Chapter 7
Towards a Blueprint for a Social Animal

Stephen Butterfill and Elisabeth Pacherie

Abstract In this chapter, we attempt to answer the question, By what steps could members of a group capable of acting together with a purpose, coordinating flexibly, communicating cooperatively and deceiving competitors be constructed from creatures with minimal social skills and cognitive abilities? The method we use is creature construction: the idea is to adopt the perspective of a designer tasked with specifying a sequence of creatures, where each is independently viable and has the capacities of its predecessors together with some new capacity which enables it to overcome limits its predecessors faced. In creature construction, the aim is not to characterise actual species, nor to describe actual evolutionary or developmental processes. Instead the aims are to understand how various forms (or prototypes) of joint action are related to, and diverge from, each other; and to identify limits on what can be achieved with a given set of cognitive and social skills.

We start with Alphonso and his kin, whose social cognition is limited to tracking the goals of others' actions. We show that despite little cognitive sophistication, the salience and triangulation heuristics enables them to initiate simple joint actions requiring coordination. One group of their descendants, Beki and her kin, develop abilities to produce pointing gestures and object-directed vocalisations, that enable them to enlist others not yet as partners but as social tools, thus extending the range of situations in which they can rely on the salience and triangulation heuristics. Another group of Alphonso's descendants, Bemi's kin, learn the art of strategic deception, acquiring increasingly elaborate tactics for manipulating others' action possibilities. This advantages them in competition. Finally, the Kimi, who are mixed descendants of both the Beki and the Bemi, inherit the former's communicative abilities and the latter's abilities for tactical deception. Progressively integrating the two allows them to develop new capacities of selective deception.

We argue that although our creatures do not yet have all the cognitive capacities classical accounts imply are needed for joint action, they have proxies for some of these capacities. These proxies allow them to coordinate in a limited but useful

S. Butterfill (✉)
Department of Philosophy, Social Sciences Building, University of Warwick, Coventry, UK
e-mail: s.butterfill@warwick.ac.uk

E. Pacherie
Institut Jean Nicod, ENS, EHESS, CNRS, Département d'Etudes Cognitives, Ecole Normale Supérieure, Paris, France

range of ordinary circumstances. Further, relying on such proxies provide ways of avoiding both omni-doxasticity and omni-intentionality when acting together.

7.1 Introduction

Human social and societal life is built on thoughts, intentions, motivations and feelings that bind us and our actions together. One essential aspect of this is joint action. Despite its foundational role in all social and cultural life, joint action has only recently become a topic of inquiry in philosophy and the cognitive sciences (for example, see Bratman 2014; Gilbert 2013; Tomasello 2009; Rakoczy 2017). A number of classical accounts of joint action make substantive demands on cognitive abilities (as has been argued by Butterfill 2012; Pacherie 2011; Pacherie 2013; Pacherie and Dokic 2006; Tollefsen 2005; Obhi and Cross 2016). By contrast, we pursue a minimalist agenda with the aim of isolating some undemanding forms of joint action (or prototypes for joint action) and their more modest cognitive requirements.

Our method is creature construction, which has a distinguished history in philosophy (Bratman 2000, 2014; Grice 1974; Strawson 1959). The idea is to adopt the perspective of a designer tasked with specifying a sequence of creatures, where each is independently viable and has the capacities of its predecessors together with some new capacity which enables it to overcome limits its predecessors faced. As Grice described it,

> The method [...] is to construct (in imagination, of course) [...] a sequence of types of creature, to serve as [...] models for actual creatures. [...] The general idea is to develop sequentially the psychological theory for different brands [...] (which of course is unlikely ever to be more than partial). (Grice 1974, 37)

Whereas Grice's project was 'to compare what one thus generates with the psychological concepts we apply to suitably related actual creature', our aims here are more modest. We will not propose that any of our constructed creatures resembles actual individuals very closely. Our claim is merely that the creatures we construct exhibit possible cognitive profiles. These fictional creatures are, however, loosely inspired by discoveries about developmental, comparative and cognitive psychology, as our footnotes reveal.

We will use the method of creature construction to attempt to answer the question, By what steps could members of a group capable of acting together with a

purpose, coordinating flexibly, communicating cooperatively and deceiving competitors be constructed from creatures with minimal social skills and cognitive abilities? The aims are to understand how various forms (or prototypes) of joint action are related to, and diverge from, each other; and to identify limits on what can be achieved with a given set of cognitive and social skills.

We take for granted in what follows that the creatures we are constructing have some ability to track cause and effect. This is, after all, a blueprint for a social animal.

7.2   Alphonso's Kin

Consider first Alphonso and his kin, whose social cognition is limited to tracking the goals of others' actions. This enables them to distinguish between, say, grasping and pushing actions. Their goal tracking is pure in the sense that it does not depend on representing intentions or other mental states (see Gergely et al. 1995; Csibra et al. 2003 for one account of how this is possible). What joint actions are they capable of?

Suppose Alphonso is walking in the mountains when he comes across one of his kin pushing against a small boulder that is blocking their paths. There is no way around the boulder: the only option is to move it. But it is manifestly too heavy for any individual to move. Alphonso joins in pushing the boulder. At first they are both pushing as hard as they can but it doesn't move much. Gradually they fall into a rhythm of pushing and releasing simultaneously, rocking it harder and harder until eventually it gives way altogether.

Moving the boulder counts as a joint action in at least a minimal sense. Minimally, a joint action is an event grounded[1] by two or more agents' actions (compare Ludwig 2007). This is a very broad notion of joint action. Whenever two or more agents' actions have a common effect and there is an event comprising the actions and their common effect, the actions will ground this event. Because Alphonso's and his kin's rocking actions have a common effect, namely moving the boulder, this is sufficient for their actions to ground a joint action.[2]

That Alphonso and his kin perform actions directed to a single goal is not entirely accidental because the environment has provided a single obstacle which they must overcome. But their both simply pushing is not enough: to move the boulder, coordinated action is required. How could their actions be coordinated? For Alphonso and his kin, coordination is nonintentional. In this case, the actions of Alphonso and

---

[1] Events D1, ... Dn *ground* E just if: D1, ... Dn and E occur; D1, ... Dn are each part of E; and every event that is a part of E but does not overlap D1, ... Dn is caused by some or all of D1, ... Dn. (This is a generalisation of the notion specified by Pietroski 1998.)

[2] Joint action of this kind is found in nonhuman primates (e.g. Suchak et al. 2016; Visco-Comandini et al. 2015) and human infants (e.g. Rakoczy 2017).

his kin could be coordinated thanks to a combination of two things: the joint affordance a large, barely movable rock presents them[3]; and entrainment, the process of synchronizing two or more rhythmic behaviours with respect to phase.[4]

Because their actions are coordinated in this way, there is a sense in which the joint action performed by Alphonso and his kin is purposive. Their actions are directed to moving the boulder. Importantly, because they are coordinated in such a way as to make the boulder's moving more likely, this is not just a matter of each agent's action being directed to moving the boulder. Rather their actions are collectively directed to this goal: it is thereby a joint goal *in the following minimal sense*.

A *goal* is an outcome towards which an action is directed (so not a mental state). And the actions which comprise a joint action can be *collectively directed to an outcome* in this sense: there is an outcome to which these actions are directed and this is not, or not only, a matter of each action being individually directed to that outcome.[5] A *joint goal* of a joint action is an outcome to which the actions comprising it are collectively directed. This sense of joint goal is broad enough to allow that the jointness may not be at the level of intentions or representations of the subject; it can instead be found at the level of the coordination mechanisms. So for Alphonso and his kin's actions to have a joint goal does not yet imply that they are aware of this, nor that the joint goal is an any sense a purpose they share in acting.

The boulder is a relatively simple case because the environment provides a single most salient goal to pursue. There is also just one most salient means to pursue this goal, and the demands on coordination are limited to the need for actions to be synchronized. But what happens when there is more than one goal to pursue, none more salient than the other? In such situations there will often be uncertainty about which goal any individual will pursue. On different trees there are two large bunches of jackfruits to harvest and efficient harvesting requires several people to participate, so that it would be futile for Alphonso and his kin to pursue different goals. How could Alphonso and his kin overcome such uncertainty, even without being aware that there is uncertainty about goals? If any of them makes a move to harvest jackfruits from one of the trees, she will be aware that some additional contribution to this action is needed (although she need not appreciate what the contribution is; compare Vesper et al. 2010). How is the additional contribution to be secured?

---

[3]A *joint affordance* is an affordance for the agents of a joint action collectively. That is, it is an affordance for these agents and this is not, or not only, a matter of its being an affordance for any of the individual agents. For evidence that joint affordances exist, (see M. Richardson et al. 2007; Davis et al. 2010; Doerrfeld et al. 2012).

[4]Entrainment is found in many species' behaviours (e.g. Backwell et al. 1998). In humans, entrainment can occur automatically—that is, independently of the subject's tasks and motivations (e.g. Varlet et al. 2015)—and without awareness (e.g. Richardson et al. 2005). For a review of emergent forms of coordination in joint action, (see Knoblich et al. 2011).

[5]Given how we have defined joint action (in terms of two or more agent's actions grounding a common effect), it follows that not all joint actions have joint goals. This is because for an outcome to be a joint goal, it is not sufficient that it be a common effect of two agent's actions. Rather, the actions must be collectively directed to this outcome in the above sense.

Suppose Alphonso goes ahead unilaterally and attempts to start harvesting jack-fruits as if on the assumption that the additional contribution will be forthcoming.[6] When he acts as if on this assumption, others can detect the goal of his actions. Providing they do so, his acting on this assumption makes the goal to which his actions are directed more salient or more attractive to others. This in turn makes it more likely that the assumption will turn out to be true, which can mean that acting on the assumption is reasonable.

Here, then, is a simple strategy that has the effect, not always but often enough, of enabling coordination when there is no single most salient goal to pursue (for example when several bunches of jackfruits could be harvested):

1. Pick a goal to pursue.
2. If it is not too costly to end up acting alone, go ahead and act as if on the assumption that any necessary additional contribution will be forthcoming.
3. If this turns out not to be the case, change your objectives.

A similar strategy is used by Alphonso and his kin when there are complementary roles and it is uncertain who is to do what. For example, to harvest the jack-fruits, one agent needs to climb the tree and break the stems while the other should stand below and catch the falling bunches. Alphonso proceeds by selecting and performing a role, as if on the assumption that the other will perform the complementary role. Where this turns out not to be the case, he will eventually change role or objectives.

However, there are situations in which it is too dangerous or otherwise too costly to act as if on the assumption that others' contributions will be forthcoming. To illustrate, suppose Alphonso and his kin have dispersed in the woods to forage for mushrooms when a large pig comes by. This is an opportunity for them to hunt the pig, but their actions will need to be tightly synchronized as tackling the animal alone would be too dangerous. If only they had common knowledge, they would be able to safely rely on the assumption that each will join pursuit of the pig. But Alfonso and his kin do not attribute, and are not aware of, knowledge states. So they do not have common knowledge. To avoid both missing an opportunity and risking disaster, they need a proxy[7] for common knowledge, one that requires no awareness of knowledge states.

---

[6] We write 'as if on the assumption' because Alphonso need not actually assume that the additional contribution will be forthcoming. Rather, his actions may be produced in such a way that they often rely for their success on others' contributions without Alphonso himself having any view on the matter. For comparison, an agent's actions can be produced in such a way that they rely on unsupported objects falling without the agent herself having any view on this. Note that this is possible even when there are significant exceptions (not all unsupported objects fall when underwater); compare Perry (1993, 202).

[7] To say that A is a *proxy* for B is to say that, at least within a useful if limited range of circumstances, A can fulfil some of the functions of which B would fulfil were B present in those circumstances. Where there are two proxies for B, one may be a *better proxy* than the other insofar as it can serve more of the functions associated with B, or insofar as it can serve functions associated with B in a broader range of circumstances.

One crude proxy for common knowledge involves stimuli such as loud noises or strong smells. Suppose the pig is making a loud and distinctive noise; or, more generally, that Alphonso and his kin are in a position to perform actions directed to goals which specify an object or event associated with a loud noise or strong smell. Suppose further that the noise or smell is salient enough that there could be little doubt that everyone nearby had picked it up, so that its existence would ensure common knowledge among people who were capable of having common knowledge.[8] Where such salient stimuli identify the pig, the risk to Alphonso of relying on the assumption that his kin's goal will be the goal of hunting the pig is reduced. So Alphonso and his kin can use salience as a proxy for common knowledge. This is the *salience heuristic*: where they would not ordinarily rely on the assumption that additional necessary contributions will be forthcoming, Alphonso and his kin will rely on this assumption where an object or event associated with the target of potential actions is both sufficiently salient and sufficiently more salient than any other object or event.[9]

Relying exclusively on salience in this way would stringently limit the range of situations in which Alphonso and his kin can perform joint actions. There will be many cases in which an event is not very salient but others appear likely to be affected by it. For example, they may exhibit characteristic responses to it, such as a startled response or a certain twitching of a nose; or they may have a line of sight to it.[10] Suppose Alphonso, observing the pig, can observe that he and the pig are each linked to his kin, where this link is a matter of the thing (himself or the pig) causally influencing his kin, or else it is a matter of his kin having a line of sight to the thing (to himself or to the pig). This triangular situation ensures that if Alphonso acts, there is less of a risk that additional contributions will not be forthcoming. This is the *triangulation heuristic*.[11]

Although neither is infallible, the salience and triangulation heuristics pick out situations in which, often enough, there could well be common knowledge among more sophisticated individuals. So Alphonso and his kin can use these heuristics as proxies for common knowledge.[12]

---

[8] At least some nonhumans are sensitive to what others hear (Santos et al. 2006), at least within limits (Bräuer et al. 2008).

[9] Note that it is salience itself, not a belief about, or representation of, salience, which drives the salience heuristic. This heuristic adapts Lewis' suggestion that common knowledge can be arrived at through public events (1969). He suggests an event E is the basis for common knowledge that P if (i) E is public and (ii) E indicates that P. For instance, a pig's grunting is a public event and indicates the presence of a pig and is therefore a basis for common knowledge.

[10] Line of sight calculations, in at least limited form (but perhaps along with more sophisticated forms of perspective taking), appear widespread in nonhumans. See, for example, (Kaminski et al. 2006; Bräuer et al. 2004; Bugnyar et al. 2004; Okamoto-Barth et al. 2007).

[11] The triangulation heuristic is to joint attention as the salience heuristic is to common knowledge (compare Rakoczy 2017).

[12] Our approach is similar in spirit to that of Carpenter (2009) who discusses 'common knowledge, in the sense of what is known or has been experienced together' (p. 383). Note, however, that she is committed to characterising common knowledge rather than merely a proxy for it.

These heuristics demonstrate that Alphonso and his kin can achieve a range of joint actions. And since the salience and triangulation heuristics work with larger groups as well as dyads, they can even achieve joint actions involving many individuals. But there are also limits on when they can act together, and on when they can avoid acting together. To illustrate, consider two situations. First, Alphonso and one of his kin hear a pig approaching and hide in order to ambush it. But the pig they heard turns out to be not one but two pigs. This defeats the salience heuristic, as neither pig is sufficiently more salient than the other. And they have no other way to determine which pig to attack. Second, sometimes members of Alphonso's kin come across a resource, such as a berry patch, which they would ideally exploit alone. If others are alerted, the original discoverer will get little. They have no way to prevent more dominant individuals from pilfering berries. While Alphonso and his kin are unequipped to overcome either of these problems, his descendants are more fortunate.

7.3 Beki's Kin

At this point in our construction, Alphonso's descendants divide into two groups, Beki's and Bemi's. Beki and her kin are frequently confronted with situations like the 'two pig' situation described above: situations in which there are multiple, equally salient possible goals achieving which would require joint action. In these situations they desire to act, and are both excited by a potential target of action and frustrated by their inabilities to act. But Beki and her kin develop abilities to produce pointing gestures and object-directed vocalizations. Initially, the gesture or vocalization is not a thoughtful attempt to communicate but merely an expression drawn out of them by a situation. Among the causes of their gestures and vocalizations is a combination of excitement at a potential target of action and frustration at not being able to act.

None of Beki's kin understand these gestures and vocalizations except perhaps as expressions of frustration. Despite this, on some occasions the gestures and vocalizations do function to draw others' attention to objects. They thereby have the effect, unintentionally, of extending the range of situations in which Beki and her kin can rely on the salience heuristic. For example, in the 'two pig' situation, Beki vocalizes at one of the pigs thereby making this pig more salient than the other to her kin. At this point they can rely on the salience heuristic (a proxy for common knowledge introduced above), and so capture the pig.

The gestures and vocalizations also extend the range of situations in which Beki and her kin can rely on the triangulation heuristic. To illustrate, consider the 'hidden pig' problem: there is a pig nearby which only Beki is linked to, and which is too dangerous to be tackled alone. Frustration at missing an opportunity to catch the pig triggers a vocalization directed to it. The vocalization draws Beki's kin's attention to the pig, thereby linking them to it as well. In this way, the conditions necessary for relying on the triangulation heuristic are met, and so Beki's and her kin capture the pig.

At this first stage, Beki's kin's communicative actions have limited effects because they are unintentional responses to exciting objects in frustrating situations (among other things). But over time they observe and become familiar with the causal effects of their gestures and vocalizations on their kin and on other animals around them. This allows them to produce gestures and vocalizations with the goal of bringing about their familiar effects. What was merely an expression of excitement and frustration (among other things) has become an instrumental action resembling a communicative act. This greatly expands the range of actions they can perform. Initially their goals were limited to acting on the physical environment. Now a new set of goals is open to them, namely influencing the behaviour of their kin. They can, for example, call to make another come towards them, or vocalize towards an object in order to direct another's actions to that object. This amounts to their coming to recognize each other as social tools (Warneken et al. 2012).

Over time, Beki's kin come to make and distinguish a range of gestures. Some of these center around danger. What was initially an involuntary vocal response to danger gradually becomes an action they produce with the goal of bringing about its familiar effects, so that it functions as a warning.[13]

Merely involuntary vocalizing and gesturing enabled Beki's kin to solve the 'two pig' problem sporadically, when their actions happened to make the pig more salient or linked it to their kin. But having voluntary control over their gestures and vocalizations enables them to gesture and vocalize strategically, suppressing gestures and vocalizations when these could hinder success, and using them to enlist others as social tools when they promote success. This gives them a systematic way of solving the 'two pig' problem. And, more generally, it means that when uncertainty and the costs of acting alone would otherwise prevent action, communicative actions can put them in a position to act.

7.4   Bemi's Kin

Bemi's kin occupy a region in which resources are scarce. This confronts them with challenges quite different from those facing Beki's kin, who inhabit a region with abundant resources. Competition for food becomes intense, and those who come across ripe berries must either be able to defend their find or consume it before others notice. This is especially challenging for weaker individuals: to stray too far from the group is dangerous, but to forage too close to others means retaining little of the food found.

The weaker of Bemi's kin acquire abilities for tactical deception (Byrne and Whiten 1985). When finding berries, the weaker individuals will occasionally

---

[13]Vervet monkeys have a range of danger calls for eagles, for pythons and, for leopards. Infant vervet monkeys are not very discriminating: they will produce the alarm call whenever they see a big bird in the sky; only later do they become able to distinguish eagles from nonthreatening big birds (Seyfarth et al. 1980).

refrain from exploiting the food source while others are around. Although such restraint will increase their chances of getting the berries for themselves, their actions need not be performed with any intention to gain an advantage. It may be, for example, that rising anxiety associated with anticipation of conflict over food immediately suppresses any inclination to feed. But at some point, the weaker among Bemi's kin learn that they can avoid having food stolen when foraging in proximity to others by delaying consumption of food for as long as possible while they are present. Now the acts (or omissions) of tactical deception are performed with an intention to avoid theft. But they are not yet performed with any insight into others' mental states.

Over time the value of freezing up or refraining from exploiting food is reduced as competitors come to associate these behaviours with the presence of food. This in turn leads to an escalation of tactical deception. The weaker among Bemi's kin begin to act as if the food was absent and walk on as others pass by.[14] And what at first was a tendency not to eat when others are around gradually becomes a tendency not to eat when others are linked to you or to the food. (As stipulated above, being linked is a matter either of having a line of sight to, or else of being causally influenced by, the food). They also become discriminating in when tactical deception is used, relying on it against stronger but not weaker competitors.

Eventually some of Bemi's kin realise that being linked to food is a precondition not only for stealing it but also for performing any action concerning it; and that what goes for food goes for any kind of object. So they appreciate that, for example, if someone is not linked to a sleeping snake she will not be able to avoid it. In this situation, one seizes the other's head and forcibly links her to the snake.

Bemi's kin take an important further step when their experience of manipulating whether competitors and partners are linked to things gradually clues them into the realisation that you need not be linked to something right now in order to act on it; in many cases, it is sufficient to have been linked to it at some point in the recent past. This enhances their abilities to prevent more dominant individuals from stealing their berries. By concealing a cache of berries from a dominant competitor regardless of whether they are currently hungry enough to steal, members of Bemi's kin can avoid theft when the competitor later becomes hungry again.

The story of Bemi's kin is one of gradually elaborating tactics for manipulating others' action possibilities. Merely involuntary freezing in the presence of competitors enabled Bemi's kin to protect some of their food discoveries from theft. Involuntary freezing was sometimes ineffective and occasionally even led to missing

---

[14] Compare De Waal (2016)'s description of what happened when experimenters hid grapefruits on the island where the colony of chimpanzees spend the day: 'After releasing the apes onto the island, a number of them passed over the site where we had hidden the fruits under the sand. Only a few small yellow patches were visible. Dandy, a young adult male, hardly slowed down when he ran over the place. Later in the afternoon, however, when all the apes were dozing off in the sun, he made a beeline for the spot. Without hesitation, he dug up the fruits and devoured them at his leisure, which he would never have been able to do had he stopped right when he saw them. He would have lost them to dominant group mates' (De Waal 2016 Chap. 2).

opportunities to eat, but it did provide Bemi's kin with the experiences necessary to acquire more refined tactical deception. Learning that freezing is associated with having food stolen less often, they began intentionally to delay gathering or consuming food when others were currently linked to them. This further enriched their experiences of how others' actions are associated with facts about what they are, or have recently been, linked to, allowing them to manipulate those links. This in turn created new opportunities to learn about when others' actions succeed and when they fail.

7.5   The Kimi

Climate change forces Beki's and Bemi's groups to migrate, and they end up in close proximity. In their new environment, food comes mostly from large animals. Tight action coordination involving multiple complementary roles is therefore needed for catching prey. But food is also scarce enough that a find needs to be protected from pilfering by others, which will often require strategic deception.

Success in hunting large animals requires tight coordination among a fairly large number of individuals, a division of roles, and the ability to anticipate another's complementary action in order to coordinate your own with it. Through repeating successful behaviours, the joint actions of Beki's and Bemi's kin come to follow conventional patterns in the way they unfold and in who does what. These conventional patterns resemble action scripts and function as precursors of planning. Here Beki's kin have an advantage: as they identify patterns in their past successful behaviours, they can use their communicative abilities to assign roles. But in communicating, Beki's kin often alert Bemi's kin, and so end up losing much of the food.

For their part, Bemi's kin rely on the triangulation and salience heuristics to coordinate in capturing an animal, and so rarely succeed unless they are together when they encounter an animal. The infrequency of their successes means that their joint actions involve less conventional patterns than those of Beki's kin, which further widens the gap between their hunting and that of Beki's kin. But when Bemi's kin do capture an animal, their tactical deception means that they rarely suffer from pilfering.

Both Beki's and Bemi's kin are thus struggling to survive, although for different reasons. As their societies disintegrate, members of the two groups occasionally reproduce and raise children who become skilled in both communication and tactical deception. These children belong to neither Beki's nor Bemi's group but are outcasts. As outcasts they are often thrust together, but without thereby forming a group in their own right. They are however sometimes forced to act together, as surviving alone is impossible. And in acting together they have an advantage over Beki's and Bemi's kin: they are simultaneously communicators, who enable coordination, and tactical deceivers, who can avoid pilfering. Repeated successes in acting together results in them forming stable groups. These are the Kimi.

Despite their advantages, the Kimi are initially vulnerable because their abilities to communicate are not fully integrated with their abilities for tactical deception. When a potential target appears, the Kimi will gesture to link other group members to the animal, so enabling cooperative action. In doing this they attract competitors, especially Bemi's kin, who, being much less successful at obtaining food themselves, rapidly learn to follow the Kimi and steal from them.

Eventually some of the Kimi realise that their gestures and vocalizations are drawing competitors to them at just the wrong moment. Just as their ancestors began to intentionally delay foraging, so they come to suppress gestures and vocalizations when competitors are around. Although this initially makes things better for them, Bemi's kin become so dependent on them that the new strategy only means the Kimi are rarely apart from their competitors. It is only when there is some pressing danger that Bemi's kin will not pursue the Kimi.

Through following the Kimi so closely, Bemi's kin gradually come to associate the Kimi danger calls with danger, running away whenever these calls are made. On detecting the association between danger calls and their competitors' flight, Kimi groups have the opportunity to put danger calls to a new use. Some now begin to use danger calls to scare Bemi's kin away. Initially these fake danger calls cause both Bemi's kin and the Kimi to respond as if there was danger. At this stage, they are only useful where a Kimi has found a food source and does not need to cooperate with her kin. But as more and more Kimi come to use or encounter fake danger calls, perhaps observing apparently anomalous combinations of danger calls followed by feeding behaviour, they gradually come to produce and respond to the danger calls in a more nuanced way. Meanwhile Bemi's kin, who are not so close, lack opportunities to observe the anomalous combinations and so fail to learn to differentiate genuine from tactically deceptive danger calls. As the new practice of tactical deception takes hold among the Kimi, they come to accompany danger calls with nonvocal communication about food, thereby ensuring cooperation from their group co-members nearby even while giving a danger call. The Kimi have now combined Bemi's tactical deception with Beki's communicative abilities. They can selectively deceive.

7.6 Conclusion

There in fact are no kin of Alphonso, Beki, Bemi or Kimi among us, we assume. But fiction is a philosophical tool that can help us to understand by what steps members of a group capable of acting together with a purpose, coordinating flexibly, communicating cooperatively and deceiving competitors could be constructed from creatures with minimal social skills and cognitive abilities. We started with Alphonso and his kin, whose social cognition is limited to tracking the goals of others' actions. Despite little cognitive sophistication, the salience and triangulation heuristics enables them to initiate simple joint actions requiring coordination such as gathering and hunting. But they are dependent on the environment to provide favourable

circumstances for coordination, and vulnerable to pilfering by dominant individuals. One group of their descendants, Beki's kin, are equipped to coordinate in less serendipitous circumstances. Beki and her kin develop abilities to produce pointing gestures and object-directed vocalizations. They can suppress gestures and vocalizations when these could hinder success, and use them to enlist others not yet as partners but as social tools, thus extending the range of situations in which they can rely on the salience and triangulation heuristics. Meanwhile, another group of Alphonso's descendants, Bemi's kin, learn the art of strategic deception, acquiring increasingly elaborate tactics for manipulating others' action possibilities. This advantages them in competition. Finally, the Kimi, who are mixed descendants of both the Beki and the Bemi, inherit the former's communicative abilities and the latter's abilities for tactical deception. Progressively integrating the two allows them to develop a new capacity for selective deception.

Although the creatures we have been constructing do not yet have all the cognitive capacities classical accounts imply are needed for joint action, they have proxies for some of these capacities. Alphonso's kin already have proxies for common knowledge and joint attention. These proxies allow them to coordinate in a limited but useful range of ordinary circumstances. Whether or not the things they do together are strictly speaking cases of joint action, they will at least appear very much like joint actions.

Further, having these proxies continues to be useful for the descendants of our creatures. It frees them from relying on explicit beliefs and sophisticated forms of reasoning in many ordinary situations. As Perry says, even though the descendants may have greater cognitive sophistication, their designers will want to avoid 'omnidoxasticity'. Instead, '[a] more efficient way for Mother Nature to proceed is to fit our psychology to the constant factors in our environment, and give us a capacity of belief for dealing with the rest' (Perry 1993, 202). Of course Perry's focus is an individual acting alone: for much of the time, at least, it would be unfortunate to have to rely on beliefs about gravitational forces in reaching for a glass of water, say. The proxies Alphonso, Beki, Bemi and Kimi rely on provide ways of avoiding both omnidoxasticity and omni-intentionality when acting together.

As our creatures become more complex, their social environments become more complex. There is more variability and less constancy, which makes it more likely that the limits of what can be done with the proxies will matter. Take the triangulation heuristic, for example. According to this heuristic, if you observe that another is linked to you and to the target of a potential joint action, where the joint action is mutually desirable, it is safe to perform the joint action on the assumption that the other will participate (see [sec:alphonso]). With their communicative capacities, Beki's kin are positioned to overcome many limits of this heuristic (as illustrated with the 'two pig' problem in [sec:beki]). However, their descendants, the Kimi, end up using their communicative capacities for deceptive ends. This may create a need to distinguish deceptive from sincere communication and to keep track of how frequently others deceive. Such a need might be met by the emergence of a crude

theory of mind.[15] No less importantly, as our creatures and their descendants get better at navigating social complexity, they are increasingly likely to run up against another kind of limit. For example, action scripts and flexible role assignments allow them to exploit frequent and predictable events. But the better they get at coordinating around these events, the more need they may have to rapidly change their approach. When some of the Kimi chance on a better way of tracking their prey, there is no way for the group to exploit this fortuitous discovery. Introducing the better way of tracking prey cannot be done directly through conventions and action scripts, which can change only gradually. Instead this would require some form of planning ability. As these examples suggest, our creatures' proxies for common knowledge, joint attention and the rest may therefore be drivers of development in this sense: needs for greater cognitive sophistication arise from hitting their limits.

References

Backwell, P., Jennions, M., Passmore, N., & Christy, J. (1998). Synchronized courtship in fiddler crabs. *Nature, 391*(6662), 31–32. https://doi.org/10.1038/34076.

Bratman, M. E. (2000). Valuing and the will. *Noûs, 34*(supplement 14), 249–265. https://doi.org/10.1111/0029-4624.34.s14.13.

Bratman, M. E. (2014). *Shared agency: A planning theory of acting together*. Oxford: Oxford University Press. http://0-dx.doi.org.pugwash.lib.warwick.ac.uk/10.1093/acprof:oso/9780199897933.001.0001.

Bräuer, J., Call, J., & Tomasello, M. (2004). Visual perspective taking in dogs (Canis Familiaris) in the presence of barriers. *Applied Animal Behaviour Science, 88*(3), 299–317. https://doi.org/10.1016/j.applanim.2004.03.004.

Bräuer, J., Call, J., & Tomasello, M. (2008). Chimpanzees do not take into account what others can hear in a competitive situation. *Animal Cognition, 11*(1), 175–178.

Bugnyar, T., Stöwe, M., & Heinrich, B. (2004). Ravens, Corvus Corax, follow gaze direction of humans around obstacles. *Proceedings of the Royal Society of London B: Biological Sciences, 271*(1546), 1331–1336. https://doi.org/10.1098/rspb.2004.2738.

Butterfill, S. (2012). Joint action and development. *The Philosophical Quarterly, 62*(246), 23–47.

Byrne, R. W., & Whiten, A. (1985). Tactical deception of familiar individuals in baboons (Papio Ursinus). *Animal Behaviour, 33*(2), 669–673. https://doi.org/10.1016/S0003-3472(85)80093-2.

Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science, 1*(2), 380–392. https://doi.org/10.1111/j.1756-8765.2009.01026.x.

Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27*(1), 111–133.

---

[15]The emergence of theory of mind might involve the proxies in a further way. It may be that Beki's kin come to recognise the conditions specified by the salience and triangulation heuristics as sufficient conditions for goal-directed action to occur. So it is no longer simply that the heuristics describe what they do: they themselves identify the heuristics as things which should guide their actions. This involves them taking the first steps towards becoming mindreaders. In recognising conditions specified by the heuristics as sufficient for action to occur, they begin to appreciate how which actions someone performs can be influenced by which objects she has a line of sight to and which objects are causally influencing her.

Davis, T. J., Riley, M. A., Shockley, K., & Cummins-Sebree, S. (2010). Perceiving affordances for joint actions. *Perception, 39*(12), 1624–1644. https://doi.org/10.1068/p6712.

De Waal, F. (2016). *Are we smart enough to know how smart animals are?* New York: WW Norton & Company.

Doerrfeld, A., Sebanz, N., & Shiffrar, M. (2012). Expecting to lift a box together makes the load look lighter. *Psychological Research, 76*(4), 467–475. https://doi.org/10.1007/s00426-011-0398-4.

Gergely, G., Nadasky, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165–193.

Gilbert, M. P. (2013). *Joint commitment: How we make the social world*. Oxford: Oxford University Press. http://0-dx.doi.org.pugwash.lib.warwick.ac.uk/10.1093/acprof:oso/9780199970148.001.0001.

Grice, P. (1974). Method in philosophical psychology (from the banal to the bizarre). *Proceedings and Addresses of the American Philosophical Association, 68*, 23–53.

Kaminski, J., Call, J., & Tomasello, M. (2006). Goats' behaviour in a competitive food paradigm: Evidence for perspective taking? *Behaviour, 143*(November), 1341–1356. https://doi.org/10.1163/156853906778987542.

Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: Theory and data. In B. Ross (Ed.), *Psychology of Learning and Motivation* (Vol. 51, pp. 59–101). San Diego: Academic.

Lewis, D. (1969). *Convention: A philosophical study*. Cambridge, MA: Harvard University Press.

Ludwig, K. (2007). Collective intentional behavior from the standpoint of semantics. *Nous, 41*(3), 355–393. https://doi.org/10.1111/j.1468-0068.2007.00652.x.

Obhi, S. S., & Cross, E. S. (2016). *Shared representations: Sensorimotor foundations of social life*. Cambridge, MA: Cambridge University Press.

Okamoto-Barth, S., Call, J., & Tomasello, M. (2007). Great apes' understanding of other individuals' line of sight. *Psychological Science, 18*(5), 462–468. https://doi.org/10.1111/j.1467-9280.2007.01922.x.

Pacherie, E. (2011). Framing joint action. *Review of Philosophy and Psychology, 2*(2), 173–192. https://doi.org/10.1007/s13164-011-0052-5.

Pacherie, E. (2013). Intentional joint agency: Shared intention lite. *Synthese, 190*(10), 1817–1839. https://doi.org/10.1007/s11229-013-0263-7.

Pacherie, E., & Dokic, J. (2006). From Mirror neurons to joint actions. *Cognitive Systems Research, 7*(2–3), 101–112.

Perry, J. (1993). *The problem of the essential indexical: And other essays*. Oxford: Oxford University Press.

Pietroski, P. M. (1998). Actions, adjuncts, and agency. *Mind, New series, 107*(425), 73–111. http://www.jstor.org/stable/2659808.

Rakoczy, H. (2017). Development of collective intentionality. In K. Ludwig & M. Jankovic (Eds.), *The Routledge handbook of collective intentionality* (pp. 407–419). London: Routledge.

Richardson, M. J., Marsh, K. L., & Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance, 31*(1), 62–79. https://doi.org/10.1037/0096-1523.31.1.62.

Richardson, M. J., Marsh, K. L., & Baron, R. M. (2007). Judging and actualizing intrapersonal and interpersonal affordances. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 845–859.

Santos, L. R., Nissen, A. G., & Ferrugia, J. A. (2006). Rhesus monkeys, Macaca Mulatta, know what others can and cannot hear. *Animal Behaviour, 71*(5), 1175–1181. https://doi.org/10.1016/j.anbehav.2005.10.007.

Seyfarth, R. M., Cheney, D. L., & Marler, P. (1980). Vervet monkey alarm calls: Semantic communication in a free-ranging primate. *Animal Behaviour, 28*(4), 1070–1094. https://doi.org/10.1016/S0003-3472(80)80097-2.

Strawson, P. F. (1959). *Individuals*. London: Meuthen.

Suchak, M., Eppley, T. M., Campbell, M. W., Feldman, R. A., Quarles, L. F., & de Waal, F. B. M. (2016). How chimpanzees cooperate in a competitive world. *Proceedings of the National Academy of Sciences, 113*(36), 10215–10220. https://doi.org/10.1073/pnas.1611826113.

Tollefsen, D. (2005). Let's pretend: Children and joint action. *Philosophy of the Social Sciences, 35*(75), 74–97.

Tomasello, M. (2009). *Why we cooperate*. Cambridge, MA: MIT Press.

Varlet, M., Bucci, C., Richardson, M. J., & Schmidt, R. C. (2015). Informational constraints on spontaneous Visuomotor entrainment. *Human Movement Science, 41*(June), 265–281. https://doi.org/10.1016/j.humov.2015.03.011.

Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks, 23*(8–9), 998–1003. https://doi.org/10.1016/j.neunet.2010.06.002.

Visco-Comandini, F., Ferrari-Toniolo, S., Satta, E., Papazachariadis, O., Gupta, R., Nalbant, L. E., & Battaglia-Mayer, A. (2015). Do non-human Primates cooperate? Evidences of motor coordination during a joint action task in macaque monkeys. *Cortex, Special issue: Neurocognitive mechanisms of social interaction, 70*(September), 115–127. https://doi.org/10.1016/j.cortex.2015.02.006.

Warneken, F., Gräfenhain, M., & Tomasello, M. (2012). Collaborative partner or social tool? New evidence for young children's understanding of joint intentions in collaborative activities. *Developmental Science, 15*(1), 54–61. https://doi.org/10.1111/j.1467-7687.2011.01107.x.