Cooperation, Cognition, and the Elusive Role of Joint Agency

Stephen M. Downes, Department of Philosophy, University of Utah

Patrick Forber, Department of Philosophy, Tufts University

Joshua Shepherd, Department of Philosophy, ICREA/Universitat Autònoma de Barcelona

Abstract: We propose an approach to the evolution of joint agency and cooperative behavior that contrasts with views that take joint agency to be a uniquely human trait. We argue that there is huge variation in cooperative behavior and that while much human cooperative behavior may be explained by invoking cognitively rich capacities, there is cooperative behavior that does not require such explanation. On both comparative and theoretical grounds, complex cognition is not necessary for forms of joint action, or the evolution of cooperation. As a result, promising evolutionary approaches to cooperative behavior should explain how it arises across many contexts.

1. Introduction

How is complex, coordinated social action possible? The question can be pursued from a philosophical perspective by analyzing the mental states, reasons, or commitments required to distinguish genuine cases of group action or joint agency from individual action or agency. The question can also be pursued from an evolutionary perspective by identifying the historical factors responsible for a transition from solitary to group living, or from a competitive and individualistic life way to a more social and cooperative one. These two approaches generate a potential conflict: philosophical approaches tend to focus on the conditions (some constellation of necessary or sufficient) that distinguish true joint agency from individual agency, while evolutionary approaches provide explanations for the origins of complex traits and behaviors, and these explanations involve tracing transitions from ancestral to descendant traits. To put it bluntly, philosophy identifies bright lines between joint and individual agency, whereas evolution shows how those lines are crossed, often in messy and complicated ways.

If we grant that humans can and often do engage in joint agency, we can ask whether this capability is unique to our lineage, or whether we can find precursors, homologies or analogies in other branches of the tree of life. Articulating this hypothesis in a testable way requires bridging the gap between "thinner" evolutionary accounts of the strategic dynamics of cooperative behavior and "thicker" philosophically loaded, cognitively rich accounts of cooperation that invoke uniquely human cognitive capacities. In effect, we must tackle the thorny question whether, and to what extent, social animals engage in complex, joint agential behavior, and what cognitive capacities these animals possess to enable their behavior. We suspect that this gap presents a more significant hurdle than commonly supposed.

In what follows, we defend a continuity thesis with respect to the cognitive capacities implicated in joint agential cooperation: the features of cognition that allow individuals to engage in joint agency are assembled incrementally over evolutionary time and we should expect to see some of these features occurring in non-human animals. In contrast to more exclusive "bright line" views of joint agency, we argue that our approach fits better with an evolutionary perspective on social behavior and the cognitive capacities that produce it. [1] The contrast helps identify the standards any evolutionary account of joint agency needs to meet to provide a viable explanation for the emergence of prosocial cooperative behavior. To make our case we focus on an exemplar of complex, coordinated social action: group hunting and foraging. This case is particularly important because many have invoked group hunting and foraging as a crucial activity for hominin social and cognitive evolution (Sterelny 2012; 2021; Tomasello et al. 2012; Tomasello 2014), and cooperative hunting can be found in many non-human animals, providing a good basis for a comparative analysis. In section 2 we start by taking a closer look at complex hunting tasks in two social mammals, wolves and dolphins.

The complexity and fluidity of social interactions found in these lineages suggests that forms of joint agency may arise in different ways and exhibit subtle differences across species that become relevant for understanding the relationships between the variety of evolutionary pathways that produce cooperative behaviors. This raises a worry about the fruitfulness of any

---

[1] "Bright line" accounts invoke specific criteria about the intentional states and rationality of the subjects involved and this characterizes much of the work in philosophy of action (see, e.g., Tuomela and Miller 1988; Bratman 1992; Gilbert 1992; 2009; McMahon 2005; Pettit and Schweikard 2006; Rovane 2019; see also Schweikard and Schmid 2021 for discussion).

approach that aims to explain how some lineage somehow crossed a "bright line" and became capable of true joint agency (as opposed to some lesser or simpler form). For, as we discuss in section 3, there are several different explanatory dimensions of joint agency and cooperative behavior. Since these may map in different ways to the different forms of cooperative behavior and joint agency distributed in nature, any satisfying evolutionary explanation must be clear about the dimensions in view. In section 4 we give a comparative approach that we argue is better suited to studying the various forms of cooperation that arise in many animals. In section 5 we turn to the human case, and to identify the stakes in debates over human cognitive evolution and the emergence of joint agency, we review some approaches to the evolution of cooperation and sociality in the lineages leading to Homo sapiens. The continuity thesis we defend here puts pressure on claims that the capacity to engage in joint agential behavior is uniquely human, and our argument connects to previous continuity arguments about language (Moore 2017) and culture (Whiten et al. 2017).

2. Group hunting in wolves and dolphin

While group hunting in early Homo sapiens and late hominins (such as Neanderthals) involved a unique level of sophistication—refined weapons, communication, planning, complex strategy—group hunting behaviors occur widely across the mammalian class. When we explain these behaviors, and make inferences about the cognitive capacities of the individual animals engaged in these behaviors, there is a broad question we can ask: are these behaviors merely the aggregative product of individual decisions, or do they require recognition and coordination between members of the group? To put it in the framework of the current investigation, do non-human mammals exhibit a kind of joint agency when hunting in groups? It could be that what

looks like group hunting is best explained by opportunistic individuals joining an ongoing activity to maximize their own chances of filling their bellies. On this opportunistic recruitment interpretation, recognition and coordination between members of the group, if it occurs at all, would amount to an awareness that other competitors are near and potentially a threat to an individual's take. Alternatively, individual animals may be coordinating with each other in a way that produces a group behavior that requires an appeal to some sort of joint agency. Evidence that animals recognize one another, fill different strategic roles in the hunt, respond fluidly to the actions of both conspecifics and prey, and communicate in ways that direct the hunt would support a joint agency hypothesis for group hunting in these species. Such evidence should also show that group hunting behavior cannot be easily accounted for by simply aggregating individual, context-independent, and evolutionarily selfish decisions.

We suspect forms of joint agency occur widely in mammals, particularly in group hunting scenarios. To support this suspicion, we turn to two case studies. Taking a closer look at hunting in wolves and dolphin show that there is convincing evidence that supports genuine cooperation, and thus a form of joint agency, over opportunistic recruitment in some of their more complex foraging behaviors.

The reintroduction of wolves (*Canis lupus*) into Yellowstone National Park created a new study population, providing a rich source of observational data on wolf social behavior. As a result, biologists now have an extensive and unique data set on wolf hunting behavior. Two key observational studies, one on wolves hunting elk (MacNulty et al. 2012) and another on hunting bison (MacNulty et al. 2014), provide some of the best evidence for evaluating whether, and to what extent, wolves need to work together as a pack to bring down the different prey types.

While wolves often hunt both prey types in packs, there are differences between how the wolves pursue the different prey. Elk are easier prey and a lone wolf can sometimes bring down an elk successfully. Wolves are much more likely to hunt in packs, and their chance of success when hunting elk increases markedly. However, there is a threshold size—four wolves—where increasing the group size of the hunting party no longer increases the chance of success; in fact, there is a statistical trend of declining success probability in bigger groups (MacNulty et al. 2012, 79). The explanation given for this trend is observed free-riding in larger groups. Individual contribution drops in larger hunting groups, and the drop is more dramatic for wolves lower in the dominance hierarchy or without cubs. The study ranked individual wolves by hunting prowess, and so could evaluate whether inept hunting groups, or interference from too many hunters might explain the trend. Neither hypothesis fares well given the data. The decline in success probability occurs faster in competent groups compared to inept groups, and there is no evidence of interference increasing across group sizes (MacNulty et al. 2012, 81).

Wolves obviously benefit from working together to hunt elk and such hunts involve fluid coordination of action as they pursue their prey, a feature we would expect in genuine cooperation. Yet this evidence is not decisive. The wolves are, in an evolutionary sense, pursuing their own interests. Joining a small group increases the chance of successfully capturing elk, but as the group gets larger individuals start to ride free and avoid the costs of the cooperative hunt. This suggests that even though wolves exhibit some group level cooperation, individual opportunistic recruitment to the hunt may be the best explanation for the pattern of behavior: cooperate when it makes a substantial impact or if in substantial need, otherwise ride free.

The situation changes when wolves hunt bison. Bison present a more significant challenge since they can easily injure or even kill attacking wolves. It is extremely rare for lone

wolves to approach bison, and MacNulty et al. (2014) made zero observations of a single individual successfully bringing one down. Perhaps unsurprisingly, more complex cooperation and pack interaction is necessary to hunt the larger, more dangerous prey. MacNulty et al. (2014) studied how group size affected hunting success for wolves pursuing bison. In this case, the probability of success scales with group size up to eleven wolves. After that threshold, probability success levels off but doesn't show the same decreasing trend observed for elk (MacNulty et al. 2014, 4). Also, wolves show restraint when in small groups. In groups of four wolves or less, the probability of starting pursuit of bison is much less (approximately half) than the probability of pursuit of elk (see Figure 3 in MacNulty et al. 2014).

The study attributes these trends to two features. First, given the very low probability of success for solo hunters or small groups, hunting bison requires more cooperative investment by the pack. The individual wolves track their group size and adjust their behavior (whether to pursue) accordingly. Second, since bison are more dangerous prey, the wolves adjust their spatial configuration and responsiveness to the pack during the hunt (MacNulty et al. 2014, 5-6). Video supplemental to the study shows an example of a successful bison hunt. The wolves surround a group of bison and take turns harassing individuals over several minutes. In response to the harassment, individual bison charge at the wolves and the group of bison eventually starts to splinter with some individuals moving away to avoid the wolves. The hunt ends when a bison calf panics and bolts—the wolves immediately pursue and bring down the calf in a matter of seconds.

Wolf hunting behavior in the bison case demonstrates a higher degree of cooperation and thus suggests at least some elements of joint agency (i.e., cognitive capacities and dispositions that could explain the cooperative behavior in view). The wolves show a tendency to avoid

starting bison hunts when groups are small; the success of the hunt scales with group size up to eleven wolves; there is no consistent trend of defection by individuals in larger groups; and their behavior during the hunt is more responsive to packmates. If we want to exclude this as a case of genuine joint agency, we need to articulate the appropriate criteria for drawing a principled line between (say) uniquely human joint agency and cooperative hunting among Yellowstone wolves. There is one criterion that is a good candidate for identifying the kinds of cognitive capacities that support joint agency, that would draw the right line between wolves and humans: communication between individuals before, during, or after the hunt. While wolves are obviously tracking their packmates and responding accordingly, there is no evidence that they are engaging in explicit signaling to coordinate the hunt. Requiring complex or linguistic communication for joint agency would limit its explanatory power by setting the bar for joint agency rather high—hominin ancestors were engaging in group cooperative activities long before there is evidence of language (Sterelny 2021; Planer and Sterelny 2021). Yet one might argue that explicit communication as essential to the planning and execution of genuine cooperation, or truly joint agential behavior. However, this creates difficulties because there are cases where we see both highly cooperative hunting and specific signaling between group members in several dolphin species.

Two studies done by Benoit-Bird and Au (2009a; 2009b) document a cooperative foraging strategy in spinner dolphins (*Stenella logirostris*) off the coast of Hawaii. The strategy involves large groups of dolphins (16-28) working together to herd fish before feeding. The overall behavioral strategy has four stages and requires careful coordination in four dimensions. Dolphins start with a wide line oriented like a V in the vertical dimension of the water column. When the group encounters a shoal of fish dolphins adopt a tight line formation to bulldoze the

8

prey. The group then transitions to a circle formation, though each dolphin remains at the same depth, effectively forming a cylinder around prey. The circle then tightens and pairs of dolphins, each from an opposing position of the circle formation, take turns darting into the circle to feed. As the biologists summarize:

> "Our data show that spinner dolphins worked collectively to achieve densities of prey that did not occur in the habitat in the absence of this dolphin behavior, likely resulting in increased feeding success despite the individual costs of herding and maintaining a prey patch and other potential costs of social feeding. There was no evidence of cheating by animals either within groups or from outside groups. All of this, when considered together, strongly supports cooperative rather than simply group foraging. The remarkable degree of coordination shown by foraging spinner dolphins, the very strict geometry, tight timing, and orderly turn taking, indicates the advantage conferred by this strategy and the constraints placed upon it" (Benoit-Bird and Au 2009a, 136).

There is also some evidence that there is significant acoustic coordination, perhaps even communication between the dolphins during the search and feeding. During transitions between stages there are spikes in echolocating clicks (Benoit-Bird and Au 2009b). The evidence for communication is not definitive because disentangling communication from individual echolocation can be tricky and the whistles that tend to be used for communication are absent during the hunt. It is not clear whether the dolphins are signaling or using each other's echolocation clicks as cues for the position of conspecifics. A further concern with the studies is that position information is inferred from sonar rather than directly observed. Nonetheless, the studies document large group hunting in a marine mammal that minimally involves complex

spatial and temporal coordination, and it coheres with observations of hunting and communication in other dolphin species.

Hamilton et al. (2022) document a hunting strategy in bottlenose dolphin (*Tursiops truncates*) off the Florida Keys that they label "driver-barrier feeding." In this strategy one dolphin acts as the "driver" and herds small prey fish towards a line of waiting dolphin, the "barrier," driving them into the air where the dolphin jump to catch them. This hunting strategy exhibits role-specialization and coordination among the group. The dolphin sort into a driver and a barrier, and the dolphin in the barrier stop using echolocating clicks—only the driver dolphin uses echolocation clicks during the hunt. In addition, there is clear evidence of communication: the driver dolphin whistles before beginning the hunt in a way that is statistically different from behavior in other contexts (Hamilton et al. 2022, Figure 3). The study suggests that whistles may play a role in signaling motivation, recruiting conspecifics or coordinating behavior (Hamilton et al. 2022, 6). All of these involve communication, and the pattern of behavior fits with the role whistling plays in dolphin generally, as Hamilton et al. note. Furthermore, there is also new evidence that dolphin tailor their communication to specific individuals: mother dolphins modulate their whistling when signaling to their own offspring rather than other dolphins in the pod, effectively using "motherese" when whistling to their own calves (Sayigh et al. 2023). In short, dolphins regularly communicate, and they do so during complex group tasks.

Evidence indicates, then, that sophisticated forms of cooperative behavior, highly suggestive of joint agency have evolved outside of the human lineage. The two cases we discuss here exhibit different forms of coordination deployed to achieve a joint action with the specific form of coordination facilitated by the affordances in each lineage. If we were to expand the scope of inquiry across the tree of life, we would find further gradations and variation in

coordinated behavior. This places pressure on any bright line accounts that aim to distinguish "true" joint agency, and the sophisticated cognition capacities that produce it, from the cooperative hunting behaviors in wolves and dolphin and the cognitive capacities implicated by their behavior. One might respond, of course, that philosophers and evolutionary biologists are simply pursing different explanatory projects.[2] This may be so, but this response leaves important questions unresolved regarding the forms (joint) agency involved in cooperative behavior may take, and what types of cognitive capacities might underlie these behaviors. We think these are questions worth answering as we argue in the next section.

3. Thick versus thin and the diversity of group action

We have argued that wolves and dolphins exhibit forms of joint agency, in the sense that their behavior requires recognition of others, communication, and coordinated, goal-directed adjustment of behavior in light of these. The following important questions arise at this point. What sort(s) of joint agency or cooperation are we studying in animals like wolves or dolphins? How does it compare to the form(s) of joint agency or cooperation we see in other animals,

---

[2] Some philosophical accounts are recruited to serve as evolutionary explanations, particularly for human social cognition. For instance, Tomasello (2019, 341) provides a "partial list of the most important uniquely human psychological outcomes." We think that several of the listed capacities (specifically joint attention, cooperative/referential gestures, conventional linguistic communication, instructed learning, coordinated decision-making, and joint commitment) are implicated in the examples of wolf and dolphin cooperation, and therefore put real pressure on his human uniqueness claim. We return to this point in section 5.

especially humans? How are we to explain the similarities and differences between forms of joint agency found across nature? Evolutionary explanations primarily seek to trace changes in forms of cooperative behavior, and the emergence of more sophisticated forms of joint agency. Such explanations have to be sensitive to these questions of categorization and variance (see Moore 2017).

The current literature offers no consensus on these questions: the diversity of group behavior is mirrored by a range of approaches to explaining joint agency and cooperation. But various approaches to categorizing forms of joint agency can be mapped along three orthogonal dimensions. The first concerns the level of sophistication of the joint agency or cooperation in view. One way to approach this issue is to focus on the complexity or flexibility of various cooperative behaviors. Much of the philosophical literature focuses less on this, and more on the conditions for 'bona-fide' joint agency (and therefore special, real, or truly collaborative cooperation) of the sort that one sees in paradigmatic human cases (see, e.g., Joyce 2006; Kaufmann 2015; 2016; Ludwig 2020). And some approaches seek to characterize more basic forms of joint agency and cooperation. Sometimes these are called 'proto' joint agency or 'minimal cooperation' (Paternotte 2014).

A second dimension concerns the cognitive capacities assumed or argued to be critical for joint agency or cooperation. Some approaches make minimal assumptions about the cognitive capacities of the agents involved. This is true of game-theoretic accounts, for example (Gold and Sugden 2007, see Section 4 below). Other approaches attempt to make explicit the cognitive capacities required. This is true of more sophisticated as well as more minimal accounts of joint agency. Butterfill and Pacherie (2020), for example, engage in an exercise of creature construction, and argue persuasively that forms of joint agency and cooperation could

emerge in an agent that lacked any capacity to attribute mental states, provided that the creature could track the goals of conspecifics, and provided that certain heuristics (e.g., shared perceptual mechanisms of salience-tracking as a proxy for common knowledge) typically hold. They then progressively add cognitive capacities to their creature, and they link these more sophisticated capacities to more sophisticated forms of cooperation.

A third dimension concerns a distinction between internalist and externalist approaches to explaining social phenomena in general. Internalist views appeal to internal items like mental states or representational contents (Bratman 1993; Gilbert 2009). Externalist views appeal to extra-mental items like laws and contracts, material resources, or the organizational structure of groups and group roles (Epstein 2015; Ritchie 2020; Westra and Andrews 2022).

We highlight these dimensions primarily to isolate explanatory emphases and strategies one finds in recent literature. To some extent, treatments of joint agency and cooperation may draw on all three dimensions. Both externalists and internalists, for example, can and sometimes do draw on accounts of the cognitive capacities involved in cooperation. And interestingly, philosophical accounts of sophisticated joint agency and cooperation tend to be heavily internalist, while approaches that seek to characterize more minimal forms of joint agency and cooperation often appeal, at least in part, to externalist features. Arguably, however, evolutionary explanation of different forms of joint agency and their relationships cannot afford to ignore any of the above dimensions. Indeed, in our view understanding these different dimensions of explanations of joint agency affords the ability to notice explanatory resources and opportunities that a reliance on leading philosophical accounts of bona-fide joint agency may hide.

If one follows philosophers who seek to explain the most sophisticated forms of joint agency, one might think that these most sophisticated forms are the most relevant target for

evolutionary explanation. We argue, instead, that from the perspective of evolutionary explanation, a normative dispute about the conditions for bona fide joint agency is less important than an investigation into the fitness of different forms of group activity in different types of circumstances, and how these may lead to the evolution of more sophisticated forms of sociality. Different forms of cooperation can be found across the animal kingdom (Moore 2018).

If joint agency and forms of cooperation vary across different dimensions, then the features and capacities constitutive of joint agency may vary across species, rendering the drawing of any bright line between bona fide and proto joint agency less than fruitful for evolutionary explanation. While philosophical accounts of joint agency can be useful for some purposes, our understanding of the evolution of coordination can benefit from more incrementalist (e.g., Moore 2017), and more pluralist, approaches.

4. An evolutionary approach to joint action

Evolutionary game theory provides a natural framework for analyzing joint action in a comparative way. Joint action, especially the sort we see in complex cooperation or divisions of labor, involves strategic interactions between multiple individuals. Game theory provides a way of representing the incentives of the participating individuals and, combined with a dynamics, identifies when cooperative behavior will invade, remain stable, or collapse. Most thinkers on human evolution have embraced game-theoretic characterizations of the cooperative and coordination challenges that faced human ancestors and that an appeal to joint agency purports to solve. One path towards defending the existence and importance of joint agency involves identifying complex coordination and cooperation problems that the capacity enables organisms to solve. Ideally, we would be able to characterize these as evolutionary problems imposed by

the challenges of an organism's social environment, and game theory provides abstract and general tools to do this.

That said, using the formal framework of evolutionary game theory immediately raises a problem for "bright line" accounts of joint agency: strategic interaction and cooperation are found across the tree of life from microbes to humans and sometimes these behaviors are mediated by complex cognition (e.g., wolves, dolphins) and sometimes not (e.g., microbes). So, some lineages can solve these problems without expensive cognition. Are there any features of the strategic set up that require cognition, or is it the case that some lineages have the expensive equipment for other reasons and simply deploy it to handle strategic interactions? In effect, we must confront a version of the problem that Gold and Sugden (2007) raised for philosophical accounts of shared or collective intentions that produce group action. They point out that, according to many philosophical accounts, two individuals engaged in Nash equilibrium behavior look like they possess a collective intention to produce the Nash equilibrium. At a Nash equilibrium players are making simultaneous best responses to the strategies of their interaction partners; thus, no one player can change their strategy and improve their specific payoff. What counts as "rational" action and a solution to the game depends on the strategy choices of all the players involved. This interdependence of action seems to meet the criteria for joint intentions or joint agency. Yet many "bright line" accounts would classify these cases involve a sort of pseudo-coordination and are not candidates for true group action, nor do they implicate joint agency despite the player's best strategy choice depending on the choices of other players.

Consider Gold and Sugden's key example: a simple competition over resources as represented by the Hawk-Dove game. In this game each player has a choice between an aggressive hawk strategy and a deferential dove strategy. Hawk beats dove and takes all the

resource; doves split the resource evenly; dove loses to hawk but avoids any cost or injury; and finally, hawk-hawk interactions lead to the worst outcome, an escalating conflict that risks injury. Against a dove, you would rather play hawk, but against a hawk, you prefer to play dove and avoid conflict. This game has three Nash equilibria—pairs of strategies such that each player chooses the simultaneous best response to their opponent. There are two pure Nash equilibria where one plays hawk and the other dove and one mixed where each player randomizes between hawk and dove in a way that makes their opponent indifferent between their strategy choices. Gold and Sugden (2007, 111-112) point out that pairs of individuals playing either of the pure Nash equilibria look to have a joint intention to bring about a group action: one player wins the resource and the other walks away with nothing. Given their opponent's play, each prefers their respective strategy of hawk and dove, and neither can do better by unilaterally changing their strategy. Thus, the pair seems to collectively intend that they play the Nash equilibrium and act on these intentions to bring about a joint action. Yet this seems strange, especially for the dove player who loses the resource.[3] Evolutionary dynamics provide further support for Gold and

---

[3] The presumption that "true" joint action involves cooperative or prosocial outcomes, while intuitive and often made in this debate, might itself be interrogated. The Hawk-Dove game is motivated by a different vignette in the economics literature that involves cooperating to shovel a snowdrift. The payoff ordering for the Snowdrift game is identical to the Hawk-Dove game, yet the motivation for exploring this game is whether cooperation—where both individuals play a dove-like strategy and help shovel—can be sustainable. The motivating vignette for a game shouldn't affect our analysis of whether joint agency is implicated. Deception is another interesting case. For instance, corvids routinely hide food when they suspect conspecifics are

Sugden's diagnosis, as neither pure Nash equilibrium is evolutionarily stable. The challenge Nash equilibria behavior poses for joint intention based accounts of group action is even more pressing for the classic game of cooperation, the (one-shot) Prisoner's Dilemma. This game has but one Nash equilibrium: both players defect. Yet accounts of joint action are usually invoked to explain how we can *cooperate* in these sorts of scenarios, not defect (see, e.g., Tuomela 1992).

Perhaps another game would better capture the nature of joint agency in cooperation and avoid this unpalatable consequence. Many evolutionary theorists prefer the Stag Hunt as a representation of early cooperation in human ancestral lineages (Skyrms 2004; Tomasello et al. 2012; Sterelny 2012). As Rousseau describes the scenario, each individual has a choice between cooperating with their fellows to bring down a stag or defecting to hunt rabbits on their own. The stag is the preferred quarry but your partner must also hunt stag. Hunting rabbit guarantees prey but of less value. In this game cooperation is risky but stable, if achieved. Many have argued that the mutual interest of cooperation captured by the Stag Hunt best characterizes the problems human ancestors faced with respect to collective foraging, defense, and reproductive cooperation. This led us to become obligate cooperators and produced the scaffolding necessary to solve more difficult cooperative problems. That the Nash equilibrium for hunting stag counts as joint action seems reasonable, for this game represents a paradigm coordinated activity: group hunting. However, there are more Nash equilibrium solutions to the game. In particular, hunting hare is a Nash equilibrium and this looks like the paradigm of individual, not joint, action.

---

watching (Bugnyar et al. 2016). While we suspect that this is not joint agency, it is a social action in response to another individual's attention that raises an interesting puzzle about where to draw boundaries.

Furthermore, simply solving a Stag Hunt is not yet evidence that any particularly sophisticated cognitive capacity is at work, as chimps seem to solve the game in a way different from humans (Duguid et al. 2014). Any account of joint action requires the resources to sort Nash equilibria behavior into true and spurious joint actions, if evolutionary game theory is to be of any use.

One obvious solution to this challenge is to make some form of cognition necessary for joint action. While this would cohere with the philosophical approaches to joint agency, it is not an ideal solution for an evolutionary approach for two reasons. First, human cognitive evolution is tightly bound up with human cooperation and sociality. If certain cognitive capacities are constitutive of or necessary for joint action then such an ability cannot serve the purpose of explaining the *origins* of complex cooperation and sociality. Joint action would have had to have evolved earlier in response to different selection pressures. Second, and relatedly, making cognition constitutive of joint action compromises the generality and power of evolutionary game theory for explaining the origins and stability of cooperative behavior. A key feature of almost all hypotheses concerning human evolution involves identifying the strategic contours of cooperation to set in motion a feedback loop or ratchet mechanism capable of producing smart, social hominins. The strategic features of behavior should have *something* to do with the evolutionary pressures on cognition and joint agency, so it would be costly theoretical move to set evolutionary game theory aside as orthogonal. Having rich cognitive capacities—to form the intention to hunt stag and to communicate this to potential partners—can make it easy to achieve cooperation in the Stag Hunt. Yet we also must confront the fact that such capacities are by no means necessary. Evolutionary dynamics produce (possibly 'minimal') cooperation in the Stag Hunt with regularity when, for example, simple signaling or correlated interaction is introduced (Skryms 2004). Bacteria can solve the comparatively more difficult public goods scenarios

(Inglis et al. 2012), so solving a Stag Hunt is well within microbial reach. In fact, one of the attractions of *evolutionary* game theory is that it does not presume the rich rationality and common knowledge constraints common to economic game theory (Skyrms 1996). Evolutionary dynamics find the stable cooperative equilibrium through either biological (sex and death) or cultural (learning and plasticity) means. Attributing joint agency to these systems is feasible, if we adopt a thin, deflationary account of the notion but this gives up on the task of connecting joint agency to some suite of cognitive capabilities. [4]

One strategy for avoiding such a thin characterization of joint agency yet adhering to a broadly comparative approach involves identifying criteria that determine a spectrum from "aggregative" group action to more "interactive" group action (Conradt and List 2009). The movements of bird flocks and fish schools are classic examples of aggregative group action, whereas symbiotic relationships across species or coordinating reproductive roles in cooperative breeding count as interactive. For the wolf and dolphin examples above, the opportunistic recruitment interpretation would classify hunting closer to aggregative, whereas the careful coordination between individuals and communication, as actually observed in the wolf-bison or the dolphin prey herding behaviors, provide evidence that group action here is much more interactive. On this approach, true joint agency would be implicated in interactive group action. This is a promising way to connect more extrinsic contours of the strategic interaction with cognitive capacities that mediate behavior.

---

[4] Grueneisen et al. (2017) present results that suggest that on a more minimal conception, such as the one articulated by Butterfill and Sinigaglia (2023), chimpanzees and bonobos count as engaging in joint action.

Such an approach does not (yet) commit to specific cognitive capacities for joint agency but requires identifying the criteria for interactive group action. Some characterization of cooperation would certainly be a part of this account, but additional criteria could help exclude many of the problematic Nash equilibrium cases. Drawing on the examples, careful or specific attention to the number and position of pack or pod mates, fluid and adaptive responses by the group to the situation with respect to prey type or other conditions, and communication between individuals are all good candidate criteria for distinguishing interactive from aggregative group action. Another set of candidate criteria involve causal interaction, spatial integration and other features of *biological individuality*. A community of microbes engaged in a public goods game may each be pursuing their individual evolutionary interests and only cooperating or engaging in joint action in an aggregative way. Interactive joint action would involve, say, the tight, specialized coupling of metabolisms between host and symbiont. This combines the approach to biological joint agency with the evolutionary features that involve the emergence of new biological individuals. Okasha (2018) defends this connection most explicitly, arguing that agents are biological individuals that exhibit sufficient "unity of purpose" understood in terms of the selected functional integration of their components. This effectively mirrors a move seen in the philosophical literature that links group action to the metaphysics of agency (Rovane 2019).

This contrast between thin and thick accounts of joint agency presents a dilemma: *either* a (thinner) form of joint agency plays a broad explanatory role in the evolution of animal social lifeways, *or* (thicker) joint agency is the product of a rich social cognition that is a rare, perhaps unique, evolutionary accomplishment. Based on the considerations raised thus far, opting for an account that can play a broad explanatory role is the best path forward. Forms of joint agency and cooperation may partake of diverse traits, cognitive capacities, and environmental resources.

And if (as we have suggested) joint agency and cooperation exist on a spectrum, determining the cognitive capacities and environmental situations that support more and less sophisticated joint agency and cooperation is a difficult task, and the role of joint agency in evolution becomes a moving target. No single peculiar cognitive capacity would be necessary for joint agency. Instead, a range of capacities implicated in attention, communication, recognition, and sociality will matter to differing degrees in different cases.

5. Joint agency and human cognition

A common inferential pattern found in the study of animal cognition involves identifying—through a combination of designed experiments and careful observations—behavioral patterns then postulating the minimal cognitive capacities necessary to produce those patterns. To take just one example, researchers designed experiments for the tool-using New Caledonian crows in an effort to assess their cognitive capacities for learning and solving novel problems (Taylor et al. 2010). Based on the crows' performance, researchers can make inferences about what sorts of learning mechanisms may be implicated in crow cognition. Much like tool use, cooperative behavioral patterns, even complex interactive kinds of cooperation, are not unique to the human lineage. Therefore, unless we opt for the thinnest of accounts, the existence of cooperative behavioral patterns alone is not sufficient for there to be selection pressures in place to support joint agency. So, what else is required for joint agency to emerge or be assembled? What sort of selection pressures are required?

At this point, some skepticism that joint agency is the product of a single, unified cognitive capacity is warranted. Consider what cognitive competences joint agency minimally requires in the context of a group hunt. An individual band member needs to recognize and keep

track of other band mates and their positions, to be able to plan or at least anticipate the actions of their fellow hunters, and to communicate those plans, anticipations, and positions in real time. Once the action starts, individuals may need to track gaze, to respond rapidly to changes in circumstance, to remember the plan, and to react accordingly. This is, by no means, a simple task, and it implicates a range of cognitive capacities that are subject to different sorts of selection pressures and are deployed in a wide range of tasks beyond cooperation and social interaction. What's more, as Butterfill and Pacherie (2020) demonstrate, similar forms of cooperation may be stabilized by importantly different cognitive capacities and strategies.

Rather than taking some class of capacities as a canonical package that is evolutionarily novel in humans, we hold that we owe an evolutionary account of the large number of cognitive capacities that make up the purported package. Furthermore, taken individually, components of the joint agency package are not unique to humans. Gaze tracking, for example, is well documented in several mammals and is prevalent in many primates (see, e.g., Tomasello et al. [2007]; and the overview in Zuberbühler [2008]; for corvids, see Bugnyar et al. [2016]). In Tomasello's work on this topic, he and his collaborators pick out differences between the gaze following of humans and the great apes, pointing to the need for an evolutionary story that accounts for such differences (Tomasello et al. 2007). This would not provide an evolutionary account of the emergence of joint agency but rather, an evolutionary account of one part of the package of cognitive capacities that make up joint agency.

There are several alternate accounts of the evolution of the cognitive capacities underlying cooperative behavior in humans. Most are laid out in the context of accounts of the evolution of human uniqueness. The first step of this approach is to pick out a human trait of interest, say elaborate cooperative behavior, and point to examples illustrating that the version of

22

this behavior in humans is distinct. Further, it is then argued that this distinct behavior, or behavioral suite, is key to our uniqueness. Finally, evolutionary accounts are offered for the relevant trait.

We are not challenging human evolutionary uniqueness here—we acknowledge that we deploy language, culture and norms in ways that are not replicated anywhere in evolutionary history, as any cursory examination of the tree of life will reveal. However, we do challenge the claim that joint agential cooperation is the hallmark of human uniqueness.[5] Evaluating a uniqueness hypothesis requires a comparative approach. Some point, for example, to chimpanzees and their relative inability to cooperate to the degree of sophistication of human children. For instance, human children coordinate better in riskier strategic scenarios than chimpanzees due to their ability to communicate more effectively (Duguid et al. 2014). Yet, as section 2 demonstrates, evolution has solved group hunting problems time and time again, sometimes using communication to help. This fits a general pattern that we see across the tree of life, where evolution has solved all sorts of cooperation problems, often in elaborate and important ways, including building complex multicellular individuals out of simpler unicellular organisms (Buss 1988; Maynard Smith and Szathmáry 1995; Griesemer and Shavit 2023).

---

[5] In Calcagno and Fuentes (2012) evolutionary anthropologists present 11 different key factors that led to human uniqueness. There are many more evolutionary accounts that point to one key factor resulting in human uniqueness. Sterelny (2012) argues that none of these "magic bullet" approaches will do the job, rather he says that we should consider all the many and various evolutionary pathways that led to modern hominids.

Our perspective differs from some leading accounts of the evolution of joint agency, for which markers of human uniqueness feature. Tomasello (2016; 2019), for example, maintains that rich cognitive capacities are required for cooperative behavior and so accounting for the evolution of such behavior requires an account of the evolution of the relevant rich cognitive capacities. The capacities he has in mind are those emphasized by leading philosophical accounts of bona fide joint agency – capacities for shared or collective intentionality (Schweikard and Schmid 2021).

Such capacities can be understood in different ways, of course. Bratman (1993) accounts for Jill and Sue's co-parenting in terms of their shared intention to co-parent. Their shared intention is two (or more) individual states of the form "I intend that we J" and the relations between these states. In contrast, Gilbert (1992; 2009) proposes that ascribing intentions to groups presupposes a "unification of agency." Joint commitments of the form "I will co-parent, if you will co-parent" are group commitments on this account rather than individual commitments. When joint commitments hold, intentions and beliefs are properly attributed to the group rather than individuals in the group.

Tomasello invokes both Bratman and Gilbert in introducing his own account of joint intention. He also understands the evolution of joint intention as a key step in our uniquely human evolutionary trajectory. Developing partnerships to engage in collaborative activities such as antelope hunting led to the formation of joint commitments and this in turn, "created an evolutionarily novel form of moral psychology" (Tomasello 2016, 5). Thus, "human morality is a form of cooperation, specifically, the form that has emerged as humans have adapted to new and species unique forms of social interaction and organization" (Tomasello 2016, 2).

For Tomasello, joint intention and other cognitive capacities required for the evolution of human hyper-sociality, are uniquely human traits.[6] Tomasello's evolutionary conjecture is that the capacity to engage in uniquely human joint action, and form commitments to undertake specific roles to bring about a joint end, emerged against the backdrop of early hominin obligate cooperation and facilitated a transition to more extreme human sociality and specialized cooperation. Some strategic cooperative interactions had already been undertaken in the human lineage, but the peculiarly human forms had yet to emerge (Tomasello et al. 2012; Tomasello 2016; 2019). On this view, uniquely human joint agency is supposed to be both a distinct achievement of the human lineage as well as a capacity (or set of capacities) that facilitates increasingly complex cooperative behavior culminating with morality: "The ability to form shared agencies derives from a uniquely human social psychology of shared intentionality" (Tomasello 2020b, 3).

---

[6] Any account of the evolution of these traits must explain how they arose in humans and why they have not appeared in other social species. Tomasello draws support from his work on animal cognition, specifically his work on great apes. He concludes that "[The] empirical facts establish beyond a reasonable doubt, in our opinion, that human beings are biologically adapted for collaboration in a way that other great apes are not" (Tomasello 2016, 77). So, if we were to see two chimpanzees carrying a log together, or making a ladder to escape their enclosure (Held 2019), we cannot explain this behavior in terms of joint intentions. Arguably, an implication of Tomasello's account is that such behavior should not be identified as joint action, or as true cooperative behavior.

This approach divides cooperative behavior into two categories: strategic cooperation of a sort that we see across the biological domain versus human cooperation facilitated by uniquely human joint agency. As Tomasello puts the thought, "The central claim of the current account is that the skills and motivation to construct with others an interdependent, plural-agent "we" […] are what propelled the human species from strategic cooperation to genuine morality" (Tomasello 2016, 3-4). More recently, Tomasello spells out his evolutionary hypothesis about joint agency in even starker terms: "the claim is that these abilities emerged first in human evolution between collaborative partners operating dyadically in acts of joint intentionality, and then later among individuals as members of a cultural group in acts of collective intentionality" (Tomasello 2019, 7). If joint agency is uniquely human, then it is likely a late arriving synthesis of cognitive capacities, thus playing relatively little role in the evolution of those very capacities implicated in joint agential action. This contrasts sharply with the view we defend here, which points to the possibility of numerous evolutionary precursors to the cognitive capacities central to Tomasello's account.

There are several alternate approaches to Tomasello's that also prominently feature the evolution of human uniqueness. For example Hrdy (2009) and Wrangham (2019) provide alternate accounts. We briefly look at each in turn. Kristen Hawkes identifies a key component to Hrdy's account of the evolution of human cognitive traits: her adoption of a life history approach. Hrdy's account of the evolution of the cognitive differences between human infants and apes places the evolution of the relevant traits in ancestral infants: "preferences for participation, engagement and sharing attention and intentions with others that emerge in infancy, evolved as survival responses to the cognitive ecology of ancestral infancy" (Hawkes 2020, 2). Just as the tadpole life stage of frogs faced different selection pressures than adult

frogs, ancestral human infants faced different selection pressures to ancestral human adults. Hrdy and Tomasello share an explanatory target: the difference between human infant cognitive activity (and human social cognition in general) and that of the great apes. Hrdy agrees with Tomasello that humans have unique social cognition (as do we), but Hrdy provides an alternate answer to the question of how our social cognition evolved. According to Hrdy, one part of the human social cognition package evolved in ancestral infants who "arrive[d] into a socioecology where mothers, depending on allomothers, bear another offspring before the previous one is independent" (Hawkes 2020, 6). A key difference between ancient human socioecology and ape socioecology is the presence of grandmothers (see Hawkes et al. 1998; Hawkes 2004). Ancient infants were forced to attract the attention of caregivers other than their birth mothers.

Another alternative approach, due to Wrangham (2019), treats the cooperative communication behaviors that are hallmarks of joint agency as initially emerging as a *byproduct* rather than a directly selected effect. Wrangham argues that the ability to tolerate direct gaze and spend time in close proximity to conspecifics—a necessary feature of close cooperation and joint action—evolved as part of a process of self-domestication, or selection against reactive aggression. Wrangham offers a comparative argument that selection for domestication (and against reactive aggression) produces a similar domestication syndrome in mammals, and the ability to follow gaze and tolerate individuals are paedomorphic (juvenile) traits accompanying this syndrome (2019, 172-189). He concludes that "[i]n domesticated animals, increased cooperative ability emerges as an incidental consequence of selection against reactive aggression" (Wrangham 2019, 189), though he is quick to make clear that this origin is

compatible with subsequent ratchet-like selection for increasing cooperation in the human lineage.[7]

This brief sample illustrates not only are alternate evolutionary approaches available to account for a wide range of cooperative behavior, but there are a wide range of alternate evolutionary approaches available to account for "uniquely human" cooperative behavior. In fact, the continuity of cooperative behaviors across the tree of life undermines the idea that there is a distinct category of uniquely human cooperation—human cooperative behavior simply differs in degree and sophistication from non-human cases, underscoring our earlier suggestion that identifying a normative account of bona fide joint agency or cooperation is not the most useful move for evolutionary explanation. Further, human cooperative behavior may well be supported by a cognitive package that some characterize as joint agency, but that package has most likely been assembled from disparate cognitive capacities. The evolution of each of the components of this package may call for different accounts and these accounts may create tradeoffs or conflict amongst the array of selection pressures.

The various evolutionary accounts from Hrdy, Tomasello, and Wrangham differ in many ways that reflect attention to different facets of human social life and the conditions that made such sociality possible (See also Calcagno and Fuentes 2012). Unsurprisingly, they reach different conclusions about what features are responsible for human evolution. The formal framework of evolutionary game theory provides a way to unify these approaches as various

---

[7] According to Sterelny (2012), Tomasello, Hawkes, Hrdy and Wrangham all present "magic bullet" accounts, whereas he prefers a co-evolutionary approach that factors in all of these considerations and more.

ways of solving general cooperation problems. This unification can provide some insight, but it does not help gain traction on the cognitive capacities implicated in the various solutions, for the generality of the framework extends across the tree of life. This compromises any explanation for human uniqueness that the relies on solving strategic problems associated with complex cooperation. Perhaps more important, there are drawbacks to attempting to give an evolutionary account of the origin of some new cognitive package, such as joint agency, that unlocks human sociality. Such accounts downplay the many components involved in such a package and their likely separate evolutionary histories. These evolutionary histories are almost certain to interact in ways that compromise, synergize and complicate the emergence of joint agency.

6. Conclusion

Collective action, joint action and cooperation all call for explanation. Game theorists can provide a general account of cooperation but their approach is often criticized for falling short because some cooperation appears to require sophisticated cognitive activity to occur. Such cases of cooperation or joint action seem more amenable to accounts that spell out the requisite cognitive capacities. Philosophical accounts of joint agency and joint intention can serve as a candidate explanations here. Evolutionary theorists are interested in these phenomena, too. If human cooperation is achieved via sophisticated cognitive capacities, how these capacities evolved is an important question. Some argue that uniquely human joint agency evolved in humans and it is an evolutionary precursor to our hyper-social lifeway. We have urged a different approach to the evolutionary question. There are many types of cooperative behavior and many non-human animals achieve various levels of cooperation. For many of these cases, evolutionary game theory, with its minimal assumptions about the cognitive capacities of agents,

can make headway in explaining how such behaviors arise and evolve. The evolution of

cooperative behaviors does not *require* a rich suite of cognitive capacities. Further, as the cases

of collaborative hunting in wolves in dolphins illustrate, there are cases of cooperation that do

involve impressive cognition and carry the hallmarks of joint agency yet without the particular

and extensive array of capacities we associate uniquely with humans. When we take a

comparative approach and canvas alternate answers to the question of how our hyper-social

human lifeway evolved, the proposal that the evolution of a sophisticated cognitive capacity,

such as joint agency, provides a 'magic bullet' explanation becomes less and less persuasive.[8]

First, it seems unlikely that a cognitive capacity to mediate cooperation evolved as one

mechanism, because the relevant behavior is supported by many mechanisms acting in concert.

Social cognition involves tracking cooperators, gaze following, fluid response patterns, and

memory, among other cognitive capacities. These cognitive activities are exhibited in other

animals. As a result, when we tackle human uniqueness, the evolutionary question becomes:

what evolutionary forces led to the different versions of these capacities present in great apes and

humans (and beyond)? Second, we should encourage a broader range of evolutionary answers to

this question. Once we acknowledge that cognitive capacities can be understood in terms of

many sub-capacities also present in other animals, we can then look to aspects of ancestral

human life that provide selection pressure to change these capacities. Such selection pressures

need to be specific to humans and may also impinge upon specific human life stages, for

---

[8] This point converges with an argument from Baron et al. (2023), where they urge us to look for

"major transitions" in the evolution of animal nervous systems rather than focusing on providing

evolutionary accounts for specific cognitive capacities.

example, selection for shared attention and gaze tracking in ancestral human infants. Also, some relevant human capacities may have arisen as evolutionary byproducts rather than as direct products of selection. As we see it, cooperation is diverse, widespread, and evolved in many different ways. Some forms of sophisticated human cooperative behavior are best accounted for by appealing to a suite of cognitive capacities. The relevant evolutionary questions are how each of these capacities were shaped in human evolutionary history and whether (or to what extent) we share a version of these capacities with great apes or other mammals. We reject the idea that one complex cognitive mechanism, such as joint intention or agency, evolved de novo and gave rise to uniquely human cooperative behavior.

References

Angus, Simon D. and Jonathan Newton. 2015. "Emergence of Shared Intentionality Is Coupled to the Advance of Cumulative Culture." *PLoS Computational Biology* 11: e1004587.

Baron, Andrew B., Marta Halina and Colin Klein, 2023. "Transitions in Cognitive Evolution." *Proceedings of the Royal Society B* 290: 20230671

Benoit-Bird, Kelly J. and Whitlow W. L. Au. 2009. "Cooperative Prey Herding by the Pelagic Dolphin *Stenella longirostris*." *The Journal of the Acoustical Society of America* 125:125-137.

Benoit-Bird, Kelly J. and Whitlow W. L. Au. 2009. "Phonation Behavior of Cooperatively Foraging Spinner Dolphins." *The Journal of the Acoustical Society of* America 125:539-546.

Bratman, Michael E. 1992. "Shared Cooperative Activity." *The Philosophical Review* 101: 327-341.

Bratman, Michael E. 1993. "Shared Intention." *Ethics* 104: 97–113.

Bugnyar, T., S. A. Reber, and C. Buckner. 2016. "Ravens Attribute Visual Access to Unseen Competitors." *Nature Communications* 7: 10506.

Buss, Leo. 1988. *The Evolution of Individuality.* Princeton: Princeton University Press.

Butterfill, Stephan and Elisabeth Pacherie. 2020. "Towards a Blueprint for a Social Animal." In *Minimal Cooperation and Shared Agency: Studies in the Philosophy of Sociality,* vol 11, ed. Anika Fiebich. New York: Springer.

Butterfill, Stephen and Corrado Sinigaglia. 2023. "Towards a Mechanistically Neutral Account of Acting Jointly: The Notion of a Collective Goal." *Mind* 132: 1-29.

Calgano, James M. and Agustín Fuentes. 2012. "What Makes Us Human? Answers from Evolutionary Anthropology." *Evolutionary Anthropology* 21: 182-194.

Conradt, Larissa and Christian List. 2009. "Group Decisions in Humans and Animals: A Survey." *Philosophical Transactions of the Royal Society B* 364: 719-742.

Duguid, Shona, Emily Wyman, Anke F. Bullinger, Katharina Herfurth-Majstorovic and Michael Tomasello. 2014. "Coordination Strategies of Chimpanzees and Human Children in a Stag Hunt Game." *The Proceedings of the Royal Society* 281:20141973.

Epstein, Brian. 2015. *The Ant Trap: Rebuilding the foundations of the social sciences*. Oxford: Oxford University Press.

Gilbert, Margaret. 1992. *On Social Facts*. Princeton: Princeton University Press.

Gilbert, Margaret. 2009. "Shared Intention and Personal Intentions." *Philosophical Studies* 144:167–87.

Gold, Natalie and Robert Sugden. 2007. "Collective Intentions and Team Agency." *The Journal of Philosophy* CIV: 109-137.

Griesemer, Jim and Ayelet Shavit. 2023. "Scaffolding Individuality: Coordination, Cooperation, Collaboration and Community." *Philosophical Transactions of the Royal Society B* (378): 20210398.

Grueneisen, Sebastian, Shona Duguid, Heiko Saur, and Michael Tomasello. 2017. "Children, Chimpanzees, and Bonobos Adjust the Visibility of Their Actions for Cooperators and Competitors." *Scientific Reports* 7: 8504.

Hamilton, Rebecca A., Stefanie K. Gazda, Stephanie L. King, Josefin Starkhammer, and Richard C. Connor. 2022. "Bottlenose Dolphin Communication During a Role-Specialized Group Foraging Task." *Behavioural Processes* 200:104691.

Hawkes, Kristen, James O'Connell, Nicholas Blurton-Jones, H. Alvarez, Eric L. Charnov. 1998. "Grandmothering, Menopause, and the Evolution of Human Life Histories." *Proceedings of the National Academy of Sciences* 95: 1336-1339.

Hawkes, Kristen. 2004. "The Grandmother Effect." *Nature* 428: 128-129.

Hawkes, Kristen. 2020. "Cognitive Consequences of our Grandmothering Life History: Cultural Learning Begins in Infancy." *Philosophical Transactions of the Royal Society B: Biological Sciences* 375: 20190501

Held, Amy. 2019. "Chimpanzees Build Makeshift Ladder, Break Free from Zoo Enclosure." NPR.org. https://www.npr.org/2019/02/11/693560383/chimpanzees-build-ladder-to-break-free-from-belfast-zoo-enclosure

Hrdy, Sarah B. 2009. *Mothers and Others: The Evolutionary Origins of Mutual Understanding.* Cambridge: Belknap Press.

Inglis, R. Fredrik, Sam P. Brown, and Angus Buckling. 2012. "Spite versus Cheats: Competition Among Social Strategies Shapes Virulence in *Pseudomonas aeruginosa*." *Evolution* 66:3472-3484.

Joyce, Richard. 2006. *The Evolution of Morality*. Cambridge: MIT Press.

Kaufmann, Angelica. 2015. "Animal Mental Action: Planning Among Chimpanzees." *Review of Philosophy and Psychology* 6:745-760.

Kaufmann, Angelica. 2016. "Collective Intentionality: A Human—Not a Monkey—Business." *Phenomenology and Mind* 2:98-105.

Ludwig, Kirk. 2020. "What Is Minimally Cooperative Behavior?" In *Minimal Cooperation and Shared Agency: Studies in the Philosophy of Sociality*, vol 11, ed. Anika Fiebich. New York: Springer.

MacNulty, Daniel R., Douglas W. Smith, L. David Mech, John A. Vucetich, and Craig Packer. 2012. "Nonlinear Effects of Group Size on the Success of Wolves Hunting Elk." *Behavioral Ecology* 23:75-82.

MacNulty, Daniel R., Aimee Tallian, Daniel R. Stahler, and Douglas W. Smith (2014). "Influence of Group Size on the Success of Wolves Hunting Bison." *PLoS One* 9:e112884.

Maynard Smith, John and Eors Szathmáry. 1995. *The Major Transitions in Evolution.* Oxford: W. H. Freeman Spektrum.

McMahon, Christopher. 2005. "Shared Agency and Rational Cooperation." *Noûs* 39: 284-308.

Moore, Richard. 2017. "Social Cognition, Stag Hunts, and the Evolution of Language." *Biology and Philosophy* 32:797-818.

Moore, Richard. 2018. "Gricean Communication, Joint Action, and the Evolution of Cooperation." *Topoi* 37(2):329-341.

Okasha, Samir. 2018. *Agents and Goals in Evolution*. Oxford: Oxford University Press.

Paternotte, Cédric. 2014. "Minimal Cooperation." *Philosophy of the Social Sciences* 44(1):45-73.

Pettit, Philip and Christian List. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.

Pettit, Philip and David P. Schweikard. 2006. "Joint Actions and Group Agents." *Philosophy of Social Sciences* 36: 18-39.

Planer, Ron and Kim Sterelny. 2021. *From Signal to Symbol: The Evolution of Language.* Cambridge: MIT Press.

Ritchie, Katherine. 2020. "Minimal Cooperation and Group Roles." In *Minimal Cooperation and Shared Agency: Studies in the Philosophy of Sociality*, vol 11, ed. Anika Fiebich. New York: Springer.

Rovane, Carol. 2019. "Is Group Agency a Social Phenomenon?" *Synthese* 196:4869-4898.

Sayigh, Laela S., Nicole El Haddad, Peter L. Tyack, Vincent M. Janik, Randall S. Wells, and Frants H. Jensen. 2023. "Bottlenose Dolphin Mothers Modify Signature Whistles in the Presence of Their Own Calves." *Proceedings of the National Academy of Sciences* 120(27), e2300262120.

Schweikard, David P. and Hans Bernhard Schmid. 2021. "Collective Intentionality." In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/fall2021/entries/collective-intentionality/>.

Skyrms, Brian. 2004. *The Stag Hunt and the Evolution of Social Structure*. Cambridge: Cambridge University Press.

Sterelny, Kim. 2012. *The Evolved Apprentice: How Evolution Made Humans Unique*. Cambridge: MIT Press.

Sterelny, Kim. 2021. *The Pleistocene Social Contract: Culture and Cooperation in Human Evolution.* Oxford: Oxford University Press.

Taylor, Alex H., Douglas Elliffe, Gavin R. Hunt, and Russell D. Gray. 2010. "Complex Cognition and Behavioural Innovation in New Caledonian Crows." *Philosophical Transactions of the Royal Society B: Biological Sciences* 277:2637-2643.

Tollefsen, Deborah. 2014. "Social Ontology." In *Philosophy of Social Science: A New Introduction* eds. Nancy Cartwright and Eleonora Montuschi, 85–101. Oxford: Oxford University Press.

Tomasello, Michael. 2014. *A Natural History of Human Thinking*. Cambridge: Harvard University Press.

Tomasello, Michael. 2016. *A Natural History of Human Morality*. Cambridge: Harvard University Press.

Tomasello, Michael. 2019. *Becoming Human: A Theory of Ontogeny*. Cambridge: Harvard University Press.

Tomasello, Michael. 2020a. "The Adaptive Origins of Uniquely Human Sociality." *Philosophical Transactions of the Royal Society B* 375(1803): 20190493.

Tomasello, Michael. 2020b. "The Moral Psychology of Obligation." *Behavioral and Brain Sciences* e56:1-58.

Tomasello, Michael, Brian Hare, Hagen Lehmann, and Josep Call. 2007. "Reliance on Head versus Eyes in the Gaze Following of Great Apes and Human Infants: The Cooperative Eye Hypothesis." *Journal of Human Evolution* 52:314–320.

Tomasello, Michael, Alicia P. Melis, Claudio Tennie, Emily Wyman, and Esther Herrmann. 2012. "Two Key Steps in the Evolution of Human Cooperation: The Interdependence Hypothesis." *Current Anthropology* 53: 673– 92.

Tuomela, Raimo. 1992. "On the Structural Aspects of Collective Action and Free-Riding." *Theory and Decision* 32:165-202.

Tuomela, Raimo and Kaarlo Miller 1988. "We-Intentions." *Philosophical Studies* 53: 367-389.

Westra, Evan and Kristen Andrews. 2022. "A Pluralistic Framework for the Psychology of Norms." *Biology and Philosophy* 37: 1-30.

Whiten, Andrew, Francisco J. Ayala, Marcus W. Feldman, and Kevin N. Laland. 2017. "The
Extension of Biology Through Culture." *Proceedings of the National Academy of
Sciences* 114(30):7775-7781.

Wrangham, Richard. 2019. *The Goodness Paradox: The Strange Relationship Between Virtue
and Violence in Human Evolution*. New York: Pantheon Books.

Zuberbühler, Klaus. 2008. "Gaze Following." *Current Biology* 18: R453–R455.