



Animal freedom and its importance in the lives of persons

Joe Gurrola

To cite this article: Joe Gurrola (27 Jul 2026): Animal freedom and its importance in the lives of persons, *Philosophical Psychology*, DOI: [10.1080/09515089.2026.2708115](https://doi.org/10.1080/09515089.2026.2708115)

To link to this article: <https://doi.org/10.1080/09515089.2026.2708115>



Published online: 27 Jul 2026.



Submit your article to this journal [↗](#)



Article views: 1



View related articles [↗](#)



View Crossmark data [↗](#)



Animal freedom and its importance in the lives of persons

Joe Gurrola

Department of Philosophy, University of Maryland, College Park, MD, USA

ABSTRACT

Humans have freedom, do other animals? This paper argues that they do. Specifically, it advances an account of a kind of freedom which I term animal freedom. Animal freedom is constituted by an agent's pursuit of the things that matter to them. My account depends on two central considerations. The first is that many things matter to us independently of whether we endorse or consider their importance for our values or projects, because they reflect who we are. The second is that many things matter to other animals in just this way – that is, in a manner that reflects who they are. Since other animals can act in pursuit of the things that matter to them, they have animal freedom. But animal freedom is not restricted to the psychological lives of other animals, nor is it something which is easily contrasted with the way that we think of ourselves as free, such that we can say that ours is the freedom of persons while theirs is the freedom of animals. Instead, much of our own lives is also made up of instances of animal freedom, ones which we would regard as genuinely constituting or realizing our free agency. There is therefore considerable overlap in the free lives of human and other animals. Recognizing this fact stands to enrich our conception of ourselves as free agents, as well as our understanding of the psychological lives of other animals.

ARTICLE HISTORY

Received 13 February 2026
Accepted 9 July 2026

KEYWORDS

Free will; nonhuman animals; Caring; value; agency; Compatibilism

“Alas! Why does man boast of sensibilities superior to those apparent in the brute; it only renders them more necessary beings. If our impulses were confined to hunger, thirst, and desire, we might be nearly free; but now we are moved by every wind that blows and a chance word or scene that that word may convey to us.”

Frankenstein, Mary Shelley.

1. Introduction

Philosophers of agency are generally unfriendly to the idea that other animals are free, let alone that they and humans are free in the same way.¹ Frankfurt (1969, 1971) famously defends the capacity for freedom without the ability to do otherwise. Along the way, he draws a contrast between persons and non-persons, or wantons. Only persons are free because only

persons care about what their will is to be. Caring about one's will, according to Frankfurt, is manifest in the formation of second-order volitions, or desires about which first-order desires move the agent to act. Since other animals lack the requisite mental faculties to form attitudes of the second order, they are classed along with other kinds of wantons as beings who lack free will.² Accounts that one might expect to be more ecumenical about the presence of freedom across species, for example, those that defend the claim that two-year-old humans and severely demented persons possess the same kind of agency as human persons generally, similarly withhold this kind of agency from other animals (Jaworska, 2007). With few exceptions, the landscape of compatibilist accounts is populated by accounts that take for granted that other animals are not free, and that therefore humans and other animals are very different kinds of agents.³

This article takes the idea of animal free will seriously. It argues that there is a shared basis for free will between humans and other animals, found in the capacity to act in accordance with what we care about. Part one concerns the question of what is required for free will. It begins by considering a set of views that take freedom to consist in our ability to act in accordance with endorsed desires and reflective judgments (Frankfurt 1971; Watson, 1975). These views preclude animals from the set of free agents based on their inability to authorize attitudes to speak for themselves via reflective endorsement. I conclude part one by reviewing well-documented problems with these views. In part two, I turn to a different set of views, ones that take free will to consist in our capacity to act in accordance with what we care about (Jaworska, 2007; Sripada, 2016a). This second set of views is more amenable to the idea that other animals might have free will, but the views present problems because they consider capacities for reflection – though not their exercise – necessary for caring and therefore for free will. In response, I argue for the claim that capacities for reflection are not required for an agent to care. I propose a minimalist account of care that says that cares are mental states that 1) provide intrinsic motivation for action, including in cases where it is costly, and 2) dispose an agent to a set of affective responses based on the fortunes of the cared-for thing.

Part three is dedicated to establishing the plausibility of the claim that other animals can care. It reviews behavioral evidence that suggests that other animals act in ways that reflect what they care about. Part four discusses the value of animal freedom, and how it comprises much of our own free lives. At the end of the article, I discuss a set of objections, including the question of whether persons can be free in a special way that is not available to other animals. I suggest that persons have a capacity for freedom called the freedom of persons. However, this kind of freedom is not foundational to our capacity to act freely, nor is it fundamentally valuable. The value of the freedom of persons is explained

by the value of acting in accordance with what we care about, and acting with the freedom of persons is made possible only by the exercise of our animal freedom. It is better to think of the freedom of persons as a set of capacities that expand the paths we can take to be animal-free (Sripada, 2016b), rather than as a fundamentally different kind of freedom.

2. Free will and identification

One basis for denying freedom to other animals is that other animals lack something like a deep self. We can think of a deep self as being the part of an agent with which actions can be identified, such that we can say that, by virtue of issuing from this part of their psychology, their actions flow from them and not from something that is alien to them. A classic case demonstrating the importance of this distinction is Frankfurt's willing and unwilling addicts, in which he shows that two agents who are equally addicted to a drug, and therefore cannot do other than take the drug, can nevertheless differ with respect to the freedom of their wills (Frankfurt, 1969). Frankfurt argues that the unwilling addict's desire to take the drug is alien to him, and therefore not identifiable with him, because he does not desire to act based on this desire and does all he can to struggle against it. Consequently, his will is unfree. By contrast, the willing addict's desire to take the drug is identifiable with him and his will free because he endorses the influence of his addictive desire and would do everything in his power to retain it. Frankfurt's contention that some of an agent's attitudes speak for her and support free action while others do not has served as the foundation for deep-self accounts that attempt to locate the conditions under which attitudes and actions can be identified with an agent by virtue of originating from the right part of her psychology, as opposed to from some other part, and thereby support the attribution of freedom and responsibility.

Because of their cognitive limitations, other animals are frequently thought to fall short of the requirements for a deep self, and therefore identification. For example, since they are thought to lack the capacity for higher-order thought (thought about one's own mental states), they fail to qualify as free on accounts that require the capacity to endorse one's attitudes (Frankfurt, 1969).⁴ And since they lack the capacity to engage in explicit reflection over which of their motivations is best, they lack the capacity to output all-things-considered judgments, which other sourcehood accounts require for identification (Watson, 2004; Wolf & Schoeman, 1987). In short, it does not appear, at least according to these accounts, that the minds of other animals admit of the internal/external distinction required for sourcehood: the attitudes that are present in the minds of other animals are neither internal nor external, since other animals are not the kinds of beings with psychologies containing an

inner and outer component (Wolf & Schoeman, 1987). If freedom requires a deep self, then according to these accounts other animals are not free.

In what follows, I accept for the purpose of argument that other animals lack the capacities required for the type of identification grounded in the exercise of higher-order reflection and endorsement. The accounts that require this are primarily those that take identification to be a function of authority. According to these accounts, an action or desire is identifiable with an agent when it is authorized to speak for her. When, for example, the unwilling addict's desire for the drug expresses an unfree will, this is because the desire is not authorized to speak for her, since she has explicitly disavowed it.

While compelling in many respects, these accounts are also famously beset by problems. For starters, there are many cases where agents appear to have acted in ways that express who they are while acting contrary to their endorsed desires (Arpaly & Schroeder, 2014). A prime example here is what Arpaly and Schroeder (2014) term “inverse akrasia,” where agents with faulty moral knowledge fail to act in accordance with their faulty moral knowledge and thereby do the right thing (Liberti 2026). These accounts also have a difficult time making sense of actions that appear to reflect our free and responsible selves despite being motivated without reflective endorsement (Scanlon, 1998; Smith, 2005, 2008). Examples include failures of attention, spontaneous emotions like jealousy or contempt, and cases of forgetting (King & Carruthers, 2012). Sripada (2016a) argues that “common sense says we can be ... morally responsible for an enormous variety of spontaneous conduct. Indeed, it seems most natural to say that we are responsible for this conduct precisely because such conduct is deeply expressive of our selves” (Sripada 2016a, p. 1214). Finally, these accounts cannot account for the potentially free actions of agents who lack the capacity for reflection, such as very young children and agents with significant cognitive impairments (Jaworska, 2007). For these reasons, philosophers have sought an alternative basis for identification that can accommodate the concerns raised by authority-based views (Betzler, 2024; Sripada, 2016a). I turn to these in the next section.

2.1. Care-based views

Some philosophers have argued for a second basis for identification besides authority. Following Shoemaker (2015a), I will call this second basis *authenticity*. An action or desire is authentic with respect to who an agent is when it reflects her character (Shoemaker 2015a; Watson, 2004), what matters to her (Creasy, 2025; Wolf, 2015), or what she cares about (Arpaly & Schroeder, 2014; Jaworska, 2007; Sripada 2016a; Shoemaker 2015a, 2015b).

While it is reasonable to think that other animals may act in ways that reflect a stable character,⁵ discussions of authenticity as a basis for attributability most often focus on actions or attitudes that are identifiable with an agent because they reflect what she cares about or what matters to her (Arpaly & Schroeder, 2014; Jaworska, 2007; Sripada, 2016b).⁶ For this reason, I focus primarily on the relationship between cares and attributability in what follows. Not just any actions that express an agent's first-order desires qualify as expressing what she cares about, since caring about something is different than having a first-order desire for it. To care about something, we must be affectively tethered to the object and its fortunes and disposed to respond in the right kind of way to the fortunes of that thing. For example, we feel fear or anger at the prospect of its being harmed or elation at its being saved from harm, and we are happy to see the thing is flourishing and disappointed or worried when it appears to flounder (Jaworska, 2007; Shoemaker, 2003; Sripada 2016a).⁷ It is reasonable to think that a person who is not disposed to respond with strong emotions to the harm of an object cares little for that object, even if they say they care or have otherwise endorsed its importance.

Cares should also motivate us to action for their object's own sake, and not merely for the sake of something else we value (Sripada 2016a). If the only reason we are ever motivated to act in the service of *a* is that *a* is important to other things that matter to us, then we might reasonably wonder whether *a* is something we care about or whether it is something we value in a merely instrumental way. Finally, we should be willing to engage in costly behaviors to ensure an object's wellbeing (Arpaly & Schroeder, 2014; Creasy, 2025).⁸ For example, if we say we care about something but only act in the service of that thing when it is easy to do so, and abandon all pursuit of it every time it becomes difficult, it is reasonable to think we might not care about that thing after all.

2.2. *Cares and reflective judgment*

What we care about does not always align with our reflective judgments about what we care about (Sripada 2016a; Jaworska, 2007; Kelly, 2022).⁹ For example, there are many things that matter to us even though we do not want them to. Here is an instructive case from Arpaly (1999):

Imagine a person, Lynn, who discovers that she is a lesbian and is deeply disturbed by that discovery. Her homosexual desires conflict with her values and her sense of her identity. She does not want her desires to motivate her into action under any circumstances – the very thought scares her more than anything else. If Lynn were to go to her favorite college professor for help, she would likely be told that she should try to accept herself for who she is, refrain from attempts to suppress her true self, and so on. If, on the other hand, she were to read the moral psychology literature and

believe its claims, she would probably conclude that she was right and her homosexual desires are not truly her own.

It is reasonable to think that Lynn's lesbian desires are not mere appetites but rather express something she cares about. Despite her disavowal of the desire, it continues to exert influence over her attention and her affective states. We can imagine that though Lynn feels prospective fear about acting in accordance with this desire, she also feels sadness when it is frustrated (for example, when she successfully suppresses it). It is also reasonable to think that sometimes Lynn's desires exert a motivational force that she must suppress in order to do what she thinks is right.

Though Lynn would disagree that she cares about this thing, it is reasonable to think that based on their affective-motivational profile, her sexual desires matter to her and express something she cares about. They play the right kind of role in her psychological life as a valuing agent. To say otherwise would be to suggest that conative states that consume much of our attention and have strong affective traction on us do not matter, but we typically think these are good indicators that something matters to us. Lynn's case suggests that authority, grounded by endorsement or explicit judgments of value, is not the only basis for identification. For not only can we imagine judging that were Lynn to act on her lesbian desires she would be acting in a way that expresses her deep self; it also makes sense to describe Lynn as acting with freedom, and not merely freedom from external constraints.

At least part of the reason for this is that some of the things we think we care about are adopted as something like self-governing norms that take the form of intellectual commitments to or endorsements of the value or goodness of something, which we have good reason to think are different kinds of mental attitudes than cares as they have been discussed here (Carruthers, 2024; Kelly, 2022).¹⁰ These commitments are unlikely to motivate apart from their relevance to something that we care about (Carruthers, 2024; Kelly, 2022).¹¹ Nor are they apt to elicit the affective responses indicative of cares without being relevant to something else we care about. While it is reasonable to think that these attitudes reflect what we value in some broader sense (Mitchell-Yellin, 2025), the fact that they lack the motivational and affective profile of cares suggests that endorsements of or commitments to the value of something or its goodness are not sufficient for generating corresponding cares, nor are they corresponding cares themselves. Considerations about the rather limited power of belief and commitment to change our underlying affective structures suggests that we do not have ultimate say over what we care about, and that our cares do not always reflect our commitments and endorsements.

We can extend this line of thought by considering that many of the things we care about are products of unreflective processes for value acquisition rather than reflective commitment. Research in evolutionary and normative psychology suggests that many of our caring attitudes may be innately specified, such as the value we place in our bodily health (Arpaly & Schroeder, 2014) or the approval of our group members (Carruthers, 2024, 2025). Other cares might not themselves be specified by our evolutionary history but are nonetheless best explained by the unconscious processes that facilitate norm acquisition and internalization without any need for conscious thought about the norm (Carruthers, 2024, 2025; Kelly, 2022). Research in normative psychology suggests that humans, and perhaps some other social creatures, possess psychological adaptations that facilitate norm acquisition and internalization since such adaptations would have proved fruitful for individuals whose survival depended upon other people (Westra & Andrews, 2022). Crucially, all of this can be done without explicit instruction or commitment to a norm in reflection.

The above suggests that we can have caring attitudes without endorsing them. Despite this, some philosophers of care have argued that certain sophisticated cognitive capacities are nevertheless required for care (Sripada 2016a; Jaworska, 2007). These capacities include the capacity to form judgments about the goodness of something (Sripada 2016a) or to recognize it as important (Jaworska, 2007), and the capacity to remain committed to the caring attitude itself (Sripada 2016a). In the next section, I argue that two of these cognitive capacities – valuing the caring attitude itself and the capacity to form judgments about the goodness of the cared-for object – are not required for caring. In part three, I argue that the third capacity, the ability to recognize something's importance, is present in other animals.

2.3. Against the evaluative judgment condition

Sripada (2016a) argues that the disposition to form judgments that cast the object of care in a normatively favorable light is part of the syndrome of dispositions that make up the functional profile of cares (1210). The thought is that when we care about something, as evinced by our affective responses to the fortunes of that thing and our motivations to preserve it, we are also disposed to form judgments that take that thing as good. Moreover, these judgments often help explain at least part of the reason why we choose to act as we do when we reason about what to do. Our cares dispose us to judge that the cared-for object is good, and make these judgments available to reflection when we reflect over how to act. Were an agent not disposed to form these kinds of judgments, we would have good reason to think that they do not care about the thing.

Importantly, the requirement does not say that evaluative judgments are sufficient for picking out what one cares about. (This is the position of the authority-based views discussed above.) Nor does it say that positive evaluative judgments must be present for someone to care about something. The reason is that there are many factors that might act as “masks” or “finks” that prevent the manifestation of the disposition to form positive evaluative judgments. For example, a person might be disposed to endorse the goodness of something but never consider the question of whether it is good. In such a case, the person would still possess the disposition to evaluate the object of care as good, and therefore care on Sripada’s account.

While it is true that caring often disposes us to form positive evaluative judgments of the cared-for object, we should not think that this is a necessary condition for caring. Consider cases of very young children (Jaworska, 2007) or perhaps people who have become cognitively impaired with the result that they have lost the capacity to make reflective judgments about the goodness or value of something though they can unreflectively appraise something as good and be intrinsically motivated to act in the service of this object. In cases like these, perhaps only the affective and motivational dispositions of the functional profile of care are present, yet it seems that these agents retain the capacity to act in ways that reflect what matters to them.

There seem to be other cases where the disposition to form judgments that cast something in a normatively favorable light is absent, yet we think that actions issuing from an intrinsically motivating attitude that takes this thing as its object express what an agent cares about. For example, we sometimes have cares that have been cultivated and nurtured all our lives, but which now run afoul of our newly acquired moral commitments. Consider cases like the newly-minted vegetarian who has come to believe that eating meat is wrong. Such a person will likely continue to be intrinsically motivated by the appearance and smells of meat-based dishes, and they may also feel disappointment at ordering the tofu on the menu instead of their preferred meat-based dish. Yet given their new moral beliefs and the ways these have reshaped their set of reasons concerning the goodness of eating meat, it seems they’d also lack the disposition to endorse the goodness of eating meat. It’s true that they might feel slightly pulled to endorse the goodness of eating meat by considerations about the pleasure to be gained from eating meat and the conceptual association between pleasure and goodness, but it is easy to imagine the force of these considerations being outweighed by considerations that take eating meat to be morally bad. This person would therefore lack the underlying disposition to form judgments about the goodness of eating meat, and the explanation for this is not a mask or fink but rather that their present set of reasons does not output the judgment that eating meat is good. Regardless, it seems that if this person

were to eat a steak at their next meal, we would say that this reflects something they care about.¹² This suggests that the disposition to form judgments about the goodness of a particular thing is not required for a person to care about that particular thing, and that an agent might care if they possess only an affective connection to an object and are intrinsically motivated by it. This line of argument provides independent reasons to think that very young children and those with significant cognitive impairments can care.

The above suggests that the kind of attributability that is grounded in cares bears a complicated and tenuous relationship to the capacity for reflective reasoning and judgment. If we can imagine agents like Lynn and the newly-minted vegetarian who act in ways that express who they are while acting in ways antipodal to their endorsed judgments, and other agents like children and people with significant cognitive impairments acting in ways that express who they are while lacking the capacities required to form these kinds of judgments, it becomes unclear what role, if any, the capacity for reflective thought plays in underwriting free actions of this kind.

I therefore suggest that we adopt the following minimalist conception of cares:

Minimalism about Cares: Cares are mental states that 1) provide intrinsic motivation for action including in cases where it is costly, and 2) dispose an agent to a set of affective responses based on the fortunes of the cared-for thing.

According to this account, the capacity for reflective judgment is not necessary for the capacity to care. This suggests that many kinds of creatures might care. If, as many philosophers think, free will consists in the capacity to act in accordance with what we care about, then many creatures besides ourselves might also have free will.

2.4. The capacity to conceive of things as important

Jaworska (2007) argues that though the capacity for reflection is not required for caring – and that therefore very young children and severely demented adults can care – other animals nonetheless lack the capacity to care. Her primary reason for this position is that other animals do not have the required emotional capacities to care. Jaworska distinguishes between primary and secondary emotions. Primary emotions are things like fear, which she argues are relatively inflexible responses to preset stimuli, whereas secondary emotions, like grief, require the capacity to track objects across time. The capacity to track objects across time is required because it is one of the main ways that we represent things as important. According to Jaworska, cares are comprised of a set of varied emotional dispositions

that 1) take something as their object, 2) extend over time, and 3) play a role in organizing the agent's psychic life, where this consists in being motivated to act in the service of the same thing over time and the presence of planning behaviors (Jaworska, 2007, p. 564). She argues that though other animals can exhibit some emotions directed at objects, these emotions do not extend over time in the right kind of way, nor do they play a role in organizing the agent's psychic life.

According to Jaworska, the affective responses of other animals are largely "stereotypical," since they exhibit only a "rudimentary pattern of attachment" that does not track the well-being of the object across time and circumstances. Her primary supporting example is child-rearing in other animals. She argues that though other animals act as if they care for their offspring, they do not care for them "beyond biologically prescribed reactions" (564). This is evidenced by the fact that other animals cease to exhibit affective reactions related to the wellbeing of their offspring once the offspring mature. She also notes that the affective bonds other animals have to their offspring do not support planning and policy-making behaviors, and that this is best explained by the fact that they do not perceive their offspring – the object of their affective states – as important in a way that extends across time.

While I am sympathetic to Jaworska's conditions on caring – it makes sense to think that cares have a temporal component as well as affective and motivational components – her analysis of the emotional lives of other animals leaves much to be desired. For example, the mere fact that an animal will cease to exhibit affective attachment toward its offspring at some future time hardly implies that they do not care about their offspring when they are raising them. As we will see in the next section, other animals form long-term friendships across time, some of these lasting for many years, and are particular about whom they befriend. They console social partners who have recently lost fights. Most strikingly, they grieve, and do so in ways that can hardly be described as fitting a stereotypical pattern of behavior. I turn now to this evidence.

3. Caring animals

Research on the social and emotional lives of other animals suggests that, far from being inflexible agents who act exclusively on innately specified desires, they possess capacities for learning and forming relationships that make up their capacity to care (Benz-Schwarzburg & Wrage, 2023; Monsó et al., 2022; Seyfarth & Cheney 2012; Vincent et al., 2018). Recognizing this fact is important, since some philosophers of agency have argued that other animals cannot care since they lack the capacity for emotion "beyond fixed patterns of emotional responses to specific, immediately sensed features of

one's environment" (Jaworska, 2007, p. 555).¹³ In this section, I present evidence of animal behaviors that suggest that animals are capable of being caring agents.¹⁴

Perhaps the best evidence for cares in other animals can be found in the long-term friendships they form.¹⁵ Evidence suggests that primates, dolphins, and elephants form long-term social bonds with one another that can last for many years (Seyfarth & Cheney, 2012). As just one example, some female baboons have been shown to maintain friendships with one another for upwards of five years. The strength of these bonds is measured by the amount of time the baboons spend with one another and the frequency with which they engage in grooming practices with their "friends" compared with other baboons.¹⁶ Friendships of this length are not typically formed with all or most members of their communities. Instead, most female baboons in the community form "weak, impermanent social bonds" with other females in their communities (Seyfarth & Cheney, 2012, p. 156). This suggests that friendship behaviors, while evolutionarily adaptive, are far from inflexible automatic behaviors and that baboons base their choice of friends on specific features. Along these lines, the best predictor of whether two females shared a strong bond was matrilineal kinship, especially for mothers and daughters.¹⁷ However, female baboons also frequently formed friendships with other baboons they were not related to. Other reliable predictors included age-similarity and status-similarity, suggesting that baboons' choice of friends are sensitive not only to genetic relatedness but also the social features of other baboons.¹⁸ It is striking that reliable predictors of whether two baboons in a community will be friends mirror many of the factors that determine whom we spend the most time with in our own lives.

It is difficult to explain enduring friendships of this kind without reference to other animals' capacities for learning from and developing reliable long-term values toward objects within their environments. Many of these relationships include very different cooperative behaviors that are spaced far apart. For example, grooming and supporting each other in coalitions are very different activities that do not regularly co-occur in time, yet the animals in these relationships reliably manifest both kinds of behaviors, seemingly with sensitivity to what has previously occurred in their interactions (Seyfarth & Cheney, 2012). This suggests that not only are these animals behaving in ways that reflect their evaluative stances (e.g., by choosing to make or keep a friendship with one conspecific over another), but that they also do so in a way that takes into account previous experiences (cf. Jaworska, 2007).

Consider next that other animals have been found to exhibit consolation and grieving behaviors. Consolation behaviors are actions aimed at alleviating distress or providing comfort. For example, chimpanzees spend more

time and console social partners who have been in and lost fights (De Waal, 2007). These behaviors are exhibited in many animals, including elephants, monkeys, rodents, canids, and cetaceans (Monsó et al., 2022). Moreover, at least in some animals they appear to be mediated by the same chemical pathways as in humans. Prairie voles who are administered an oxytocin agonist do not show consolation responses to bonded partners in distress, while those not administered the agonist will reliably console (Burkett et al., 2016). The authors of this study took these results to suggest that the motivational pathways for consolation are similar between humans and prairie voles, since oxytocin helps mediate empathetic responses in humans.

One prominent discussion of animal grief in the literature gives the following description of the dimensions of grief in other animals:

Survivors alter their behavior in the wake of a death in significant ways that indicate intense distress. Depending on the species, these changes may include atypical patterns of eating or sleeping; withdrawal from social activities; and expression of upset at or near the body through vocalizations, facial expressions or body language. (King, 2013)

Unlike consolation behaviors, most of the evidence we have of grieving behaviors in other animals is anecdotal. Given the emotional significance of grief for humans, many scientists are often reluctant to apply the term to other animals' behavior. However, the numerous examples we have of such behavior are striking. Here is one from Barbara King, describing the behavior of an orca whale after losing her calf:

In July 2018, a female orca named J35 captured world-wide attention for her unprecedented vigil. J35, also known as Tahlequah, is a member of the closely monitored Southern Resident population of orcas in the Salish Sea, off the coast of Washington State and British Columbia. She had just given birth, following a nearly year-and-a-half-long gestation period. It was her second offspring, a daughter, and the first live birth in the declining Southern Resident community in three years. But 30 minutes after birth, the calf died. J35 would not let her baby go. With great effort, she swam with the tiny body on her head and made deep dives to retrieve it when it slipped off. Other members of her pod registered her distress: at one point, a group of females gathered in a tight circle around J35, an act of apparent emotional attunement that lasted at least two hours. Seventeen days and 1,000 miles passed before J35 finally released her daughter's corpse for good. (King, 2019)

Examples like this suggest that other animals can form strong bonds with each other that mirror the bonds that we feel to those close to us.¹⁹ Behaviors like those depicted in the story of J35, especially the prolonged period of deviation from normal behaviors, also suggest that we respond to the loss of those we care for in similar ways. Moreover, even if one assumes that the lives of other animals are made up primarily of comparatively simple behaviors, it is difficult to read grieving behavior in this way. Instead, given that the behavior is typically in response to the loss of agents

with whom the grieving animal has shared a social bond, and that the behavior removes the animal from her normal course of action, we have reason to think that this behavior is indicative of not just flexibility in response to the agent's environment, but also a capacity for grasping the loss of something that is of significance to her.

As a final piece of evidence for the claim that other animals possess the capacity to care, consider evidence that they manifest various helping behaviors apparently motivated by emotional responses. The helping behavior of rats is especially instructive since rats are often excluded from the set of other animals whose cognitive capacities might qualify them as agents. Studies consistently show that rats who have the opportunity to free a fellow rat will do so, even if they are unable to interact with the freed rat after the fact (Bartal et al. 2011; Monsó et al., 2022). Additionally, rats are more likely to help a rat in a distressing circumstance if they have previously experienced the same unpleasant conditions (Monsó et al., 2022). Lesion studies also suggest that helping behavior is motivated by rats' affective systems. Studies on subjects with lesioned amygdalae show that they are less likely to exhibit helping behavior. For example, in a task where a subject can choose to secure food for themselves and a conspecific or only for themselves, rats who have intact amygdalae reliably choose to secure food for both parties while rats with lesioned amygdalae fail to reliably do so (Hernandez-Lallement et al., 2016; Monsó et al., 2022). Monsó et al. (2022) interpret these findings in light of evidence that links damage to the amygdala to impaired sensitivity in emotional processing and social cognition in humans, and suggest that the rats with lesioned amygdalae failed to help their partners due to a deficit in emotional processing.

That the subjects in these experiments rely on emotional processing to motivate their helping behavior is yet further evidence that other animals appear to have the affective machinery to care.²⁰ They reliably engage in behaviors that help their partners over other behaviors that might be just as appealing, and seemingly lose the capacity to engage in these behaviors once their affective systems are damaged.

These data suggest that other animals have the capacity to care about things in the sense of care laid out in section 2. Other animals appear to respond flexibly to their environments in ways that suggest that they take some things to be more important than others. They also appear to be motivated to behave in these ways based on learned emotional and not merely "instinctual" processes. Perhaps most importantly, other animals appear to care for one another as depicted by behaviors that we ourselves would manifest toward others we care about. Of course, whether any token action expresses an animal's care will be hard to discern, but this is also true of our ability to discern cares in the actions of other humans as well as our ability to discern cares in our own actions.

4. Animal freedom in human lives

Let's take stock. I have argued that acting freely is independent of the capacity for reflection since we can act in ways that express our cares without possessing or ever exercising such a capacity. I have also argued that other animals have the capacity to care, which, in conjunction with the first claim, suggests that they have the capacity to act freely. But we might wonder whether this kind of freedom is something we should care about. Many philosophers are committed to the view that humans and other animals are fundamentally different with respect to their status as free agents (Doris, 2015; Frankfurt, 1969; Sripada, 2016b). This suggests that even if there is a kind of freedom available to both humans and other animals, it isn't the kind of freedom we should be especially concerned with since it isn't what makes us the distinctively free beings we are. So, should we care about a kind of freedom that is not distinctive to us?

Yes, we should care. To see why, recall that many of our most valuable relationships, those that reflect some of the things we care about most, are not the product of reflective reasoning over what we should care about, but are rather products of more general value-internalization processes (Arpaly & Schroeder, 2014; Carruthers, 2025; Kelly, 2022). Consider our relationships with our family members. We do not begin to care about these people because we think we should; rather, in many cases we come to love them as the result of the love and care that they show us. Furthermore, our actions toward the people we care for are often motivated not by explicit thoughts about what we should do or what we want our will to be, but rather by first-order appraisals of the states of affairs involving them that, in conjunction with our set of beliefs about available courses of action, help us to output decisions that lead us to act in ways that express our care for them (Carruthers, 2024, 2025). Finally, though some of us may reflect over whether we want to continue loving mom, dad, or our children, this question likely never occurs to many of us. Even if it did, it is unlikely that our decisions about whether to love will extricate us from our affective bonds to these people or change what we care about (Creasy, 2025).²¹

We should also reflect on the fact that reflectively endorsed actions can sometimes be inconsistent with manifesting a caring attitude toward our objects of care. Consider Williams (1981) case of a husband diving into the ocean to save his drowning wife only after reflecting over whether saving her instead of a complete stranger who is also drowning is permissible. In this case, it is reasonable to think that the husband's action, though consistent with his reflectively endorsed value for his wife, fails to manifest the proper care for her, since it occurs only after a process of reflection over what to do. By contrast, had the husband

been moved immediately to jump into the ocean upon seeing his drowning wife without any further thought, we would be more likely to describe this as the proper motivation.

There are also many social values and norms that reflect what we care about that are acquired in similar ways to the cares we have for our loved ones. For example, many of the norms that govern how we navigate the social world are not reflectively chosen but rather adopted by us unreflectively when we are young and then internalized through a process of reward (Carruthers, 2024; Kelly, 2022). Some social norms may be rooted even more deeply, such as those governed by disgust responses toward incest; in these cases, we are apt to feel that certain things are “just wrong” and will be strongly motivated to avoid them (Haidt 2001; Kelly, 2011). These norms motivate strong responses in us when they are violated or otherwise harmed, and we are often moved by them even when we explicitly disavow their importance.

This might strain credulity, since it suggests that virtually any value can reflect what we care about regardless of our control over it, so long as it manifests the right kind of motivational/affective profile. Someone might argue that if these kinds of values can reflect what we care about despite being outside our control, then even our biologically prescribed values express what we care about. Though this sounds problematic, there are cases, which suggest that biological values can express what we care about. Consider that though social expectations play a role in shaping some of the affective attitudes parents have toward their children, it is reasonable to think that some of these affective responses are biologically prescribed. However, we do not think that just because some of our affective responses are so prescribed we care about our children any less, or that our love for them is any less a reflection of what matters to us.²² Likewise, when someone is suffering from starvation, it makes perfectly reasonable sense to say that they care about procuring food despite this care being grounded in a care for something like homeostasis (Arpaly & Schroeder, 2014).²³

These observations suggest that many of our cares are developed via processes that do not require reflection over what we care about, and therefore when we act on them, we exhibit animal freedom. Additionally, many – perhaps some of the best – actions that reflect these cares are similarly unburdened by our capacity to reflect over what we care about. This suggests that a large and important part of our lives as agents who act in ways that reflect what we care about is – or can reasonably be – made up by animal freedom. This gives us a good reason to care about animal freedom, since it is the kind of freedom we manifest in service of many of the things we care about most.

5. Objections

5.1. *Are all desires cares?*

Does this account make freedom too widespread by reducing the constraints on caring attitudes to include exclusively first-order conative states? Put another way, can we embrace the idea that cares can be first-order states without endorsing the idea that every first-order desire is a care?

There seem to be two sets of concerns here. The first is that my account suggests that even mere preferences, for example, my desire for chocolate over vanilla ice cream, might be considered cares. The second is that if cares are first-order desires, we collapse the distinction between desires that are important to us – in the sense that they form our deep selves or reflect what matter to us – and those that are not, but such a distinction is important for understanding ourselves and others as valuing agents.

Many preferences are not indicative of our care for the preferred object, but rather are something like realizer desires (Arpaly & Schroeder, 2014). Consider my preferences for some flavors of ice cream over others. I have a weak preference for chocolate over vanilla, and I am therefore slightly disappointed whenever I get vanilla instead of chocolate. But I am also happy all things considered to order vanilla and nonetheless am very happy with the vanilla as I'm eating it. One way to think of this case is that both chocolate and vanilla are good paths to realizing what I care about – perhaps just ice cream or dessert, or, more generally, fatty and sweet foods. If that were so, then a mere preference for chocolate over vanilla does not indicate that I care about chocolate.

However, some preferred things might also be things we care about, though very little. For example, it might be the case that I have a preference for chocolate ice cream over vanilla because I have a standing value for chocolate that does not exist for vanilla; in this case, I value chocolate for its own sake, but value vanilla ice cream only because of its relevance to something else I care about.²⁴ However, I might care about chocolate very little, as measured by the emotional resonance that the availability of chocolate has for me, the amount of time and attention I devote to it, and the motivational force the care exerts, especially in situations where it would be costly to obtain chocolate over vanilla.²⁵ Both in terms of its intrinsic dispositional properties to motivate me to act, and in terms of its relevance among my set of cares more broadly, my care for chocolate may be comparatively unimportant. But this would not change the fact that there exists a state within my mental economy capable of providing intrinsic motivation for action that also generates affective responses based on the fortunes of chocolate.

It makes sense to think we can care about things while caring about them very little. Ordinarily when we commit to the goodness of something, we do

not thereby come to care about it for its own sake very much. Rather, the commitment tends to motivate us by virtue of tapping into other things that we care about, such as being people who do what they think is right, or who follow through on our commitments (Carruthers, 2024; Kelly, 2022). This is why, for example, it is difficult to convert to vegetarianism. If we are raised to believe that animals have no intrinsic value or moral standing, then we are unlikely to care about their wellbeing or deaths. When we learn that eating meat is wrong for reasons of animal welfare, we form new beliefs about the permissibility of eating meat, but the change in our underlying set of cares is likely minimal (Arpaly & Schroeder, 2014; Carruthers, 2024). Once a care is generated, it remains dispositionally weak and provides only a very small amount of motivational force relative to other values. This is part of the explanation for why it can be so difficult to act in accordance with newly acquired norms. Without motivational scaffolding provided by other things we care about, like the value we place in being good people or in doing the right thing, we are unlikely to be sufficiently motivated to act in accordance with our commitments (Kelly, 2022).

Not all desires are cares, then, since not all desires possess the affective-motivational profile that cares have. Some have argued that this means that cares are relatively rare, since most of our actions can be explained by reference to a very small set of cares (Sripada 2016a). While it is true that cares are less numerous than the many desires we have, and that most of our actions may be motivated by instrumental desires, it may be the case that cares are still more numerous than we think. To see why, note that while most actions are instrumentally motivated, it is also true that things that we valued only instrumentally at one point in time often come to be valued for their own sake down the road (Carruthers, 2024). When we try a new food to be adventurous, we may dislike the taste but enjoy the experience because it serves our standing care to be adventurous. Over time, and with enough positive reinforcement, we may come to find the food desirable for its own sake, and not merely for the sake that it represents something adventurous. The same is true for many things that are initially valued only for the sake of something else.

5.2. *Concerns about attributability*

A second objection one might have to my view is that it identifies us too closely with attitudes whose contents we do not endorse. Implicit biases are an instructive example. These biases influence reason and attention outside of our awareness, such that things can appear to be reasonable and in line with our reflective commitments while being grounded in the influences of biases we would not endorse (Gurrola, 2023). They also dispose us toward reflexive affective reactions and judgments we would not endorse that

inform our behavior. On the story I've told, at least one type of implicit bias, affective bias, might be considered a care if it exerts strong enough motivational and attentional force. Is this a problem?

My initial thoughts here are that this is not a problem; recall that one way to think about what we care about is to consider the affective responses we have toward things. If we are consistently afraid of people of color, or are consistently hateful toward women, this suggests that we have a care toward those things that takes them to be bad, perhaps by virtue of their belonging to our outgroup. Therefore, at least in some cases of implicit bias, our biases do reflect what we care about. However, these biases may not reflect the extent of our caring attitudes toward justice, such that we can say of our implicitly biased friend that they do not care about the wellbeing of all people or egalitarianism. It is plausible that people who genuinely care about these things also sometimes act in implicitly biased ways and that their biased actions are motivated by cares. In such cases, their implicitly biased actions nevertheless reflect what they care about, though they may not reflect all of what they care about.

5.3. *Are persons special?*

Philosophers are generally committed to the idea that we have a kind of freedom that other animals lack. This kind of freedom is what supports holding persons responsible while excluding other animals and very young children from the class of morally responsible agents. My view suggests that animal freedom is shared between persons and other animals. Does this mean that there is nothing special about the kind of free will that persons have?

While there is nothing special about the kind of free will persons have, they do possess two capacities for augmenting their capacity for animal freedom (that is, for acting in accordance with what they care about). This does make them special kinds of free agents, and perhaps special in the manner required for moral responsibility. However, this does not make them fundamentally different kinds of free agents, since the kind of freedom they enjoy, even when it is facilitated by their additional capacities, is the very same kind of freedom that other animals enjoy.

What makes persons the special kinds of free agents they are is not that they can act in accordance with what they care about, since this is shared with other animals, but that they can act in accordance with other kinds of attitudes whose contents fail to be reflected in their set of cares. These are attitudes like commitments and endorsements. Our capacity to be motivated by these attitudes, when their contents are not reflected in our set of cares, is what underlies what I will call the Freedom of Persons. This capacity is distinctive of humans because

other animals appear to lack the reflective capacities required to form commitments and endorsements. This capacity affords us a distinctive form of free will because it augments our capacity to act in accordance with what we care about.

Unlike other animals, we are capable of recognizing that something is good or otherwise of value when we do not yet care about that thing. When we endorse the goodness or value of something, or commit to its preservation and well-being, these attitudes cue related things that we care about, and these provide us with motivation to act on our commitments and endorsements. This process allows us to bridge the gap between what we care about and what we believe we should care about but do not yet care about. It therefore allows us to exert some rational control over our actions, even when the ultimate motivational force of our actions always comes from what we care about.²⁶

Though the freedom of persons is what distinguishes us from other animals, and is therefore interesting in its own right, it is not the only kind of freedom we should be interested in. To see why, consider that the value of the freedom of persons depends on the value we place in acting in accordance with what we care about, or animal freedom. We care about having the freedom of persons because of what it allows us to do in terms of animal freedom. (Why do we value our capacity to act in accordance with our commitments that fail to be reflected in our set of cares? Presumably, part of the answer is that it facilitates our ability to act in accordance with something we do care about, and we care about acting in accordance with what we care about.) (Enoch, 2022)

Moreover, it is not as if the freedom of persons is a *sui generis* kind of freedom unbounded by animal freedom. Rather, when we form a commitment or endorsement, we do so only in light of other things that we care about. This means that the freedom of persons is made possible by and is a manifestation of animal freedom. While it will be interesting to investigate the links between the freedom of persons and animal freedom, there is no talking about the former without also talking about the latter.

Finally, as I discussed in [section 4](#), our actions often reflect cares that we developed unreflectively and that are largely recalcitrant to reflective judgments about their value. Many of these cares support some of our most important relationships and ways of interacting with the world. It would be one thing if the most important parts of our lives as caring agents were grounded in cares that reflect the freedom of persons (in such a case, it would make sense to prioritize the freedom of persons as the “interesting” kind of freedom). But this isn’t the case, and therefore in understanding the kind of valuing agents we are, understanding animal freedom is just as important as understanding the freedom of persons.²⁷

Notes

1. For reasons of space, I limit my discussion here to compatibilist accounts of freedom of the will that focus on sourcehood, since they do not require any special kind of agent causation. Other compatibilist accounts which focus on control, such as Fischer and Ravizza's Reasons-Responsiveness account, also contain criteria that exclude other animals from the set of free agents. For an account of incompatibilist freedom that attributes freedom to both humans and other animals based on a special kind of agent-causation, see Steward (2012).
2. According to Frankfurt, the class of wantons includes agents who possess the capacities required to form desires of the second order, but who fail to exercise such a capacity because they do not care what they will.
3. One potential exception is the emerging literature on animal agency that argues for the capacity of animals to behave rationally (i.e., for reasons) and not as mere automata (Melis & Monsó, 2024; Sebo, 2017). This work is important in its own right. However, philosophers in the compatibilist tradition usually do not take the capacity to act for a reason to be sufficient for acting freely. For example, a person could be in a state of mania and still be capable of manifesting means-end reasoning. Despite their rational behavior, we would not consider such a person to be acting freely. Therefore, it is reasonable to think that something more should be said in support of the idea that the rational nature of other animals supports attributions of free will. In arguing for the free will of other animals, I take myself to be contributing to the broader literature on animal agency.
4. This is a controversial claim, and there is an ongoing debate about the extent to which other animals possess metacognitive capacities (Andrews, 2020; Carruthers & Williams, 2019; Proust, 2013). However, as the goal of this article is to argue for the possibility of caring without the capacity for reflective endorsement, it is a claim I accept solely for the purposes of argument. What I aim to show is that regardless of whether other animals possess the capacity for reflective thought, they can act in accordance with what they care about and therefore act freely.
5. For example, a dog who consistently braves a rushing river to save his owner might reasonably be considered courageous, and therefore his actions would reflect this fact about him. For other examples, see King (2023).
6. While some philosophers discuss caring about something as being equivalent to its mattering to us (Sripada 2016b), this is not true of every discussion of the nature of cares. However, it is reasonable to think that caring about something implies that it also matters to us, and therefore I will move between these two ways of talking about caring in what follows.
7. Accounts of care often introduce other conditions, such as the motivational, commitment, and evaluative conditions proposed by Sripada (2016b), or the condition that caring agents be able to grasp, if only implicitly, the notion of importance (Jaworska, 2007).
8. This final condition is arguably less central than the others, since it will not be met in cases where we care about something very little. For example, when we are forming a new care, it is unlikely that such a state will motivate very costly behaviors, though it may motivate us to costlier behaviors than would a merely instrumental desire.
9. This is not to say that what we care about is never within our evaluative control. Creasy (2025) makes the point that both evaluative commitments and strong affective responses can reflect what we find meaningful, since both contribute to a meaningful life, though in different ways.

10. There is another kind of commitment, which is commitment to act a certain way (Carruthers, 2015, 2024). These kinds of commitments do provide intrinsic motivation and therefore get us closer to the idea that our commitments become reflected in our cares. However, it is less clear that they generate the kind of affective tethering required for the commitment to be considered a care.
11. There is debate about the extent to which norm-guided behavior is the product of purely instrumental motivation or intrinsic motivation that involves valuing the norm for its own sake (Kelly et al., 2025). Care-concordant action is a rather broader category than norm-concordant action, insofar as it involves not only internalized rules about the social world but also intrinsic desires for valued objects. However, the question of how a new value motivates is relevant to the best explanations of norm-concordant action *and* care-concordant action more generally. The view I advance here adopts a form of pluralism about the motivations involved in care-concordant action, with the thought being that newly acquired cares depend upon the motivational force of internalized cares for their motivational force, suggesting that these actions are motivated by the newly acquired care instrumentally or as realizer actions, while once internalized, newly acquired cares can motivate in their own right.
12. This is consistent with our being understanding toward them, even if we assume that being vegetarian is morally required. New cares are difficult to develop, and old cares are difficult to change. Even if we ourselves are dedicated vegetarians – or perhaps, because we are – we likely understand how difficult it can be to make the change to vegetarianism. We can therefore take the action to be attributable to the agent and therefore a freely willed action, since it expresses something they care about, while understanding that the project they've undertaken is not an especially easy one and therefore modulate our reactive attitudes toward them.
13. Jaworska is open to the possibility that her account leaves open that some very advanced animals like chimpanzees could care. However, the evidence I present here suggests that the kinds of emotional capacities that support caring responses are found even in animals like rats, which many typically consider to be relatively simple animals.
14. Most of the evidence I discuss here has previously been assembled by other philosophers, though for different purposes. I am especially indebted to Monsó et al. (2022) and Benz-Schwarzburg and Wrage (2023) for their discussions of caring behaviors in other animals.
15. Other evidence may also be found in research on long-term mating pairs of animals who are socially, though not sexually, monogamous. As King (2013) persuasively argues, we have no more reason to think that other animals who are socially monogamous but who engage in EPCs (extra-pair copulations) fail to care for their social mate than we do for thinking that humans who exhibit caring bonds but who engage in affairs outside their monogamous relationships fail to care for their partners. This is consistent, of course, with their failing to *behave* in a caring way toward their partners by failing to keep a promise of sexual fidelity.
16. Seyfarth and Cheney (2012) provide persuasive arguments for thinking that the friendships they discuss are not best explained by the Current Needs hypothesis, which claims that the apparently extended social bonds actually comprise nothing more than token events of reciprocity that meet its members' current needs.
17. The prevalence of bonds forming between mother-daughter pairs is not exclusive to baboons, with research indicating that this is the most reliable predictor of bonds in African elephants (Moss et al. 2010) among other animals.

18. Knowledge of these social categories may operate without any explicit conceptual representation of higher and lower status.
19. For extended discussion and other examples of grief behavior in nonhuman animals, such as elephants, see King (2013).
20. In addition to helping others and themselves when they could choose only to help themselves, rats have also been shown to help another rat escape a distressing situation even when they could escape instead (Carvalho et al., 2019). This observation is in line with other examples that suggest that many kinds of nonhuman animals will come to the aid of conspecifics even when it comes at a cost to their fitness (Waal & Frans, 2008).
21. All of this is consistent with the idea that there are also many relationships that we value because of the work and reflection we have had to put into them. Thanks to Chris Masciari for bringing this point to my attention.
22. An interesting example of the interaction between biological and social values and the way they inform what we care about is the case of adoptive parents. Research on the differences in parental attitudes toward and investment in biological versus adopted children suggests that the desire to parent and invest in one's child extends reliably to non-related children. Studies indicate that adoptive parents invest just as much in terms of time and resources in their children as biological parents (Hamilton et al., 2007). This suggests that the mechanisms underlying biologically prescribed forms of care can ground caring attitudes even when a genetic bond is absent between parent and child. This might go some way toward assuaging the concern that the caring attitudes manifested in biological relationships are not genuine because they are merely a product of kin selection.
23. There are independent reasons to think that higher-order reflection is the wrong place to look for control. Rowlands (2012) argues that though philosophers often appeal to the exercise of higher-order reflective processes as constituting the kind of control over our motivations that makes us moral agents, this cannot be correct. Consider that the reasons represented in higher-order conscious reflection, against which we test our first-order motivations when deciding whether to endorse them, are not themselves subjected to the same kind of scrutiny – that is, they are not themselves represented in a manner that allows them to be tested against further reasons. Rather, their authority to reflect what we consider reasonable is taken for granted. However, as should be obvious from research on the motivational impacts of implicit attitudes and situational factors, these attitudes can fail to be under our control in important respects (Doris, 2002; Gurrola, 2023). For the extent to which they are accessible, as well as the extent to which we consider them reliable, can also be due to situational factors external to us, or to psychological factors whose influence we would not endorse. This means that the second-order attitudes that are supposed to imbue us with control over our first-order attitudes also fail to be in our control. It is therefore a mystery how the exercise of higher-order thought is supposed to imbue our motivations with the kind of control required for moral agency. Rowlands labels the reliance on higher-order processes for explaining control the “miracle-of-the-meta” fallacy, since it assumes that there is something intrinsic to higher-order processes that affords us a special kind of control over our motivations.
24. I might still be OK with ordering vanilla if the restaurant is out of chocolate, but doing so would fail to satisfy one care – that for chocolate – though it might help to realize another, such as my intrinsic desire for sweets.
25. We can imagine that, were I told that the restaurant had only vanilla but that I could obtain chocolate by going to the shop next door, and that I left and returned a few

minutes later with my cone of chocolate, that we would think I care about chocolate ice cream.

26. The person who does this may be irrational in other ways. For example, it may be irrational to continue to pursue a very costly goal that one does not care about just because one has committed to it. Thanks to Chris Masciari for bringing this to my attention.
27. I am grateful to Peter Carruthers, Eric Sidel, Lizzie Schechter, Rachel Singpurwalla, Kaity Creasy, Chris Masciari, Lia Curtis Fine, Robert Ragsdale, Xiaohui Yu, and two anonymous referees from this journal for comments on earlier versions of this article. I am also grateful to Chandra Sripada for helpful correspondence about the nature of cares, and to the participants of the University of Maryland Philosophy Department Work in Progress Workshop for helpful comments on an early presentation of the ideas in this article. Thanks to Tyler Loveless for helpful discussion and for advising on the editorial process. Finally, thanks to Vanessa Singh for extensive discussion over the course of many months on the nature of agency in humans and other animals, and to all the animals whose behavior inspired these conversations.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by the Ann G. Wylie Dissertation Fellowship, awarded by the [Graduate School at the University of Maryland], College Park.

References

- Andrews, K. (2020). *The animal mind: An introduction to the philosophy of animal cognition*. Routledge.
- Arpaly, N. (1999). Praise, blame and the whole self. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 93.2(2), 161–188. <https://doi.org/10.1023/A:1004222928272>
- Arpaly, N., & Schroeder, T. (2014). *In praise of desire*. OUP USA.
- Bartal, I. B.-A., Decety, J., & Mason, P. (2011). Empathy and pro-social behavior in rats. *Science*, 334(6061), 1427–1430.
- Benz-Schwarzburg, J., & Wrage, B. (2023). Caring animals and the ways we wrong them. *Biology & Philosophy*, 38.4(4), 25. <https://doi.org/10.1007/s10539-023-09913-1>
- Betzler, M. (2023). Responsibility and the Deep self. In M. Kiener (Ed.), *Routledge Handbook of philosophy of responsibility* (pp. 177–189). Routledge.
- Burkett, J. P., Andari, E., Johnson, Z. V., Curry, D. C., de Waal, F. B. M., & Young, L. J. (2016). Oxytocin-dependent consolation behavior in rodents. *Science*, 351(6271), 375–378. <https://doi.org/10.1126/science.aac4785>
- Carruthers, P. (2015). *The centered mind: What the science of working memory shows us about the nature of human thought*. OUP Oxford.
- Carruthers, P. (2024). *Human motives: Hedonism, altruism, and the science of affect*. Oxford University Press.

- Carruthers, P. (2025). A content-general adaptation for tribal value-acquisition. *Evolution and Human Behavior*, 46.6(6), 106791. <https://doi.org/10.1016/j.evolhumbehav.2025.106791>
- Carruthers, P., & Williams, D. M. (2019). Comparative metacognition. *Animal Behavior and Cognition*, 6.4(4), 278–288. <https://doi.org/10.26451/abc.06.04.08.2019>
- Carvalho, J., Seara-Cardoso, A., Mesquita, A. R., de Sousa, L., Oliveira, P., Summavielle, T., & Magalhães, A. (2019). Helping behavior in rats (*Rattus norvegicus*) when an escape alternative is present. *Journal of Comparative Psychology*, 133(4), 452. <https://doi.org/10.1037/com0000178>
- Creasy, K. (2025). Meaning at the limits of practical agency. *Philosophical Psychology*, 39(1), 1–23. <https://doi.org/10.1080/09515089.2025.2510559>
- De Waal, F. (2007). *Chimpanzee politics: Power and sex among apes (25th anniversary)*. Johns Hopkins University Press.
- Doris, J. M. (2002). *Lack of character: Personality and moral behavior*. Cambridge University Press.
- Doris, J. M. (2015). *Talking to our selves: Reflection, ignorance, and agency*. OUP Oxford.
- Enoch, D. (2022). Autonomy as non-alienation, autonomy as sovereignty, and politics. *The Journal of Political Philosophy*, 30.2(2), 143–165. <https://doi.org/10.1111/jopp.12264>
- Frankfurt, H. (1971). *Journal of Philosophy*, 68(1), 5–20.
- Frankfurt, H. G. (1969). Alternate possibilities and moral responsibility. *The Journal of Philosophy*, 66(23), 829–839. <https://doi.org/10.2307/2023833>
- Gurrola, J. (2023). The importance of self-knowledge for free action. *European Journal of Philosophy*, 31(4), 996–1013. <https://doi.org/10.1111/ejop.12812>
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108(4), 814.
- Hamilton, L. (2007). Adoptive parents, adaptive parents: Evaluating the importance of biological ties for parental investment. *American Sociological Review*, 72(1), 95–116. <https://doi.org/10.1177/000312240707200105>
- Hernandez-Lallement, J., van Wingerden, M., Schäble, S., & Kalenschner, T. (2016). Basolateral amygdala lesions abolish mutual reward preferences in rats. *Neurobiology of Learning and Memory*, 127, 1–9. <https://doi.org/10.1016/j.nlm.2015.11.004>
- Jaworska, A. (2007). Caring and internality. *Philosophy and Phenomenological Research*, 74.3(3), 529–568. <https://doi.org/10.1111/j.1933-1592.2007.00039.x>
- Kelly, D. (2011). *Yuck!: The nature and moral significance of disgust*. MIT Press.
- Kelly, D. (2022). Two ways to adopt a norm: The (moral?) psychology of internalization and avowal. In M. Vargas & J. M. Doris (Eds.), *The Oxford handbook of moral psychology* (pp. 285–309). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198871712.013.17>
- Kelly, D., Westra, E., & Setman, S. (2025). The psychology of normative cognition. In Edward N. Zalta & Uri Nodelman (Ed.), *The stanford encyclopedia of philosophy* (Summer 2025 Edition). Stanford, CA: Metaphysics Research Lab, Stanford University. <https://plato.stanford.edu/archives/sum2025/entries/psychology-normative-cognition/>
- King, B. J. (2013). *How animals grieve*. University of Chicago Press.
- King, B. J. (2019). The ORCA'S SORROW. *Scientific American*, 320(3), 30–35. <https://doi.org/10.1038/scientificamerican0319-30>
- King, M. (2023). *Simply responsible: Basic blame, scant praise, and minimal agency*. Oxford University Press.
- King, M., & Carruthers, P. (2012). Moral responsibility and consciousness. *Journal of Moral Philosophy*, 9.2(2), 200–228. <https://doi.org/10.1163/174552412X625682>

- Liberti, M. (2026). Neoptolemus and Huck Finn reconsidered. Alleged inverse akrasia and the case for moral incapacity: M. Liberti. *Journal of Value: Inquiry*, 60(2), 395–411.
- Melis, G., & Monsó, S. (2024). Are humans the only rational animals? *The Philosophical Quarterly*, 74.3(3), 844–864. <https://doi.org/10.1093/pq/pqad090>
- Mitchell-Yellin, B. (2025). On Valuing. In D. Shoemaker (Ed.), *Oxford Studies in Agency and Responsibility* (Vol. 9, pp. 6–27). Oxford: Oxford University Press. <https://doi.org/10.1093/9780198955498.003.0002>.
- Monsó, S. & Andrews, K. (2022). Animal moral psychologies. In M. Vargas & J. M. Doris (Eds.), *The Oxford handbook of moral psychology* (pp. 388–420). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780198871712.013.22>.
- Moss, C. J., Croze, H., & Lee, P. C. (Eds.). (2010). *The Amboseli elephants: A long-term perspective on a long-lived mammal*. University Chicago Press.
- Proust, J. (2013). *The philosophy of metacognition: Mental agency and self-awareness*. OUP Oxford.
- Rowlands, M. (2012). *Can animals Be moral?* Oxford University Press.
- Scanlon, T. M. (1998). *What we owe to each other*. Cambridge, MA: Belknap Press of Harvard University Press.
- Sebo, J. (2017). Agency and moral status. *Journal of Moral Philosophy*, 14(1), 1–22. <https://doi.org/10.1163/17455243-46810046>
- Seyfarth, R. M., & Cheney, D. L. (2012). The evolutionary origins of friendship. *Annual Review of Psychology*, 63(1), 153–177.
- Shoemaker, D. (2015a). Ecumenical Attributability. In R. Clarke, M. McKenna, & A. M. Smith (Eds.), *The Nature of Moral Responsibility: New Essays* (pp. 115–140). Oxford: Oxford University Press.
- Shoemaker, D. (2015b). *Responsibility from the margins*. Oxford University Press (UK).
- Shoemaker, D. W. (2003). Caring, identification, and agency. *Ethics*, 114(1), 88–118. <https://doi.org/10.1086/376718>
- Smith, A. M. (2005). Responsibility for attitudes: Activity and passivity in mental life. *Ethics*, 115(2), 236–271. <https://doi.org/10.1086/426957>
- Smith, A. M. (2008). Control, responsibility, and moral assessment. *Philosophical Studies*, 138(3), 367–392. <https://doi.org/10.1007/s11098-006-9048-x>
- Sripada, C. (2016a). Free will and the construction of options. *Philosophical Studies*, 173(11), 2913–2933. <https://doi.org/10.1007/s11098-016-0643-1>
- Sripada, C. (2016b). Self-expression: A deep self theory of moral responsibility. *Philosophical Studies*, 173(5), 1203–1232. <https://doi.org/10.1007/s11098-015-0527-9>
- Steward, H. (2012). *A metaphysics for freedom*. Oxford University Press.
- Vincent, S., Ring, R., & Andrews, K. (2018). Normative Practices of Other Animals. In A. Zimmerman, K. Jones, & M. Timmons (Eds.), *The Routledge Handbook of Moral Epistemology* (pp. 57–83). New York: Routledge.
- Waal, D., & Frans, B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, 59(1), 279–300. <https://doi.org/10.1146/annurev.psych.59.103006.093625>
- Watson, G. (1975). Free agency. *The Journal of Philosophy*, 72(8), 205–220.
- Watson, G. (2004). *Agency and answerability: Selected essays*. Clarendon Press.
- Westra, E., & Andrews, K. (2022). A pluralistic framework for the psychology of norms. *Biology & Philosophy*, 37(5), 40. <https://doi.org/10.1007/s10539-022-09871-0>
- Williams, B. (1981). *Moral luck: Philosophical papers 1973–1980*. Cambridge University Press.

Wolf, S. (1987). Sanity and the Metaphysics of Responsibility. In F. D. Schoeman (Ed.), *Responsibility, Character, and the Emotions: New Essays in Moral Psychology* (pp. 46–62). Cambridge: Cambridge University Press.

Wolf, S. (2015). Responsibility, moral and otherwise. *Inquiry*, 58(2), 127–142. <https://doi.org/10.1080/0020174X.2015.986852>