

# A stag hunt with signalling and mutual beliefs

Jelle de Boer

Received: 21 September 2012 / Accepted: 28 March 2013 / Published online: 2 April 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** The problem of cooperation for rational actors comprises two sub problems: the problem of the intentional object (under what description does each actor perceive the situation?) and the problem of common knowledge for finite minds (how much belief iteration is required?). I will argue that subdoxastic signalling can solve the problem of the intentional object as long as this is confined to a simple coordination problem. In a more complex environment like an assurance game signals may become unreliable. Mutual beliefs can then bolster the earlier attained equilibrium. I will first address these two problems by means of an example, in order to draw some more general lessons about combining evolutionary theory and rationality later on.

**Keywords** Stag hunt game · Evolutionary game theory · Game theory · Common knowledge · Convention

## Introduction

Imagine two primitive speechless men, say Homo Habilis, side by side staring over grassland at a stag, half a mile away. They seem to be all geared up for a cooperative hunt, ready to go, but that need not prevent us from noticing some problems of an epistemological sort. Firstly, how can one creature be sure that the other is looking at the stag-as-prey, and not just as some brown animal standing there, or as just another fine example of *Cervus Elaphus*, or as something that is not a biped? So one

---

J. de Boer (✉)  
Philosophy Department, University of Amsterdam, Oude Turfmarkt 141-147,  
1012 GC Amsterdam, The Netherlands  
e-mail: j.s.deboer2@uva.nl

J. de Boer  
Philosophy Department, Delft University of Technology, Delft, The Netherlands

question is whether they look at the stag under the same description: the problem of the intentional object. Secondly, given the same intentional object for both, would it be necessary for their cooperation that either creature believes that the other creature believes that it looks at the same intentional object? And should it believe that the other believes that it believes that the other believes...? How many such iterations are required? And where do these mutual beliefs come from? This is the problem of common knowledge for finite minds.<sup>1</sup>

The aim of this paper is to have a look at these problems in tandem. I will argue that subdoxastic signalling can solve the problem of the intentional object as long as this is confined to a simple coordination problem. In a more complex environment, signals may become unreliable. Mutual beliefs can then bolster the earlier attained equilibrium. To address these questions I will continue the story of the hunters started above. I have chosen the two hunters and the stag because I will develop this line into a more complex game theoretical problem, which has lately revived some interest in both philosophy and economics, a version of a stag hunt *game*.<sup>2</sup> The story will be fictional, but it has a function. I use it as a model, as a tool to explore how signalling and mutual beliefs work together in human interaction.<sup>3</sup> During this development there will be various points of contact with empirical literature.

As said, we will address the two epistemological problems mentioned above, but there is a larger methodological issue in the background. This has to do with the question of how to combine evolutionary theory with a theory that has rationality as a central assumption, i.e. classical game theory. An important problem in classical game theory is that there are often too many equilibria. Many scholars bring in evolutionary game theory or simulations with agents with a fixed choice of strategies to reduce the equilibria, to solve the underdetermination, and they make this work. But how? Typically, one begins with some problem like large-scale cooperation in present real life, or as observed by Hobbes or Hume, then note that classical game theory has difficulty in explaining this, and then radically modify the ways of interaction. The possibility of common knowledge is shut down. Actors have only one or a few strategies at their disposal. And then they randomly pair under circumstances with some interesting condition included, for example biased mutation, which drives the whole system towards a happy result. Thus in the end, the agents appear to coordinate or cooperate just fine. But which agents are we now talking about? These latter agents are simple automata they do not look like us at all. So what is this investigation supposed to tell us about the initial question? You cannot just import

<sup>1</sup> Lewis (1969) is the seminal work regarding the statement of these problems in combination. See also Gilbert (1989), and see Radford (1969), Cargile (1970) on the problem of common knowledge for finite minds, and Heal (1978). Cubitt and Sugden (2003) is one noteworthy and fairly recent paper, also on both problems. The central question of their paper, however, is not about what the function of iterated mutual beliefs, what the relevant part of common knowledge, could be—something I will try to argue. We will get back to Lewis at the end of this paper.

The problem of the intentional object is of course a classic in analytic philosophy, with seminal work by Wittgenstein, Quine, Davidson, Kripke, Goodman.

<sup>2</sup> In philosophy, most notably by Skyrms (2004). In economics, see the literature in Sect. 3.

<sup>3</sup> For philosophical legitimating of this use of fiction, see Frigg (2010), Godfrey-Smith (2009). Fine (1993, 2009) provides an anti-theoretical account. In connection with models more specifically in economics see Sugden (2000, 2009) and Morgan (2001, 2004).

some interaction scheme with radically truncated actors somewhere in an analysis to attain a better grasp on, for example, the quandary of large-scale human cooperation. Such intervention should come at a plausible point.<sup>4</sup>

More plausible combinations of classical game theory and evolutionary theory do presume that the agents in the models are made of flesh and blood. Let me briefly mention two recent contributions. Binmore (2008) argues that conventions can often arise as a result of trial and error learning—to be analysed by evolutionary game theory—and that common knowledge is not required in such cases. It may well be the case that fully-fledged common knowledge is often too heavy a presumption to make sense of conventions, but I find it equally hard to believe that mutual beliefs would then not function at all. Gintis (2007), along with a number of related scholars,<sup>5</sup> gives prominence to both the rational actor model and evolutionary theory, whereby the latter helps to explain the content of preferences, and how beliefs, preferences, and constraints are transmitted between individuals. Here also, mutual beliefs do not enter the picture.

The proposals by Binmore and Gintis do not really combine game theory, in the classical sense, but *decision* theory with evolutionary thought. They are parametric, other people are treated as a state of nature—the strategic is left out. It seems that to avoid the heavy presumptions of classical game theory these scholars have thrown away the baby with the bathwater.

In this paper I will first address the problem of the intentional object and the problem of common knowledge by means of an example, in order to draw some more general lessons about combining evolutionary theory and rationality later on.

### Signalling to coordinate

Back to our hunters and the stag. There are no rabbits in view yet—the hunters are in a coordination problem. A zoologist bystander tells us that social mammals, including those capable of cooperative hunting, like lions, hyenas, and chimpanzees, can inform their conspecifics about their intentional objects by a combination of signals. Their head posture and the position of the eyes indicate the focus of their attention, the right coordinates: “*this, there*, however you describe it, is what I am looking at.” And a series of displays betrays the appraisal. A special grunt or a yell, a particular facial expression, some movement of the limbs correlates with anger. Another set correlates with fear, another again with hunger, and so on. The zoologist informs us that evolution has shaped such displays into good indicators of the underlying states.<sup>6</sup> If your neighbour looks angry, be prepared, then he *is* angry. Conversely, if he isn’t angry, he doesn’t look angry—no need to worry, go do

<sup>4</sup> Examples of bringing in parametric choice at an implausible (or unclear) point abound, to take a pick: Dupuy (1989), Gintis (2003), Vanderschraaff (2007) notes the trouble himself. Zollman (2005) also warns that one should be careful with drawing general lessons about human cooperation from very truncated models. Cf. the criticism by Sugden (2001).

<sup>5</sup> See for example Henrich et al. (2004).

<sup>6</sup> See Van Hooff (1972), Van Hooff and Preuschoft (2003) for chimpanzees. See most notably the work by Paul Ekman for humans: Ekman and Friessen (1971), Ekman (2003); and cf. Frank (1988, 2004); Brosig (2002).

something else. In the Darwinian struggle everybody is advised to make few mistakes, either way. Evolution has reduced false positives and negatives.<sup>7</sup>

Let us summarize this knowledge for our case by an invention, a piece of fiction, and say that a certain contraction around the eye socket plus a deep wrinkle above the nose, signify the appraisal in our hunters with sufficient reliability. Screwed up eyes and the nose wrinkle say: “I am not on a taxonomical excursion, I see this as *PREY*.” Thus, we elaborate, these eyes wrinkles + nose wrinkle together *mean* prey in the direction that the creature is looking.<sup>8</sup> We can see this signalling as a primitive kind of communication. It makes possible what Donald Davidson calls triangulation (1982). One can modify one’s perception through the gaze of others. One creature sees a stag-as-prey, but the other just coldly stares in that direction—the contraction wrinkles and the nose wrinkle are absent. This produces a second look. E.g. “Ah... it’s not a stag, it’s one of those gnus—mm... much too dangerous.”<sup>9</sup>

Trusting these signals establishes the default state, we might say. The signals tell us that the conditions for coordination or cooperation are satisfied, unless there is a clue to think otherwise. Without such basis the participants would be at a loss. Thanks to the signalling machinery animals can block epistemological bad tripping: “You look at the same coordinates all right but what do you see, perhaps a stag-as-prey, perhaps not. Then, what do you know about me? I *do* see the stag as a prey, but *you* might think I see it differently. Or you think that I think that you could see it differently.” And so on. Both these problems are circumvented because, firstly, it is clear from the special configuration of wrinkles on one’s partners face what his intentional object is: stag-as-prey. Secondly, the facial expression of your partner betrays what he is up to, his plan of action. It tells that he will just go, and that he is not occupied with belief iterations.<sup>10</sup> Therefore, *you* need not consider belief iteration. (Somewhat puzzled, the zoologist leaves the scene.)

But, we can continue, do these signals themselves not give rise to problems of interpretation? How can one creature know that the other one sees his wrinkles as indicating that he is ready for a kill, and not as conveying sadness, or sexual lust, or as just some interesting geometrical configuration of lines on his face, or...? Aren’t we back at square one? No, there is no such regress here because with our hunters

<sup>7</sup> Albeit not in symmetric fashion. Escaping some actual teeth underway is more important than the problem of jumping up for nothing.

<sup>8</sup> On a teleosemantic account of meaning, e.g. Millikan (1984, 1989). Of course teleosemantic theories of meaning, broadly construed, are not uncontroversial; but not, I suppose, for the restricted domain I am talking about here: primitive signals with an undisputed evolutionary past. Cf. Stegmann (2005).

<sup>9</sup> Compare the literature on joint attention. E.g. Tomasello et al. (2005), Eilan et al. (2005).

<sup>10</sup> Of course this plan of action is not unconditional, it depends on the other. Stag hunting, as I have supposed, is a cooperative project for our creatures. Then how might they ever get started? Stylized, the hunt might be preceded by a two-stage trigger. A) The presence of a stag plus the presence of a comrade in my vicinity causes a half way house ‘stag-as-prey’ action tendency and the corresponding signal in me. B) If this also happens in you, then you transmit the ‘stag-as-prey’ signal to me, and this brings closure, the final confirmation I need: now I am all ready to go. Compare Velleman’s analysis of a shared intention between two individuals (Velleman 1997). It is questionable whether Velleman succeeds because he discusses conditional intentions between rational actors. The crucial point here is that the cooperative project at this stage can be understood in purely causal, subdoxastic, terms.

evolution has done this work already. Differential reproduction has stabilized what the signals mean. Also, under this regime the ensuing equilibria are (local) optima. Creatures which remain through the generations confused as to how to coordinate their online actions will reproduce less, while those who can settle on a shared understanding through some signal will thrive.<sup>11</sup> In line with our hunters, those who ponder a lot about how to interpret the gazes of their possible fellows were more often confined to a lowly diet, and therefore a meager life, especially when it comes to raising a family, i.e. reproduction. Thus evolution has yielded hominoids that bring some ready-made communication equipment to the hunting ground.

### Reliability, lie deception, and mutual beliefs

Enter the rabbits. They jump around in the grassland here and there. They are also prey for our creatures, not as good as a stag, but still tempting. Suppose it takes two to hunt a stag, while one can catch a rabbit on one's own, and that each creature has the following preference ordering, from better to worse:

- I hunt stag, you hunt stag
- I hunt rabbits, you hunt stag (because that leaves more rabbits for me)
- I hunt rabbits, you hunt rabbits
- I hunt stag, you hunt rabbits (which leaves me with an empty stomach)

Then the situation is a stag hunt game, subdivision assurance game<sup>12</sup>:

	Stag	Rabbit
Stag	4,4	0,3
Rabbit	3,0	2,2

The trouble for each creature is that the original plan, the cooperative hunt, is still the best for the both of them, but that one needs some guarantee that this plan has a good chance, a measure of trust that the other one will not be too distracted halfway through by a rabbit scurrying by. Because if that can happen then maybe it would be better to chase the smaller prey, anyway—in accordance with the rule better safe than sorry. Thus there are two motivational pulls: pay off maximization and avoidance of a bad outcome. Correspondingly we have two pure Nash equilibria in this game: [stag, stag] constitutes what is known as the so-called Pareto optimal Nash equilibrium and [rabbit, rabbit] the risk dominant one.

Hence in a one-shot stag hunt game individuals face two options, and there is something to be said for both. Experiments in the lab, however, show that in a series of randomized trials (a series of anonymous one-shot games with different players)

<sup>11</sup> E.g. Skyrms (2004, 2010).

<sup>12</sup> In the other subdivision of the stag hunt game [rabbits, rabbits] has the same pay off as [rabbits, stag] and v.v.

people tend towards rabbit hunting and risk dominance.<sup>13</sup> In early rounds stag hunting is fairly common, more than 50 %, but it gradually gives way to rabbit hunting, and rabbit hunting is what most people do in the end. Why? Presumably because the would-be stag hunters who are occasionally paired off with a rabbit hunter in the early rounds will transform into determined rabbit hunters themselves, once and for all, to avoid further losses; and this will then quickly increase the proportion of rabbit hunters in the following rounds, and thereby the chance of meeting one. Apparently, pay off dominance is a fragile solution concept under these circumstances. Can it be strong otherwise? Under what conditions can [stag, stag] be expected in a randomized series?<sup>14</sup>

This difficulty cannot easily be resolved by signalling because the signal now has a serious credibility problem. As Robert Aumann has pointed out (1990), a sincere stag hunter will surely have a reason to signal ‘stag,’ in order to coordinate with his fellow hunter. However a rabbit hunter in this context will also be motivated to signal ‘stag’, albeit with the wicked intention of having the other fruitlessly stationed in ambush for a while, leaving more small game available for himself. Thus both stag hunters and rabbit hunters will want to signal ‘stag’ in the assurance stag hunt. Therefore the signal carries no information. For our case, the hunters should be wary of others who pretend to hunt stag by *intentionally* pulling the nose wrinkle face and screwing up their eyes and subsequently going after the rabbits. In evolutionary terms, the issue is a group of potentially deceitful rabbit hunters invading a population of stag hunters, signalling ‘about to hunt stag’ and then really hunting for rabbits. A lonely mutant rabbit hunter cannot invade, but a migrating group of rabbit hunters can. This will unsettle the earlier equilibrium. (Migrating groups or tribes are a common factor in the human lineage).

Does Aumann’s conjecture bear out? Charness (2000) set out to test it—that cheap talk cannot be credible in a stag hunt assurance game. He found that the “degree of coordination with a one-way signal is quite high (...). Coordination and efficiency are much lower without a signal, so that the hypothesis that the signal conveys no information is easily rejected.” Clark et al. (2001) also found that communication increases the propensity to hunt stag, but they note that it isn’t sufficient, with rabbit hunting still dominating in the final rounds. Therefore they explicitly support Aumann. The common ground between the papers however is this: signalling has some effect, but it is not strong enough to attain full-scale stag hunting.

Thus simply signalling is not sufficient. An alternative is seeing whether biologist Amotz Zahavi’s ‘handicap principle’ could work, i.e. whether a *costly* signal is possible in this situation, one that is difficult to imitate.<sup>15</sup> A costly and reliable signal would mean that the signal ‘about to hunt stag’ is more expensive in terms of fitness costs for the invading migrants than for our simply truthful stag hunters,

<sup>13</sup> For an overview see Devetag and Ortmann (2007).

<sup>14</sup> A repeated game with a small set of players and no anonymity can also generate the Pareto dominant outcome, but this result quickly unravels with more players and anonymity. Therefore the focus in the literature on randomized series, these make up the challenge.

<sup>15</sup> For a recent restatement see Zahavi (2003).

while the benefits remain the same. Migrant hunters will then fix on a lower level of signalling. This then yields the possibility for a stag hunter, who is looking for an effective companion, to discriminate between these types, a way to tell who is speaking the truth and who is lying.<sup>16</sup> Let us pursue this possibility. How could it be true that individuals from a group of deceitful rabbit hunter migrants pay more for sending the same signal? One possible answer is to say that there is neural machinery in place in the original stag hunter—with a cost already paid for during the time when the only trouble was to coordinate, and that for the new arriver, building a structure that would produce a similar signal would now come at a higher cost. This obviously raises the question of why this new cost would be higher.

To be sure, bodily expressions in real hominids (humans, many primates) *do* work well, and it is also correspondingly true that humans have developed a capacity to discriminate between a bogus signal and a real one. This is the case because, indeed, the real signal is typically produced by a neural structure different to the bogus one, thereby, again typically, yielding a noticeable difference between the two.<sup>17</sup> Therefore we can often (sometimes?) tell a spontaneous smile from a studied one. So in this area perhaps something like an arms race has been going on—real signals yielding bogus signals yielding finer powers of discrimination.

If the arms race is true then it seems that the deceivers have just had their turn. Because the question remains: would it really be so costly to mimic the ‘about to hunt stag’ signal? To put it rhetorically, what are the costs of two or three lessons in Method acting? On an empirical note, detecting lies from just bodily cues *is* difficult when the liars are really intent on getting their false message across (see Ekman 1971). So let us conclude that the signalling machinery is probably insufficiently failsafe in the assurance game context, and that it needs backup.

For an insight into what sort of form this backup might take, let us see how some real animals deal with this issue, when there is a relatively easy or cheap signal combined with conditions that favour deception. Can this combination be found? Yes, in various species.<sup>18</sup> A male barn swallow, for example, hurriedly cries an alarm call when he finds his female gone and their nest empty. The call causes all nearby swallows to take flight, thereby effectively disrupting a possible act of extra-pair mating—his wife making love somewhere else?—and thus, the evolutionary rationale, hopefully securing a degree of paternity.<sup>19</sup> Tufted capuchin monkeys sometimes yell an alarm call to send off dominant conspecifics to some hiding place in order to get hold of more food for themselves.<sup>20</sup>

Deception exists in the animal kingdom, between members of the same sort, but of course there must be an upper limit to his behaviour somehow, since it can only

<sup>16</sup> One could object here that this adding costs for one party alters the pay-offs of our original game, as these are supposed to be *net* figures, benefits minus costs. But why should that matter? The cause of this change in pay-off structure could simply be a relevant fact. There is no rule that says that an analysis of a history of interaction should restrict itself to one and only one game form from beginning to end.

<sup>17</sup> See the literature mentioned in note 5.

<sup>18</sup> For an overview see Searcy and Nowicki (2005).

<sup>19</sup> Møller (1990).

<sup>20</sup> Wheeler (2009).

obtain against a background of true signalling. It cannot be the case that everybody is going to cry wolf whenever it suits him, because then everybody else would have reason to start ignoring this signal, rendering it superfluous in due time. So what, if not the handicap principle, could be a check on deception spreading too widely in a population? Here I follow a recent work by two behavioural biologists, Searcy and Nowicki (2005). One good trick is not only to attend to the signal, and then either respond or ignore, but also keep one's eye on the signaler. Has he fooled me (or others) in the past? If so, then I'll be less inclined to take him seriously now. "Individual directed skepticism," as Searcy and Nowicki call it, with a capacity for registering and remembering individuals for their recent actions, can thus keep deceit under control and maintain the function of a cheap signal by effectively putting the deceivers on the sidelines, counting them out of the game. The exercise of this capacity has been experimentally demonstrated in a couple of studies. Domestic chickens for example, and also vervet monkeys, can quickly learn to discriminate between reliable and unreliable individuals, those who truly and fully report danger and those who cry wolf in order to profit from others' temporary absence. After a couple of such incidents, they cease to respond to the latter.<sup>21</sup>

This suggests that the stag hunting equilibrium can also be restored thanks to this mechanism by sufficiently weeding out the false-signalling invaders. That would be a genetic evolutionary effect. But there is arguably also a second equilibrium restoring mechanism, namely learning on the part of the signalers. We are now assuming that individuals in our stag hunt are capable of registering and remembering deceitful conspecifics. They respond during their lifetimes. If this can happen then it will presumably not take long until the deceivers also start responding during their lifetimes and stop lying. Then we could say that those with a risk dominant tendency who first tried out deception now *learn* from their reduced achievements in this new environment. They learn that the lying strategy goes wrong. Suppose that this is so. Then we have grounds for saying that these individuals—the deceivers, the originally deceived subsequently ostracizing these deceivers, and then the deceivers mending their ways—are together climbing the winding stairs of common knowledge. Let me picture this evolving intentionality by four subsequent stages. I use intentional language already at stage i) and it could be objected that this is not strictly necessary. From some point on, however, ascribing beliefs to individuals who take account of other individuals who take account of them becomes appropriate and not doing so increasingly cumbersome.<sup>22</sup> The crucial point is that wherever one exactly starts using intentional language, deceit and the possibility to respond during one's lifetime then introduces a *next* level, a *higher* level, and then the counter measure *again* a higher level.

- (i) Coordination. At the first stage, when there are no rabbits yet, when there is simply just the problem of coordinating on stag-as-prey, let us say about one hunter (X) who shoots a glance at the killing gaze, directed at the stag, of his

<sup>21</sup> For references see Searcy and Nowicki (2005, pp. 75, 76).

<sup>22</sup> Couldn't it all be accounted for in terms of conditioned response? See Sect. 4.



companion (Y) that:

*X believes that Y is after the stag.*

- (ii) Deception. At this stage the rabbits have entered and now Y, a risk dominant player, deceives X into believing that Y is after the stag. Therefore, given (i), Y believes this about X:

*Y believes that X believes that Y is after the stag.*

- (iii) Recognition of deception. At this stage X shuts Y out because he has learnt to see through the deception. So X understands what Y must have been busy with believing about him (X), X understands what Y believes at (ii):

*X believes that Y believes that X believes that Y is after the stag.*

- (iv) Recognition of recognition of deception. Then Y in turn understands situation (iii), and what X's believes in it (and subsequently mends his ways):

*Y believes that X believes that Y believes that X believes that Y is after the stag.*

At stage (iv) with the original deceivers about to mend their ways we have fourth order mutual beliefs. Written as *Y believes that X believes that Y believes that X believes that Y is after the stag*, it is pretty mind boggling but that has much to do with the abstract formulation. Let's formulate it differently. Consider again the once deceiver at stage iv, what makes him mend his ways? Well, the convert now understands the position of the others, the sincere stag signalers. What is their position? That they see through his plan and his corresponding thoughts. What were these? He thought that they believed that he would hunt stag. Nested, at stage iv: fourth order.

More generally, deceit and counter measures in one generation open up the possibility of higher order mutual beliefs.<sup>23</sup> Conversely, one way to answer the question of how much belief iteration is required, the problem of common knowledge for finite minds, is to begin with a ground floor of subdoxastic signalling and then ascend as far as it is rationally required for individuals to restore trust.

It could be objected that kin selection is a stronger candidate for attaining stag hunting equilibrium. To answer this, let us first suppose that in the world we began with, with speechless hominoids trying to coordinate their actions, kindred individuals *also* need signals to fix their intentional objects. The signal that says *stag-as-prey* usefully extends beyond kin relations. Now it is true that potentially invading deceitful rabbit hunters can subsequently be blocked by kin selection but then we should note that this leaves a group of former stag hunters behind (which is not the case with the present proposal), namely, those who earlier managed to hunt

<sup>23</sup> This has a foothold in the literature on the origins of human intelligence: what is known as the Machiavellian intelligence hypothesis. Scholars who wonder about the evolutionary development of human intelligence focus on deception. Social animals benefit from group life, from economies of scale, division of labour, and cooperative work. But social life also brings various sources of conflict amongst the individuals: over food, mating access, who bosses whom, who are allies, and so on. These intricacies imply that it becomes more and more important to have in mind not only what others are doing, and what they might be up to, but also what they think: beliefs about beliefs, in other words. In this field, tactical deception has become a paradigm for demonstrating such higher order beliefs. It has become a standard for cognitive performance in comparative psychology. Seminal work is Byrne and Whiten (1988) and Dunbar (2003). Cf. Sterelny (2006). In line with this, note that the stag hunt story has now moved beyond *Homo Habilis*.

stag and are not fortunate enough to be related by kinship. They will now become rabbit hunters. Since kin selection has restricted scope, it is now in this *new* subgroup of deceived stag hunters that the evolvement of monitoring, ostracizing, learning and the onset of mutual beliefs could take place. This illustrates the familiar point that, given relatively large-scale cooperation, kin selection can only provide a partial explanation.

Mutual beliefs have a larger reach. As is well known, an extension beyond the two-person case normally puts pressure on a cooperative convention. How could that be remedied under the view in development? It is beyond the scope of this essay to elaborate on it but I would be inclined to look first at low cost information transfer (instead of for example punishment, cf. Guala 2012). Someone has falsely signalled ‘stag’ in for example a five-person stag hunt? Then, if we suppose that it is like a real group hunt, the others will quickly learn this: four at a time. Assume that they pass this information on to others—gossip—so that our deceiver will also be known in many other future group compositions. With the same capacity to learn from this he then would also be well advised to mend his ways and signal honestly next time. So such information flow in larger settings combined with learning and the possibility of ostracism might considerably help to stabilize stag hunting. Note that in such a context it is not necessary that two individuals develop their mutual belief by means of direct interaction. These may originate from a third (fourth, fifth...) source. Hence the large scope of the learning-cum-mutual-beliefs mechanism, which, furthermore, is also a faster mechanism.

Zollman (2005) demonstrates on the basis of simulations that signalling combined with spatial structuring can lead to a stag hunting equilibrium. In these simulations, all individuals are on a grid and they only interact with neighbors. As for this mechanism, I also doubt whether it is strong. Evolutionary stag hunt game simulations with simple agents and spatial structuring show interesting results but I find it hard to estimate their worth and what exactly the analogy for human lineage should be.<sup>24</sup> E.g., the members of this lineage seem to be not at all spatially fixed and structured as in these simulations. Humans are very mobile, as groups and as individuals within groups.

That signalling and mutual beliefs together can produce [stag, stag] in humans is shown in the laboratory by experimental economists Duffy and Feltovich (2006). They worked with one-way signals, thus only one party could send a signal, with the other party just receiving and not sending himself. Now in one of their testing conditions people not only received a one-way signal of what the other player was currently doing, they could also see what the same player had signaled *and* done in the previous round. In other words, they could see whether their counterparts had lied or spoken honestly in the previous round. This further increased stag hunting significantly in cases when all information pointed in the same direction (e.g. previous round: signal was ‘stag’ and behaviour was ‘stag’; current round: signal is ‘stag’).<sup>25</sup> Interestingly, successful stag hunting works mainly through the senders

<sup>24</sup> To be fair, Zollman is aware of this problem (see note 4).

<sup>25</sup> Unfortunately, Duffy and Feltovich (2006), while relating to Aumann’s work (in Duffy and Feltovich 2002), did not choose a stag hunt of the assurance type. They did, however, compare a simple stag hunt with a chicken game and a prisoner’s dilemma. The trends in the results with the simple stag hunt and the

who keep to their words. Responders turn out to be not very sensitive to their counterparts' messages from the previous round. Crudely, stag hunters on the senders' side are mainly to be found among those who did a triple 'stag', while stag hunters on the responders' side are not very discriminatory.

Duffy and Feltovich offer the following interpretation: "Consequently, while senders' previous round messages do provide additional information about their subsequent actions, the value of this information is relatively small and may be overwhelmed by the cognitive costs of processing it. This suggests that the common knowledge in our experiment that lie detection is possible—via the observation of previous round messages—is sufficient to constrain the behaviour of senders to the point that the receivers need not be so careful about checking for lies." (2006: 685)

These empirical findings support my claim about how signalling and mutual beliefs work together in attaining the collectively best outcome, with the qualification of course that the structure of common knowledge was only necessary up to a point in this case, namely second order mutual beliefs.<sup>26</sup>

At this point it could be questioned whether the signalling is still of any use. Because now that monitoring, memory, and mutual beliefs have been established, cannot this new machinery do the whole job, rendering the signals superfluous? People could just observe each other's behaviour—and see who are stag hunters and who are rabbit hunters—remember, and use this knowledge when they meet. A first conjecture is that this does not seem to be true of our species. Communication has not atrophied in strategic settings, it *does* seem to be functional.<sup>27</sup> But why? One important factor is that reading behaviour comes with a measure of ambiguity. Suppose you saw me rabbit hunting in the previous round. That need not imply that I am a risk dominant chooser. After all, I could have been knowingly paired off with a rabbit hunter in that very round. In such a case the appropriate response would have been 'hunt rabbits'. So I could have genuine stag inclinations right now.

Signalling and lie detection leave less room for ambiguity. Suppose we take part in this experiment and you see that I have lied in the past. You conclude that I am a rabbit hunter in disguise. Or, you now have doubts, could it be so that I have genuine stag inclinations and that I have only lied in the previous round because I was paired off with a liar myself? But then the question is why would someone with stag inclinations respond with a lie in such a case. Deception works on stag hunters, not on rabbit hunters. From a pay-off dominant chooser's point of view lying should not make much sense. So for him it is not so that lying is the appropriate response to lying. So there seems to be less equivocation possible: detecting a lie indicates that you are dealing with a rabbit hunter.

---

Footnote 25 continued

PD are such that I think it is safe to interpolate: what is true of *both* their stag hunt and their PD is arguably true of an assurance stag hunt too.

<sup>26</sup> The laboratory context is of course different than our imagined stag hunt story but I submit that the relevant *relationship* between signalling and mutual beliefs, i.e. how these mechanisms work together, is similar.

<sup>27</sup> For a review see Sally (1995), see also Frank (2004).

## Grounds for mutual beliefs

Of course, mutual beliefs can arise from other conditions, too. For example, the two hunters, while still in the coordination phase without the rabbits, part ways, one heads west to lie in ambush, the other goes east to startle the stag into moving into the ambush. Contact is lost, but the representations of each other's focused gaze could still be reassuring enough. Then, here is the other condition, a black raven suddenly flies up somewhere in the middle of the grassland, visible to both. Is this a bad omen, one hunter thinks, giving them a reason to stop the hunt immediately? Well, he is not superstitious, nor, he thinks, is the other hunter. And he also believes that the other doesn't believe that he is superstitious. He believes that the other believes that he will hunt stag. This is a second order mutual belief.

On the other hand, not just anything will spur mutual beliefs. The black raven will only have this potential to unsettle matters if there is a fair chance that both hunters see the black raven as, indeed, a possible omen, because they may also see it as just some black bird, or as just another fine example of *Corvus corax*, or as something that is mostly not a fish. So our issues return. Firstly, do the hunters look at the raven under the same description? Secondly, the mutual beliefs iteration question arises. Would it be necessary for either hunter to believe that the other believes that it looks at the same intentional object, and so on: how many orders of this? Thus a black raven can yield mutual beliefs, but this in itself then also invites an explanation.

With the black raven flying up in the middle of the hunting ground mutual beliefs that keep on supporting the [stag, stag] equilibrium result from *other* mutual beliefs, namely about black ravens being an omen or not. The point is that one should be able to explain, at least in principle, where these mutual beliefs come from. What are their grounds? There should be a clue that our two hunters share a recent history in which black ravens were conventionally thought of as bringing bad luck, and that each now has a reason to reflect on this. With the signalling and subsequent reliability problem induced by the rabbits, I have tried to sketch an alternative case wherein *evolutionary* pressure gives rise to lie deception, ostracism, a return to honest signalling, and thereby the cognitive sophistication of mutual beliefs.

Combining: mutual beliefs originating from other mutual beliefs do not constitute a vicious regress, if some mutual beliefs are prompted by differential reproduction.<sup>28</sup>

Another issue is that a good description of what happens in our fictional story, with some hunters in a deceiving mood and others actively taking counter measures, does not really require mutual beliefs at all. One *may* ascribe propositional attitudes to the hunters, but it would be better to leave them out, if a simpler hypothesis can cover the same facts: namely trial and error learning or a history of conditioned response. Then there would be no grounds for talking in terms of mutual beliefs. On this approach a liar would be confronted with negative reinforcement when the

<sup>28</sup> Compare on the origin of convention Cubitt and Sugden (2003: 203) remarking that in the end "at least some inductive standards could be common," for example certain "innate tendencies to privilege certain patterns when making inductive inferences," with such tendencies then of course being the product of an evolutionary process.

others turn their backs on him. Then at some moment he randomly switches to honest signalling again and this, in due course, triggers positive feedback. This can be modelled by evolutionary game theory if one wants (e.g. with mutating rules as replicators). See, no need for fancy beliefs about beliefs! This is of course a well-known objection, in behaviouristic psychology (no mental concepts!) and also in ethology (mental concepts are human, if anything, no anthropomorphism!).

Here is a controversial ethological case, of primate deception and subsequent countering, in two rounds. Chimpanzee Belle was the only one who knew where a certain amount of food was hidden. The experimenters made sure they only showed her. But the other chimpanzees soon realized Belle's position, and they began to follow Belle. Thus Belle responded by sometimes leading the others away and then waiting until it was safe to go to the food. On this, one of the other chimpanzees, Rock, responded in turn by pretending he was indeed walking away, only to suddenly spin around a few steps later to see where Belle was going ('on her way to the food already?').<sup>29</sup> If this really was Rock's strategy, then he must have been involved in third order cognition!

Rock believes {Belle believes [Rock believes (Belle is not after the food right now)]}.

Rock seems to be fantastically smart, but a sceptic would query whether it is really necessary to ascribe such a higher order mutual belief to Rock. E.g. one could also explain observations like this as being outcomes of long sequences of conditioned response ('every time I go there, she goes there and this brings me food' etc.). But note that cases like deception and counter-deception can always be interpreted in the minimal style, as an outcome of a history of conditioned responses.

Such an explanation in terms of conditioning becomes less convincing, however, when subjects show the same cognitive performance across a range of situations, and when they show it in novel cases. Even then a conditioning account may still be possible but the argument of simplicity will now significantly gain weight on the side of the intentional explanation. This is what recent work on primate cognition demonstrates, and in particular the debate on whether and to what degree chimpanzees have a theory of mind. On the one hand, there are researchers (Michael Tomasello and co-workers) who conclude on the basis of their experiments that chimpanzees can understand to some extent what others have in mind, e.g. what another chimp perceives at a certain moment, or what the goal of his actions is, or what he knows about a specific situation. E.g. Tomasello et al. (2003); Tomasello and Call (2006). On the other hand, there are researchers (Daniel Povinelli and co-workers) who defend and restrict themselves to the conditioning theory. For any piece of seemingly smart chimp behaviour that may invite an explanation in mental terms from the first camp, the conditioning theorists put forward a hypothesis that is wholly framed in terms of outward behavioural properties, e.g. gaze following,

<sup>29</sup> The experiment is from 1973, described by Tomasello and Call (1997) p. 237, 238. Cf. Byrne and Whiten (1988).

acting on someone else's body or face orientation, monitoring the movements of others, and so on. E.g. Povinelli and Vonk (2006).

To illustrate with Belle and Rock, at an early stage of the experiment with Belle the conditioning theorists for example could claim that the other chimps are just following Belle's gaze until it hits an object. As this defines their target, they go there and this produces a reward. However, this story clearly cannot account for how Rock responds to what Belle did next, to her wandering about. Rock did *not* go to the end point of Belle's gaze. The conditioning defenders may then instead say that Rock follows other surface properties, e.g. Belle's movements for a certain period. Tomasello and his co-workers argue, convincingly I believe, that such a typical manoeuvre of the conditioning camp is not only ad hoc, it then also makes one to invoke different sets of causal factors for different situations, and this increases the complexity of the explanation.<sup>30</sup>

To illustrate with our hunters, suppose that some of the risk dominant oriented hunters have fine-tuned their false signalling, and choose to lie only in special circumstances, when temporarily among the old and forgetful, say. The elderly will be angry when finding out that they have been lied to, but it will be of little consequence because the next time a stag comes into view they will have forgotten the earlier deceit. Among the old and forgetful, such a sophisticated liar then believes that the others believe that he will hunt stag (which he will not). Now imagine that some of the seniors, being perhaps old and forgetful but surely not stupid, see through this: "Ah, I believe that he believes that I believe that he will hunt stag. But I don't." This is a third order mutual belief. In the past this senior has learnt the technique of nesting beliefs, and now he applies it to a new case. Such an old and forgetful subject who outsmarts a cunning liar is quite conceivable, I hope. He would provide a case in point against the conditioning hypothesis. This individual has not gradually learnt by trial and error that he should be on his guard whenever he is in a situation with only other old and forgetful subjects around him. He cannot learn this, he forgets things.<sup>31</sup> He rationally understands, he sees through the liar. He believes that the liar believes that he believes that he will hunt stag (which he will not).

The lesson is: mutual beliefs can arguably be invoked when individuals demonstrate corresponding behaviour across various contexts and in novel situations, in such cases conditioning should not be considered the default hypothesis.

I have sketched a stylized history of creatures first in a coordination problem and at a later evolutionary stage in an assurance game setting to see how signalling and mutual beliefs work together and how mutual beliefs might have evolved from signalling. A real natural history of the capacity for higher order intentional states would trace the evolutionary steps at specific times, with each step specifying the cognitive capacities as solutions to earlier environmental selection pressures.

Human beings qualify as creatures at the end point of this history: we have the cognitive equipment to apply this combination to the stag hunt game. This is

<sup>30</sup> E.g. Tomasello and Call (2006), Call and Tomasello (2008).

<sup>31</sup> I assume here that forgetfulness includes such failure of conditioned learning.

shown in the economist laboratory experiments that I have discussed. But we are arguably a little overqualified for this task. Rock and Belle, and the ethological and comparative psychological literature, make a case that humans are probably not unique in having this capacity. Now exactly which organisms are capable of higher order intentional states (chimps, dolphins, elephants ...) and with what content (someone else's perceptual state, goal of action, knowledge ...) are both the subject of current empirical research.<sup>32</sup> How this comparative psychology should inform an evolutionary account of intentionality in the human lineage is not a straightforward matter, as it seems reasonable to assume that this cognitive skill might have emerged independently in different species (think of the dolphins). And as we saw, it isn't easy to establish what are exactly the minimum cognitive requirements to solve a puzzle to begin with. So there are many important questions ahead of us.

My own ambition has been to sketch a model that shows how signalling and mutual beliefs may function together, and present this as a credible mechanism for attaining cooperation in a game like the stag hunt; and thereby developing an example of how evolutionary theory can sensibly combine with rationality theory.

## Conclusion

In *Convention*, David Lewis claimed that common knowledge requires a mutual ascription of common inductive standards and background information, rationality, and mutual ascription of rationality (1969: 56). This seems to be correct, but of course one may query where all this comes from, does Lewis not presuppose a little too much? It is probably true that English-speaking people have concordant expectations about stag, calling them by the same name 'stag', but before they can arrive at this language convention style-Lewis, they should first match their expectations on what they perceive. Stag-as-taxonomical-object possibly puts the next stag in a different category than stag-as-prey. There are always, as Lewis observed, "innumerable alternative analogies" (1969: 37–38). Hence, a particular problem of common inductive standards, our problem of the intentional object, comes first. Do people look at the same object, under the same description? One way, as we saw, to nail down an answer is subdoxastic signalling.

Common knowledge is an infinite structure in logical space. For conventions Lewis understands this infinity in terms of potential reasons.<sup>33</sup> This may be fine for logical purposes, but then the question remains how we, finite beings, connect with this.<sup>34</sup> What portion of this structure do people actually need? In our stag hunt

<sup>32</sup> See for example Tomasello et al. (2005) for a comparison between chimpanzees and human infants, and Connor and Mann (2006) for a review paper on dolphins.

<sup>33</sup> See Cubitt and Sugden (2003) on how this structure could be generated.

<sup>34</sup> Compare Sillari (2005) for a formal answer. Which portion of the infinite structure that becomes actual, Sillari says, has to do with what is deemed irrelevant, or what cannot be handled for lack of computational power, or for psychological reasons, and so on. He then develops a tool, 'awareness structures,' to take formally account of such heuristics. My approach is different but could be regarded as

signalling lays a subdoxastic foundation, it makes up the basics of trust. Then, given this basis, a possible complexity like the reliability problem induced by the rabbits requires lie deception and active ostracism to restore trust, and this establishes how much of the structure of common knowledge, how much iteration, is required. This, at least in part, answers our second problem.

I have used facial expressions as signals with an evolutionary origin, but that is just an example. I think it is not difficult to broaden my case here, and claim that many simple signals can be understood in evolutionary terms.<sup>35</sup> So, more generally, evolutionary-based signals can solve coordination problems without invoking common knowledge. Then, as typically happens in evolution, these solutions also function in new contexts, fail or partly fail, and then call for the development of new machinery. The signalling-solution of the intentional object coordination problem carries over to a new situation, like an assurance stag hunt game, in which signals are unreliable, and then mutual monitoring and mutual beliefs can restore the earlier equilibrium.

Evolutionary thinking has increasingly good press in social science. More specifically, it seems that combinations of classical and evolutionary game theory have a promising future. But we should be careful about how the two theories go together. I propose the following questions as a methodological check:

- (i) *What can plausibly be argued to have an evolutionary basis?* In my story signalling had an evolutionary basis. The signalling brought about coordination while avoiding the common knowledge problem. An exchange of subdoxastic signals was causally sufficient to attain [stag, stag].
- (ii) *How do mutual beliefs, the strategic, enter the picture?* With the rabbits, signalling involved a reliability problem, and then trust was restored through lie deception, active ostracism, and a subsequent return to honest signalling. This gave rise to mutual beliefs. Given the signalling basis mutual beliefs operated as a trust-restoring mechanism.
- (iii) *Check (ii) against the conditioning hypothesis (for which evolutionary modelling applies).* In my model story, individuals with good performance but bad memories undermined the conditioning hypothesis.

**Acknowledgments** Many thanks to Govert den Hartogh and Gijs van Donselaar.

## References

- Aumann R (1990) Nash equilibria are not self-enforcing. In: Gabszewicz J, Richard J, Wolsey L (eds) Economic decision making, games, econometrics and optimization. Elsevier, Amsterdam
- Binmore K (2008) Do conventions need to be common knowledge? *Topoi* 27:17–27

---

Footnote 34 continued

contrastive: given a subdoxastic basis, what could be, not irrelevant, but exactly *relevant* for mutual belief iteration, what could be grounds for ascending in the common knowledge structure.

<sup>35</sup> Some theoretical possibilities, not mutually exclusive: signals are products of trial and error learning (evolutionary game theory applies), signals are replicators *and* units of selection (memetics), a teleological theory of meaning restricted to more or less simple signals (cf. Sterelny 1990).



- Brosig J (2002) Identifying cooperative behavior: some experimental results in a prisoner's dilemma game. *J Econ Behav Organ* 47:275–290
- Byrne R, Whiten A (eds) (1988) *Machiavellian intelligence: social expertise and the evolution of intellect in monkeys, apes, and humans*. Oxford University Press, Oxford
- Call J, Tomasello M (2008) Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn Sci* 12:187–192
- Cargile J (1970) A note on “iterated knowings”. *Analysis* 30:151–155
- Charness G (2000) Self-serving cheap talk: a test of Aumann's conjecture. *Games Econ Behav* 33:177–194
- Clark K, Kay S, Sefton M (2001) When are Nash equilibria self-enforcing? An experimental analysis. *Int J Game Theory* 29:495–515
- Connor R, Mann J (2006) Social cognition in the wild: Machiavellian Dolphins? In: Hurley S, Nudds M (eds) *Animal rationality?*. Oxford University Press, Oxford
- Cubitt R, Sugden R (2003) Common knowledge, salience and convention: a reconstruction of David Lewis' game theory. *Econ Philos* 19:175–210
- Davidson D (1982) Rational animals. *Dialectica* 36:317–327
- Devetag G, Ortmann A (2007) When and why? A critical survey on coordination failure in the laboratory. *Exp Econ* 10:331–344
- Duffy J, Feltovich N (2002) Do actions speak louder than words? An experimental comparison of observation and cheap talk. *Games Econ Behav* 39:1–27
- Duffy J, Feltovich N (2006) Words, deeds, and lies: strategic behaviour in games with multiple signals. *Rev Econ Stud* 73:669–688
- Dunbar R (2003) Evolution of the social brain. *Science* 302:1160–1161
- Dupuy J (1989) Common knowledge, common sense. *Theory Decis* 27:37–62
- Eilan N, Hoerl C, McCormack T, Roessler J (eds) (2005) *Joint attention: communication and other minds*. Oxford University Press, New York
- Ekman P (1971) *Telling lies*. W.W. Norton, New York
- Ekman P (2003) *Emotions revealed*. Henri Holt, New York
- Ekman P, Friesen W (1971) Constants across cultures in the face and emotion. *J Pers Soc Psychol* 17:124–129
- Fine A (1993) Fictionalism. *Midwest Stud Philos* 18:1–18
- Fine A (2009) Science fictions: comment on Godfrey-Smith. *Philos Stud* 143:117–125
- Frank R (1988) *Passions within reason: the strategic role of the emotions*. W.N. Norton, New York
- Frank R (2004) Can cooperators find one another? In: *What price the moral high ground? Ethical dilemmas in competitive environments*. Princeton University Press, Princeton
- Frigg R (2010) Models and fiction. *Synthese* 172:251–268
- Gilbert M (1989) Rationality and salience. *Philos Stud* 57:61–77
- Gintis H (2003) Solving the puzzle of human sociality. *Ration Soc* 15:155–187
- Gintis H (2007) A framework for the unification of the behavioral sciences. *Behav Brain Sci* 30:1–61
- Godfrey-Smith P (2009) Models and fictions in science. *Philos Stud* 143:101–116
- Guala F (2012) Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. *Behav Brain Sci* 35:1–59
- Heal J (1978) Common knowledge. *Philos Q* 28:116–131
- Henrich J, Boyd R, Bowles S, Camerer C, Fehr E, Gintis H (eds) (2004) *Foundations of human sociality: economic experiments and ethnographic evidence from fifteen small-scale societies*. Oxford University Press, Oxford
- Lewis D (1969) *Convention*. A Philosophical Study. Basil Blackwell, Oxford
- Millikan R (1984) Language, thought, and other biological categories. MIT Press, Cambridge
- Millikan R (1989) Biosemantics. *J Philos* 6:281–297
- Møller A (1990) Deceptive use of alarm calls by male swallows, *Hirundo rustica*: a new paternity guard. *Behav Ecol* 1:1–6
- Morgan M (2001) Models, stories, and the economic world. *J Econ Methodol* 8:361–384
- Morgan M (2004) Imagination and imaging in model building. *Philos Sci* 71:753–766
- Povinelli D, Vonk J (2006) We don't need a microscope to explore the chimpanzee's mind. In: Hurley S, Nudds M (eds) *Animal rationality?*. Oxford University Press, Oxford, pp 385–412
- Radford C (1969) Knowing and telling. *Philos Rev* 78:326–336
- Sally D (1995) Conversation and cooperation in social dilemmas: a meta-analysis of experiments from 1958 to 1992. *Ration Soc* 7:58–92

- Searcy W, Nowicki S (2005) The evolution of animal communication. Reliability and deception in signaling systems. Princeton University Press, Princeton
- Sillari G (2005) A logical framework for convention. *Synthese* 147:379–400
- Skyrms B (2004) The stag hunt and the evolution of social structure. Cambridge University Press, Cambridge
- Skyrms B (2010) Signals. Evolution, learning and information. Oxford University Press, Oxford
- Stegmann U (2005) John Maynard Smith's notion of animal signals. *Biol Philos* 20:1011–1025
- Sterelny K (1990) The representational theory of mind: an introduction. Blackwell, Cambridge
- Sterelny K (2006) Folk logic and animal rationality. In: Hurlley S, Nudds M (eds) *Animal rationality?*. Oxford University Press, Oxford, pp 293–311
- Sugden R (2000) Credible worlds: the status of theoretical models in economics. *J Econ Methodol* 7:1–31
- Sugden R (2001) The evolutionary turn in game theory. *J Econ Methodol* 8:113–130
- Sugden R (2009) Credible worlds, capacities, and mechanisms. *Erkenntnis* 70:3–27
- Tomasello M, Call J (1997) *Primate cognition*. Oxford University Press, New York
- Tomasello M, Call J (2006) Do chimpanzees know what others see—or only what they are looking at? In: Hurlley S, Nudds M (eds) *Animal rationality?*. Oxford University Press, Oxford, pp 371–384
- Tomasello M, Call J, Hare B (2003) Chimpanzees understand psychological states—the question is which ones and to what extent. *Trends Cognit Sci* 7:153–156
- Tomasello M, Carpenter M, Call J, Behne T, Moll H (2005) Understanding and sharing of intentions: the origins of cultural cognition. *Behav Brain Sci* 28(2005):675–735
- van Hooff J (1972) A comparative approach to the phylogeny of laughter and smiling. In: Hinde R (ed) *Non-verbal communication*. Cambridge University Press, Cambridge
- van Hooff J, Preuschoft S (2003) Laughter and smiling: the intertwining of nature and culture. In: de Waal F, Tyack P (eds) *Animal social complexity*. Harvard University Press, Cambridge
- Vanderschraaff P (2007) Covenants and reputations. *Synthese* 157:155–183
- Velleman J (1997) How to share an intention. *Philos Phenomenol Res* 57:29–50
- Wheeler B (2009) Monkeys crying wolf? Tufted capuchin monkeys use anti-predator calls to usurp resources from conspecifics. *Proc Royal Soc* 276:3013–3018
- Zahavi A (2003) Indirect selection and individual selection in sociobiology: my personal views on theories of social behaviour' (anniversary essay). *Anim Behav* 65:859–863
- Zollman K (2005) Talking to neighbors: the evolution of regional meaning. *Philos Sci* 72:69–85