we concede that the boys' motivations were evil, it does not follow that moral subjecthood applies outside the context at hand.

The second case Rowlands discusses is that of Hitler in a world of hard determinism, where no one is morally responsible, "which may or may not be the actual world" (2017: 471).

We might ... justifiably ... refuse to blame or hold him responsible for what he does. But refusing to classify his motivations as even falling into the category of the moral is highly counterintuitive. (ibid.)

Granted, Hitler's moral motivations are abhorrent even under hard determinism. After all, we can see psychopaths' motivations as vicious while (sometimes) refraining from holding them responsible. But in what sense exactly are deterministic Hitler's motivations of the moral kind? Rowlands implies that determinism precludes control, including over one's motivations, and therefore responsibility, but that motivations do not presuppose control to be morally appraisable. However, not only is this a controversial claim in the responsibility literature but rejecting the control condition leads naturally to a view like Quality of Will and therefore to the liberal horn of the dilemma. In such cases, moral responsibility and moral motivation stand or fall together. In sum, they make the distinction between moral subjecthood and moral agency intuitively plausible, but they can hardly establish it without further argument, whereas Quality of Will has the theoretical virtue of harmonizing our judgments about the cases. Its simplicity is not decisive, but it

explains the appearance that Thompson, Venables, and Hitler deserve blame even under circumstances that would normally count as exempting conditions.

To recap, Rowlands's dilemma was either accepting that animals can be morally responsible (liberal horn) or deflating the import of moral subjecthood (conservative horn). The dilemma arises from the combination of moral subjecthood with Quality of Will. In the remainder of the paper, I zoom out and lay out groundwork to make the prospect of animal responsibility less threatening. Once we understand what it does not entail, perhaps the subject/agent distinction will lose some of its appeal. We could take the liberal horn without worrying. I highlight the costs of each option and conclude that preserving moral subjecthood requires more conceptual work.

#### **4. Animals off the hook**

Each of the two “ways out” consists in protecting animal morality from the upward pressure of responsibility, or at least its practical implications: degrees of responsibility (4.1); aspects or “faces” of responsibility (4.2).

##### *4.1. Off the hook, first pass: Degrees of responsibility*

The idea that responsibility and blameworthiness can be a matter of degree is no longer controversial. Reactive attitudes should be sensitive to the degree to which an agent is responsible (i.e., competent and/or free from coercion or other responsibility-canceling influences) and the degree to which their action expresses the relevant ground of responsibility.

Some authors who have argued that animals can be moral agents have been careful to stress the significance of degrees (DeGrazia 1996; Shapiro 2006). David DeGrazia writes, “[t]he range over which a given being is responsible is determined by the range of action possibilities for which the

being can understand a rule of conduct, roughly what its point is, the consequences of breaking it, and so on.” (1996: 203n107) Moreover, because different capacities are involved in its different aspects, agency varies according to which capacities one possesses and to what degree. (204) Whatever the required competence, it is gradable, and responsibility responses should vary accordingly (see e.g., Shoemaker 2013 on the different qualities of will).

Most accounts of responsibility are amenable to degrees of responsibility. For instance, Coates and Swenson (2013) propose to amend the Reasons-Responsiveness account (Fischer and Ravizza 1998) according to how *receptive* and *reactive* to reasons an agent is. Quality of Will accounts can adjust degrees of blameworthiness to the quality of the reasons for which agents act, namely degrees of good or ill will. Blame, praise, resentment, indignation, or gratitude vary accordingly (Tierney 2019). A proponent of animal morality could then deny that their view entails that animals must be subject to the same attitudes we direct toward moral agents. If animals are just barely competent, and their actions are minimally morally worthy, then their responsibility need not trigger the same responses as normal attributions of responsibility.

Consider affective motivations, a core component of non-reflective and sentimentalist approaches to animal morality (Andrews and Gruen 2014; De Waal 2009; Ferrin 2019; Monsó 2017; Monsó et al. 2018; Rowlands 2012). We can describe their content and appraise their quality in a graded fashion—e.g., how much concern for, or sensitivity to, the distress of others a creature’s conduct manifests; how reliably responsive to morally significant situations it is. Thus, even if animals were morally responsible, they might not be very blameworthy or praiseworthy, let alone answerable to us. The range of potential moral worth of their actions may be as limited as the range of their quality of will or reasons-responsiveness.



Though attractive, this response won't insulate animals from the outward expression of reactive attitudes. Graded responses are difficult to maintain in practice. People often express reactive attitudes toward beings who should be exempt, such as children and the mentally disabled. We also miscalibrate our responses to people with impaired agency, such as addicts and patients with personality disorders, which is why some advocate for "responsibility without blame" (Pickard 2011; 2017). Sometimes, there are good reasons for holding some reactive attitudes. Strawson distinguishes between the "objective" attitude—under which we predict, manage, or control others—and the "participant" attitude—under which we hold each other to account. And he notes:

parents and others concerned with the care and upbringing of young children ... are dealing with creatures who are potentially and increasingly capable both of holding, and being objects of, the full range of human and moral attitudes but are not yet truly capable of either. The treatment of such creatures must therefore represent a kind of compromise, constantly shifting in one direction, between objectivity of attitude and developed human attitudes. (1974: 19)

But because our attitudes are "constantly shifting", one should expect some involuntary leakage. As with children, so with animals—we might end up blaming moral subjects when we shouldn't. Indeed, moral subjecthood entices us to shed the objective attitude toward animals. A reply to this concern is that the excesses of our blaming practices are just that: unjustified, and we should seek to correct them by calling for compassion or understanding instead of blame and indignation. For instance, Hanna Pickard (2017) argues that we should refrain from blaming drug addicts while keeping them responsible, because it's important for their own sake that we do so. Though, holding animals responsible doesn't benefit them the way it does people whose agency is impaired; it isn't guided by the end of recovery or rehabilitation.

In sum, degrees of responsibility do not dissolve the dilemma. Either animals are moral subjects because they can act for moral reasons, but then they are morally responsible or we will, in practice, be tempted to express some responsibility responses; or animals can only be responsible to a benign degree, but then the content of their motivations is shallower than we might have thought. Perhaps we can mitigate the implications of taking the liberal horn of the dilemma by drawing some finer-grained distinctions.

#### 4.2. *Off the book, second pass: Faces of responsibility*

Start with a distinction between *being* responsible and *holding* responsible, or between reasons to judge that a creature is responsible and reasons, in practice, to hold them responsible. Angela Smith (2007) makes the distinction to argue that our attributions of responsibility should not be sensitive to the same considerations that count for or against responding in certain ways, typically with reactive attitudes, to someone being responsible. The question of whether a creature is responsible is distinct from whether it would be fair or appropriate to blame her, even if blameworthiness is conceptually tied to responsibility. There is also a difference between judging someone to be blameworthy and expressing blame, let alone punishing.

Taking degrees of responsibility and the distinction between responsibility *judgments* and *responses*, we might avoid the implication that we should express much by way of reactive attitudes toward animals for their morally good or bad deeds. As Watson notes, “Holding people responsible ... also involves a social setting in which we demand (require) certain conduct from one another and respond adversely to one another's failures to comply with these demands.” (1996: 229) Since Strawson, the moral responsibility literature has echoed the idea that responsibility responses presuppose a capacity to participate in interpersonal relationships and the moral community

(McKenna 2012). If animals are not implicated in this social setting, then we need not hold them responsible, even if they are. Remember the emphasis on intra-species accountability by Behdadi, Bekoff and Pierce, and Ferrin. Rowlands might argue that the objectivity of the moral facts that moral subjects are tracking allows us to appraise them even without a shared social setting. It is also plausible that the shared social setting requirement applies to responsibility but not subjecthood. Either way, more needs to be said about the ethical standards that should inform our appraisal of animals of different species.

Consider now another helpful distinction between *attributability* and *accountability* (Watson 1996). Attributability reflects what Watson calls the aretaic face of responsibility (from the Greek *arete*, meaning excellence), whereas accountability (to others) involves reactive attitudes, holding responsible, which implies believing and acting like the responsible person is accountable to us or others (231).

Suppose (*pace* Watson) that we can engage in the aretaic appraisal of animals—morally appreciating their excellences and defects, judging them as the authors of their conduct. A dog could be foolish or courageous, and it could be appropriate for us to express our approval or disapproval of their behavior, but not appropriate to hold them responsible—to *demand* that they answer to us or the moral community. For, as Watson puts it, “[t]o be intelligible, demanding presumes understanding on the part of the object of the demand. The reactive attitudes are incipiently forms of communication”, or “moral address”, which presume a form of understanding of “the basic demand” of which animals and young children are incapable. (2004: 230) “The boundaries of moral responsibility,” Watson writes, “are the boundaries of intelligible moral address.” (258)

Thus, the dog's conduct could reflect well or poorly on them, they could be a moral subject, but we may not infer that they are responsible—praiseworthy or blameworthy—for their conduct. The distinction could honor the distinct category of moral subjects. The question is, again, whether can we maintain a clear demarcation between those different kinds of judgments in practice.

Watson draws a clear line. He denies that animals are susceptible to “aretaic appraisal”, which applies to “one's purposes, ends, choices, concerns, cares, attachments, and commitments” (1996: 244), hence “presupposes moral capacity, the capacity for adopting and pursuing ends”. (245)

Fischer and Tognazzini, in their own “physiognomy of responsibility”, concur: “By asking whether the agent is open to, or is a ‘sensible target’ of, aretaic appraisal, we are asking whether the agent exercised the capacities required to make the agent *the sort of creature whom it might make sense to appraise aretaically*”, which excludes dogs. (2011: 384) (cf. Wolf 1993: 63) A dog's “viciousness” is not moral viciousness because she cannot intend to hurt or manifest a lack of moral concern for others, unlike “certain psychopaths, who can indeed have specifically moral intentions” (ibid.).

There is, however, evidence that chimpanzees, orcas, and bottlenose dolphins can intentionally hurt each other, maybe manifesting negative moral emotions such as cruelty, envy, or resentment (Monsó and Wrage 2021: 16-17). If Rowlands is correct, moral subjects possess the required capacities. They are capable of flexible, intentional behavior and moral emotions that reliably track morally relevant features of situations. If so, we should accommodate nonhuman moral subjects within our “physiognomy of responsibility.”

It's plausible that humans and animals are exempt on different grounds—psychopaths because they cannot respond to moral reasons, although their actions are still “attributable to them in an aretaic sense” (Fischer and Tognazzini: 387); animals because they can't entertain moral reasons.

Psychopaths act in ways that we find morally appalling. But Rowlands thinks the same is true of moral subjects! Unlike psychopaths, and perhaps like children, moral subjects' conduct and motivations are presumed to be *responsive to reasons*.<sup>8</sup> We can, using his tracking strategy, reconstruct rational standards for their conduct. In contrast, psychopaths have cognitive access to, and understand, but fail to properly respond to moral reasons. This reveals a tension: psychopaths could be appraised aretaically, animals not, even though they can respond to moral reasons.

More plausibly, barring excuses or justification, moral subjecthood gives us standing to hold, if not express, some reactive attitudes toward some animals. Moral subjects should earn from us more than the objective stance of the ethologist yet less than the participant stance of ordinary responsibility practices. If our responses are likely to misfire, moral subjecthood calls for revisions of our responsibility model. Still, moral subjects *are* responsible in some sense (aretaic attributability). One could also endorse a practice-based, agency cultivation model of responsibility (McGeer 2012; Vargas 2013) in which animals play no part, but even those views presuppose a quality-of-will account for blame. And if our practices purport to nurture and develop responsible agents, the question becomes what we should make of moral subjects. Our ordinary practices involve different but, importantly, pervasive kinds of interspecies interactions, from companion animals to currently and formerly farmed animals to animals in the wild. We can't just assume that no context gives rise to responsibility responses. As Vargas notes, "distinct forms of acculturation provide agents with differential capacities to recognize and respond to moral considerations in different contexts." (2013: 245) The question then becomes one of "moral ecology":

---

<sup>8</sup> Does the dilemma also apply to children? Are they either less moral or more responsible than we think? If young children can act for moral reasons, then we can draw the distinctions discussed above or say that children are at least partially responsible. We could also take one horn (conservative) for animals, another for children (liberal). Either way, we may need to recalibrate our attitudes if moral subjecthood is how we operationalize our evaluations of children and animals.

once we look beyond intrinsic features of agents to the wider set of relations that structure the various capacities of interest to us, we find that moral ecology matters. ... the circumstances that support and enable exercises of agency in ways that respect and reflect a concern for morality. (246)

Before concluding, I'd like to consider a final way to avert the dilemma. Recent work on normativity suggests that several species of primates possess normative competence (e.g., Andrews 2021; Fitzpatrick 2020; Monsó and Andrews, forthcoming; Vincent et al. 2019). The range of norms includes norms of obedience, reciprocity, care, social responsibility, and solidarity of various forms. In chimpanzees, norm compliance is not external and accidental but is internalized and rests on norm-sensitive motivations. The evidence is growing more generally that normative behavior extends far beyond apes, cetaceans, and elephants, to canids, corvids, and rodents (Monsó and Andrews, forthcoming; Monsó 2017; Monsó and Wrage 2021). This literature suggests that (some) animals respond to normative reasons, but it also offers an alternative: that animals could be *normative* without being *moral*. If animals can respond to norms and their motivations form part of their excellences, they may qualify for some sort of non-moral aretaic appraisal. This would involve replacing the category of moral subjects with that of normative animals.

These various replies defuse, to an extent, concerns about the liberal horn of the dilemma. I was initially inclined to see them as making *ad hoc* distinctions against the backdrop of continuity between humans and other animals that motivated moral subjecthood. I am now inclined to embrace them. After all, theorists of responsibility believe that different senses of responsibility track important facts about responsibility. We could avail ourselves of these distinctions and conclude that, according to Quality of Will, animals can be apt targets of aretaic appraisal but should not be held accountable for their actions. And perhaps that is how it should be. Aretaic

appraisals are less burdensome than accountability and seem less morally risky. If that is how it should be, then we might defuse the liberal horn of the dilemma. And we could do this while granting my working assumption that responsibility is burdensome. A broader concern about our psychology subsists though: we often shift within the multifarious physiognomy of responsibility unwittingly, especially when our norms are ill-defined, as they are with animals. Deep facts about responsibility notwithstanding, we should tread carefully when it comes to moral subjecthood.

## 5. Conclusion

Knowing whether animals can be moral agents is morally important: moral agents have interests in exercising their moral agency and may have obligations. Some argue that there is a middle ground between mere agency and moral agency: moral subjects, who can act for moral reasons without being morally responsible. Others argue that animals can be responsible but only within their communities. I have put pressure on both views to generate a dilemma: on the liberal horn, the demarcation between moral subjecthood and responsibility dissipates; on the conservative horn, insulating animals from responsibility deflates the significance of moral subjecthood. By drawing finer-grained distinctions, I have sketched a few ways to let animals off the hook: praise/blame asymmetry, degrees, and faces of responsibility, normativity without morality, to clarify the possibilities and identify areas where more conceptual work is needed. Whether or not animals are moral, we owe them credit where it is due, but only there.

## References

Anderson, Rajen A., Molly J. Crockett, and David A. Pizarro (2020) 'A Theory of Moral Praise'.

*Trends in Cognitive Sciences* 24, 9, 694–703. DOI: [10.1016/j.tics.2020.06.008](https://doi.org/10.1016/j.tics.2020.06.008).

- Andrews, Kristin (2020) 'Naïve Normativity: The Social Foundation of Moral Cognition'. *Journal of the American Philosophical Association*, 6(1), 36–56. DOI:[10.1017/apa.2019.30](https://doi.org/10.1017/apa.2019.30).
- Andrews, Kristin and Lori Gruen (2014) 'Empathy in Other Apes'. In H. L. Maibom (ed.) *Empathy and Morality* (Oxford: Oxford University Press), 193–209.
- Andrews, Kristin and Brian Huss (2014) 'Anthropomorphism, Anthropectomy, and the Null Hypothesis'. *Biology & Philosophy*, 29(5), 711–29. DOI: [10.1007/s10539-014-9442-2](https://doi.org/10.1007/s10539-014-9442-2).
- Argetsinger, Henry (2022) 'Blame for Me and Not for Thee: Status Sensitivity and Moral Responsibility'. *Ethical Theory and Moral Practice*. DOI: [10.1007/s10677-022-10274-z](https://doi.org/10.1007/s10677-022-10274-z)
- Arpaly, Nomy (2003) *Unprincipled Virtue: An Inquiry Into Moral Agency*. Oxford: Oxford University Press.
- Arruda, Caroline T. and Daniel J. Povinelli (2018) 'Two Ways of Relating to (and Acting for) Reasons'. *Mind & Language*, 33(5), 441–59. DOI: [10.1111/mila.12185](https://doi.org/10.1111/mila.12185).
- Back, Youngsun (2018) 'Are Animals Moral?: Zhu Xi and Jeong Yakyong's Views on Nonhuman Animals.' *Asian Philosophy* 28(2): 97–116. <https://doi.org/10.1080/09552367.2018.1453234>.
- Bartal, I. B.-A., Decety, J. and Mason, P. (2011) 'Empathy and Pro-social Behavior in Rats'. *Science*, 334(6061), 1427–30. DOI: [10.1126/science.1210789](https://doi.org/10.1126/science.1210789)
- Behdadi, Dorna (2020) 'A Practice-Focused Case for Animal Moral Agency'. *Journal of Applied Philosophy*, 1–18. DOI: [10.1111/japp.12486](https://doi.org/10.1111/japp.12486).
- Bekoff, Marc and Jessica Pierce (2009) *Wild Justice: The Moral Lives of Animals*. Chicago: University of Chicago Press.
- Birch, Jonathan (2017) 'Animal Sentience and the Precautionary Principle'. *Animal Sentience*, 16(1). DOI: 10.51291/2377-7478.1200.
- Birch, Jonathan (2018) 'Animal Cognition and Human Values.' *Philosophy of Science*, 85(5), 1026–37. DOI: [10.1086/699744](https://doi.org/10.1086/699744).



- Burgis, Jennifer Lynn (2018) *Making Covenants with Brute Beasts: Making Room for Non-Human Animals in a Contractualist Framework*. Dissertation (Ph.D.), Philosophy, University of Miami.  
<https://scholarship.miami.edu/esploro/outputs/doctoral/Making-Covenants-with-Brute-Beasts-Making/991031447234502976>.
- Byrne, Richard, P.C. Lee, N. Njiraini, J.H. Poole, K. Sayialel, S. Sayialel, L.A. Bates, and C.J. Moss (2008) 'Do Elephants Show Empathy?' *Journal of Consciousness Studies*, 15(10-11), 204–25.
- Clement, Grace (2013) 'Animals and Moral Agency: The Recent Debate and Its Implications'. *Journal of Animal Ethics*, 3(1), 1–14.
- Coates, D. Justin, and Philip Swenson (2013) 'Reasons-Responsiveness and Degrees of Responsibility'. *Philosophical Studies*, 165(2), 629–45. DOI: [10.1007/s11098-012-9969-5](https://doi.org/10.1007/s11098-012-9969-5).
- Cova, Florian (2013) 'Two Kinds of Moral Competence: Moral Agent, Moral Judge'. In Bert Musschenga and Anton van Harskamp (eds.), *What Makes Us Moral? On the Capacities and Conditions for Being Moral* (Dordrecht: Springer Netherlands), 117–30.
- DeGrazia, David (1996) *Taking Animals Seriously: Mental Life and Moral Status*. Cambridge: Cambridge University Press.
- Delon, Nicolas (manuscript) 'Animal morality and epistemic risks'
- Delon, Nicolas (2018) 'Animal Agency, Captivity, and Meaning.' *The Harvard Review of Philosophy* 25 (2018): 127–46. <https://doi.org/10.5840/harvardreview201892519>
- de Waal, Frans B. M. (2009) *The Age of Empathy: Nature's Lessons for a Kinder Society*. Toronto: McClelland & Stewart.
- de Waal, Frans B. M. (2006) *Primates and Philosophers: How Morality Evolved*. Princeton: Princeton University Press.
- Dixon, Beth (2008) *Animals, Emotions, and Morality*. New York: Prometheus Books.

- Douglas-Hamilton, Iain, Bhalla, Shivani, Wittemyer, George, and Vollrath, Fritz (2006) 'Behavioural reactions of elephants towards a dying and deceased matriarch'. *Applied Animal Behaviour Science*, 100, 67–102.
- Ferrin, Asia (2017) 'Good Moral Judgment and Decision-Making Without Deliberation?'. *Southern Journal of Philosophy*, 55(1), 68–95.
- Ferrin, Asia (2019) 'Nonhuman Animals Are Morally Responsible'. *American Philosophical Quarterly*, 56(2), 135–54.
- Fischer, John Martin, and Mark Ravizza (1998) *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.
- Fischer, John Martin, and Neal A. Tognazzini (2011) 'The Physiognomy of Responsibility?'. *Philosophy and Phenomenological Research*, 82(2), 381–417. DOI: [10.1111/j.1933-1592.2010.00458.x](https://doi.org/10.1111/j.1933-1592.2010.00458.x).
- Fitzpatrick, Simon (2017) Animal Morality: What Is the Debate About? *Biology & Philosophy*, 32(6), 1151–1183.
- Fitzpatrick, Simon (2020) Chimpanzee normativity: evidence and objections. *Biology & Philosophy*, 35(45). DOI: [10.1007/s10539-020-09763-1](https://doi.org/10.1007/s10539-020-09763-1).
- Flack, Jessica C. and de Waal, Frans B. M. (2000) "Any Animal Whatever": Darwinian Building Blocks of Morality in Monkeys and Apes'. *Journal of Consciousness Studies*, 7(1), 1–29.
- Gruen, Lori (2014) 'Death as a Social Harm'. *Southern Journal of Philosophy*, 52, 53-65. DOI: [10.1111/sjp.12064](https://doi.org/10.1111/sjp.12064).
- Hockings, Kimberly J., James R. Anderson, and Tetsuro Matsuzawa (2006) 'Road crossing in Chimpanzees: A Risky Business'. *Current Biology*, 16(17), 668-70.
- Jamieson, Dale (2018) 'Animal Agency.' *The Harvard Review of Philosophy* 25: 111–26. <https://doi.org/10.5840/harvardreview201892518>.

- Johnson King, Zoë (2020) ‘Accidentally Doing the Right Thing’. *Philosophy and Phenomenological Research*, 100(1), 186–206. DOI: [10.1111/phpr.12535](https://doi.org/10.1111/phpr.12535).
- King, Barbara J. (2013) *How Animals Grieve*. Chicago: The University of Chicago Press.
- Korsgaard, Christine M. (2006) ‘Morality and the Distinctiveness of Human Action’. In F. de Waal, *Primates and Philosophers: How Morality Evolved* (Princeton, NJ: Princeton University Press), 98–119.
- Mackenzie, Catriona (2021). Culpability, Blame, and the Moral Dynamics of Social Power. *Aristotelian Society Supplementary Volume*, 95(1), 163–182.
- Markovits, Julia (2010) ‘Acting for The Right Reasons’. *Philosophical Review*, 119(2), 201–242.
- McGeer, Victoria (2012) ‘Co-Reactive Attitudes and the Making of Moral Community’. In Langdon Rbyn, Mackenzie, Catriona (eds.), *Emotions, Imagination and Moral Reasoning* (New York: Psychology Press), 299–326.
- McKenna, Michael (1998) ‘The Limits of Evil and the Role of Moral Address: A Defense of Strawsonian Compatibilism’. *The Journal of Ethics*, 2(2), 123–142.
- McKenna, Michael (2012) *Conversation and Responsibility*. Oxford: Oxford University Press.
- Mikhalevich, Irina (2015) ‘Experiment and Animal Minds: Why the Choice of the Null Hypothesis Matters’. *Philosophy of Science*, 82(5), 1059–69. DOI: [10.1086/683440](https://doi.org/10.1086/683440).
- Monsó, Susana (2015) ‘Empathy and Morality in Behaviour Readers’. *Biology & Philosophy*, 30, 671–90. DOI: [10.1007/s10539-015-9495](https://doi.org/10.1007/s10539-015-9495).
- Monsó, Susana (2017) ‘Morality Without Mindreading’. *Mind & Language*, 32(3), 338–357.
- Monsó, Susana (2021) ‘Is Predation Necessarily Amoral?’ In A. Siegetleitner, A. Oberprantacher, ML. Frick, and U. Metschl (eds.), *Crisis and Critique: Philosophical Analysis and Current Events* (Berlin; Boston: De Gruyter), 367–382.

- Monsó, Susana and Kristin Andrews. Forthcoming. 'Animal Moral Psychologies'. In John M. Doris and Manuel Vargas (eds.), *The Oxford Handbook of Moral Psychology* (New York: Oxford University Press).
- Monsó, Susana, Benz-Schwarzburg, Judith, and Bremhorst, Annika (2018) 'Animal Morality: What it Means and Why it Matters'. *The Journal of Ethics*, 22(3), 283–317.
- Monsó, Susana and Osuna-Mascaró, Antonio J. (2020) 'Death is Common, So is Understanding it: The Concept of Death in Other Species'. *Synthese*. DOI: [10.1007/s11229-020-02882-y](https://doi.org/10.1007/s11229-020-02882-y)
- Monsó, Susana and Wrage, Brite (2021) 'Tactful Animals: How the Study of Touch Can Inform the Animal Morality Debate'. *Philosophical Psychology*, 34(1), 1–27.  
DOI: [10.1080/09515089.2020.1859100](https://doi.org/10.1080/09515089.2020.1859100)
- Musschenga, Albert W. (2015) 'Moral Animals and Moral Responsibility'. *Les Ateliers de l'éthique/The Ethics Forum*, 10(2), 38–59. DOI : [10.7202/1035326ar](https://doi.org/10.7202/1035326ar).
- Nelkin, Dana (2008) 'Responsibility and Rational Abilities: Defending an Asymmetrical View', *Pacific Philosophical Quarterly*, 89(4): 497-515.
- Piazza, Jared, Landuy, J.F., and Goodwin, G.P (2014) 'Cruel nature: Harmfulness as an Important, Overlooked Dimension in Judgments of Moral Standing'. *Cognition*, 131(1), 108-124.
- Pickard, Hanna (2011) 'Responsibility Without Blame: Empathy and the Effective Treatment of Personality Disorder'. *Philosophy, Psychiatry, & Psychology*, 18(3), 209–223. DOI: [10.1353/ppp.2011.0032](https://doi.org/10.1353/ppp.2011.0032).
- Pickard, Hanna (2017) 'Responsibility without Blame for Addiction'. *Neuroethics*, 10(1), 169–80.  
DOI: [10.1007/s12152-016-9295-2](https://doi.org/10.1007/s12152-016-9295-2).
- Plotnik, Joshua M. and Frans de Waal (2014) 'Asian Elephants (*Elephas Maximus*) Reassure Others in Distress'. *PeerJ*, 2(e278). DOI: [10.7717/peerj.278](https://doi.org/10.7717/peerj.278).

Pluhar, Evelyn B. (1995) *Beyond Prejudice: The Moral Significance of Human and Nonhuman Animals*.

Durham, NC: Duke University Press.

Railton, Peter (2014) 'The Affective Dog and Its Rational Tale: Intuition and Attunement'. *Ethics*,

124(4), 813–59. DOI: [10.1086/675876](https://doi.org/10.1086/675876).

Rowlands, Mark (2012) *Can Animals be Moral?* New York: Oxford University Press.

Rowlands, Mark (2017) 'Moral Subjects'. In K. Andrews and J. Beck (eds), *The Routledge Handbook of*

*Philosophy of Animals Minds* (New York: Routledge), 469-74.

Sapontzis, Stephen F. (1987) *Morals, Reason, and Animals*. Philadelphia: Temple University Press.

Scanlon, Thomas M. (2008) *Moral Dimensions: Permissibility, Meaning, Blame*. Cambridge, MA: Harvard

University Press.

Sebo, Jeff (2017) 'Agency and Moral Status'. *Journal of Moral Philosophy*, 14(1), 1–22. DOI:

[10.1163/17455243-46810046](https://doi.org/10.1163/17455243-46810046).

Shapiro, Paul (2006) 'Moral Agency in Other Animals'. *Theoretical Medicine and Bioethics*, 27(4), 357–

73. DOI: [10.1007/s11017-006-9010-0](https://doi.org/10.1007/s11017-006-9010-0).

Shoemaker, David (2013) 'Qualities of Will'. *Social Philosophy and Policy*, 30(1-2), 95–120. DOI:

[10.1017/s0265052513000058](https://doi.org/10.1017/s0265052513000058).

Shupe, Eli (2021) 'Punishing Moral Animals'. *Canadian Journal of Philosophy*, 51(5), 351–66. DOI:

10.1017/can.2021.33.

Smith, Angela (2007) 'On Being Responsible and Holding Responsible'. *The Journal of Ethics*, 11(2),

465–84.

Strawson, Peter (1974) 'Freedom and Resentment' [1962]. In *Freedom and Resentment and Other Essays*

(London: Methuen & Co.), 1–25.

Thomas, Natalie (2016) *Animal Ethics and the Autonomous Animal Self*. London: Palgrave Macmillan.

- Tierney, Hannah (2019) 'Quality of Reasons and Degrees of Responsibility'. *Australasian Journal of Philosophy*, 97(4), 661–72. DOI: [10.1080/00048402.2018.1550795](https://doi.org/10.1080/00048402.2018.1550795).
- Vargas, Manuel (2013) *Building Better Beings: A Theory of Moral Responsibility*. Oxford: Oxford University Press.
- Vilhauer, Benjamin (2015) 'Free will and the Asymmetrical Justifiability of Holding Morally Responsible'. *The Philosophical Quarterly*, 65(261): 772–789, <https://doi.org/10.1093/pq/pqv041>
- Vincent, Sarah, Rebecca Ring, and Kristin Andrews (2019) 'Normative Practices of Other Animals'. In Aaron Zimmerman, Karen Jones, and Mark Timmons (eds.), *The Routledge Handbook of Moral Epistemology* (New York: Routledge), 57–83.
- Watson, Gary (1996) 'Two Faces of Responsibility'. *Philosophical Topics*, 24(2), 227–48.
- Watson, Gary (2004) *Agency and Answerability: Selected Essays*. Oxford: Oxford University Press.
- Wilcox, Mark (2020) 'Animals and the agency account of moral status', *Philosophical Studies*, 177(7): 1879–1899. doi: [10.1007/s11098-019-01289-x](https://doi.org/10.1007/s11098-019-01289-x).
- Wolf, Susan (1980) 'Asymmetrical Freedom'. *The Journal of Philosophy*, 77(3), 151–66. DOI: [10.2307/2025667](https://doi.org/10.2307/2025667).
- Wolf, Susan (1993) *Freedom within Reason*. New York: Oxford University Press.