Game Theory in Biology and Anthropology

Alihan Kabalak
Elena Smirnova
Juergen Jost

SFI WORKING PAPER: 2014-06-019

Game theory in biology and anthropology

Alihan Kabalak[1], Elena Smirnova[2], and Jürgen Jost[1,3]

[1] Max Planck Institute for Mathematics in the Sciences,
Inselstrasse 22, D-04103 Leipzig, Germany
[2] German Department, Leibniz Universität Hannover,
Königsworther Platz 1, D-30167 Hannover, Germany
[3] Santa Fe Institute for the Sciences of Complexity,
1399 Hyde Park Rd., Santa Fe, NM 87501, USA

June 9, 2014

Abstract

The readiness for spontaneous cooperation together with the assumptions that others share this cooperativity has been identified as a fundamental feature that distinguishes humans from other animals, including the great apes. At the same time, cooperativity presents an evolutionary puzzle because non-cooperators do better in a group of cooperators. We develop here an analysis of the process leading to cooperation in terms of rationality concepts, game theory and epistemic logic. We are, however, not attempting to reconstruct the actual evolutionary process. We rather want to provide the logical structure underlying cooperation and understand why cooperation is possible and perhaps even necessary.

1 Introduction

What distinguishes humans from (other) animals? This is an old and much debated question. Clearly, there are both cognitive and social aspects, and these are intricately intertwined. Just think of the issue of "social cognition". In recent years, the work of Tomasello and his collaborators, see for instance [75, 76], has brought a new aspect into the discussion, the natural cooperativity of humans. As comparative experiments with apes and human infants demonstrate, a human child is naturally inclined to help another person when he or she understands what that other person wants, but does not succeed, to achieve. Apes, in contrast, behave in a much more selfish manner. Also, humans freely and spontaneously share information with others, not only through the medium of language, but also by the more elementary means of pointing [22, 70, 17, 18, 6]. Again, apes do not spontaneously point, and in turn have difficulties in grasping the helpful intention behind a pointing gesture [56, 74, 12].

It is therefore a crucial question to understand the mental mechanisms behind such human cooperativity, subconscious as they may typically be. Here, we want to utilize (and expand) tools from game theory and epistemic logic to analyze this issue. This will, in particular, enable us to identify the stages leading to this human achievement. In our scheme, these stages will be

1. Simple optimization
2. Anticipation of the reaction of others
3. Reflexive reasoning
4. Cooperative reasoning

Here, subsequent stages do not represent alternatives to, but rather refinements of earlier ones. In particular, reflexive reasoning not only anticipates the reactions of others, but also takes their reflexive reasoning into account. And cooperative reasoning is a special form of reflexive reasoning.

Going through those stages, we can also appreciate in which way human cooperativity is different from forms of cooperation among animals, be they organized as ant colonies or packs of hunting dogs.

In order to describe and analyze these stages, we shall draw upon rationality concepts as developed in game theory. We need, however, to critically examine these concepts and identify their shortcomings in our context, in order to reach our purpose. This program might immediately lead to the objection that no apes, as far as we can tell, and only relatively few humans grasp the concept of, say, a Nash equilibrium of game theory, let alone the more refined and involved rationality concepts proposed and discussed in the game theoretical literature.[1] So, how could they possibly behave according to such concepts?

Well, in a certain sense, game theory, when applied to biology, is an "as if" theory. This means that equilibria can be explained as if they were achieved by rational reasoning, but in reality, completely different underlying processes can lead to the same outcome as rational reasoning. This is precisely what has made game theory so popular in theoretical biology. One of the most prominent examples of biological versions of game theory is the concept of an evolutionarily stable equilibrium, that is, an interaction rule inside a population that is resistent against invasion by a small number of deviating mutants. Another example are the signalling games that aim at explaining how signals inside or between species can acquire their meaning. In such applications of game theory to biology, the pay-off as the objective function is given by a notion of fitness, essentially meaning the expected number of progeny.[2] Rational reasoning is replaced by evolutionary mechanisms. Suppressing certain nuances, like the difference between a Nash equilibrium and an evolutionarily stable strategy (the latter is a refinement of the former, and it is concerned not with individual players, but with repeated interactions among randomly paired members of a population), an equilibrium can be maintained both by rational reasoning and by the forces of evolution. Even better,

[1] We shall describe the game theoretical concepts needed for our discussion in the Appendix.

[2] The concept of "fitness" is very subtle, but here, we do not enter into that issue.

evolutionary theory can even explain how an equilibrium can be reached from some non-equilibrium state.

However, and this is an important point, rational reasoning can lead to an equilibrium much faster than evolution. That is, it provides a short-cut. Now, as already said, even humans do not usually employ the techniques of mathematical game theory to identify an equilibrium, in particular not in those basic situations of social cooperation that we are interested in here. Nevertheless, humans arrive at cooperative solutions in real time, and not only in the course of many generations of trial and error. Thus, perhaps the trick of evolution consists in supplying us with some subconscious mechanism that leads to results that are, at least in certain situations, the same as those that would have been produced by rational reasoning.

However, this would still fall short of producing cooperative behavior. Therefore, we need to go beyond the above scheme of evolutionary game theory. In fact, even rational reasoning when applied in a straightforward manner leads to practically intractable problems. To explain those, let us return to the various stages listed above. Simple optimization can be easy. I just select that action that I believe to yield the highest pay-off in a given situation. There are two terms in the preceding sentence, "believe" and "given situation", that need some explanation.

"Belief" expresses a *subjective* probability, or if it is not quantified, possibility. Subjective and objective probabilities can be, and typically are, different, and my beliefs may be wrong, but I will not be aware of that. Thus, beliefs come with a feeling of subjective certainty. This difference or discrepancy then can be exploited by others, and this explains biological mechanisms like mimikry or other forms of deception. Nevertheless, as we shall argue in later sections, beliefs also play a constructive role in enabling cooperation.

"Given situation" means what economists refer to by the "cet.par." clause, that is, the assumption that my own action will not change the situation, and in particular, not cause a counter-reaction by others. In contrast, in stage 2, that is, anticipatory reasoning, such a possibility is taking into account. When I am in the situation playing a game (whatever that means in a concrete context) with some opponent, I should anticipate his reaction to my action. In particular, his reaction can prevent the desired effect of my action, because that effect might be good for me, but harmful to the opponent, and through his reaction, he can in turn achieve an outcome that is better for himself, but worse for me. Thus, my choice of action should already take his possible reaction into account. Importantly, this is based on an assumption. The assumption is that the opponent also wants to maximize some pay-off, in the same way I do, although his pay-off will typically differ from mine. My gain might be his loss, and conversely. Now, when I anticipate his reaction, I could choose an action that is a reaction to his reaction. So far so good. But this assumes that I am cleverer than he is, because at this stage, I do not yet grant him the capability of also anticipating my reaction. Of course, I can remedy this incompleteness of my reasoning and assume that he can also foresee my reaction. But then, I should carry my reasoning to the next level, by in turn taking this into account in my own choice of action. And so on. The process can be iterated to arbitrarily high levels. That is what rationality would ultimately require. Of course, no human and no computer is capable of such an infinite iteration, and people can hardly go beyond the second or perhaps third level. In any case, the preceding analysis depends on a crucial assumption, common belief in or common knowledge of rationality (we shall explain these concepts in precise technical terms below). That means that I assume that my opponent is as rational as I am, and that includes that I assume that he also assumes me to be that rational, and so on. If instead, he were completely erratic at some level, I could perhaps not deduce much about what my optimal action should be. In that sense, the common belief in rationality helps me, and likewise my opponent, in our reasoning chain, because it means that many potential actions need not be taken into account in our mutual reasoning, whenever they would lead to outcomes inferior to those of other actions at some level. Moreover, such a common belief is self-confirming, in the sense that whenever everybody so believes, then everybody is also best advised to act correspondingly, indeed. And thus, being rational, he will.

In abstract terms, the common belief in rationality leads to a reduction of the complexity of the situation of the game, because then only rational (or, more precisely, rationalizable, a term to be explained below) actions of others, and not all possible ones, need to be considered in the reasoning chain. And, even more importantly, as we shall analyze in detail below, that common belief can replace an iterative type of reasoning – which would involve the difficulties just pointed out – by a recursive scheme, a finite shortcut of the infinite iterative chain. Or putting it differently, a scheme with infinitely many iteration steps may be simpler than one with finitely many iterations, because in the infinite scheme, all steps are the same, whereas in the inite scheme, each is different. Therefore, for the infinite scheme, it suffices to consider a single step, as we shall elaborate below.

Now, this can be taken further. In fact, this will lead to one aspect (but not the only one) of our analysis of cooperation. In addition to the common belief in rationality, there might be other such self-fulfilling common beliefs. In particular, when everybody believes that everybody is cooperating, then cooperation may result. The difference, however, between rationality and cooperation is, however, that usually,[3] nobody can benefit from not being rational, but in many cases, non-cooperation might be better for an individual, as long as the others still cooperate. This effect is known in the literature by names like the "tragedy of the commons". The game called the "prisoner's dilemma" is the simplest instance where this can be formally analyzed, and there are numerous papers about this game and its implications. Of course, when some players cease to cooperate, then the remaining players might follow suit, and the result will be inferior for everybody involved. The crucial point, as revealed by the prisoner's dilemma, is that when one player does not cooperate, then it is better for the other one to also not cooperate. The result is worse for both than if they both cooperated, but not as bad as it would be for a cooperating player with a non-cooperating opponent. Now, in a single shot game, the most basic paradigm of game theory, there is not much one can do about this. Even when the game is repeated a finite number of times, then some kind of backward induction tells us that in the last round, it is best to not cooperate. Given that perspective, then in the second-to-last round, one should better not cooperate either, and so on by iteration.

This is, of course, a gloomy conclusion, and granting for the moment the assumption that the prisoner's dilemma game really captures the pay-off structure underlying many human

[3]There are many exceptions, some of them discussed extensively in the game theoretical literature, and some of them relevant also for our analysis, but for simplicity of this introductory discussion, we suppress that issue here.

interactions, the question why humans nevertheless cooperate so often becomes even more perplexing. Now, however, human interactions typically do not take place in isolation, and human interaction is also in some sense open-ended, so that the assumptions underlying the preceding analysis might not hold. Of course, in the modern world, there are many one-time, single-shot interactions, but as evolutionary anthropologists argue, in most of human evolutionary history, cooperative interactions did take place in small and presumably rather stable groups, and cooperation inside such groups might have been enforced through the pressure of violently antagonistic encounters with rivalling or competing other groups. Groups consisting of, or more precisely, including some non-cooperators, according to this story, would have been wiped out by groups that were superior because of their strong internal cooperation. Thus, so goes the claim, cooperativity is built into our genes as a result of the evolutionary pressures acting upon our hunter-gatherer ancestors (see e.g. [81] for the conceptual scheme of evolutionary psychology). Be that as it may – and we don't want enter here into the intensively debated issue of group selection –, our concern here is more with the mental underpinnings of cooperative human behavior. Our question is which rational considerations (taken in the above "as if" sense) can make it advantageous to cooperate and, more precisely, to resist the temptation to stop cooperation for one's own short lived advantage. In fact, we are not simply concerned with such stability of cooperative behavior, but rather with the logical analysis of the stages leading to it. Also, in this paper, we do not explore the social mechanisms, like reputation, status, delayed reciprocity, and so on (see for instance [26, 38]), that may serve to stimulate and maintain cooperative behavior.

There is another principle underlying our discussion: *Simpler solutions are better*, and therefore, when we want to infer some mechanism underlying a cognitive or social phenomenon, we should look for the simplest solution. There is a twist, however, to this kind of Occam's razor applied to cognitive or social processes. Often, a simple solution is not so easily found, or better, depends on certain prerequisites that themselves can be complex. And it is our aim to identify those prerequisites. To do so, we shall critically examine the tools of game theory and also utilize concepts from epistemic logic.

2 Decision making and learning

Let us start with stage 1, that of simple optimization. An individual, let's say me for concreteness (as we are talking about an individual perspective), can choose between several actions each of which yields some result that possibly depends on certain contingencies of the situation. Those results can be quantitatively evaluated, and therefore compared. In game theory, they are called "pay-offs". I want to maximize my pay-off, and when I am rational, I shall discard any action that will always lead to a lower pay-off than another action. In technical terms, I discard any strongly dominated action. The pay-offs might not be certain, but when I know the probabilities with which they occur for any possible action, this kind of reasoning still applies. I simply have to work with expected pay-offs. When even the probabilities of the pay-offs for the various actions are unknown – an economist would speak of Knightian uncertainty here –, then of course there no longer is a basis for such reasoning, and other criteria, like risk aversion or so, should take over. But we shall not explore that here, and assume that at least the probabilities of the pay-offs are known.

But how do I know the pay-offs or their probabilities? From a biological perspective, this might not be an appropriate question. What counts in the end is only that I take the right decision, that with the highest (expected) pay-off. And that could come about by

1. *Evolution*: I have inherited the genes that dictate my choice of decision in some given situation from ancestors who, by choosing that action in such a situation, had achieved higher pay-offs than others that had chosen a different action.

2. *Learning*: I have tried out various actions and have gained experience about the resulting pay-offs. Like evolution, such individual learning works by trial-and-error.

3. *Cultural transmission*: I have been told by others what the best action is. Thus, I can learn from the experience of others.

4. *Anticipation*: I foresee the consequences of my actions.

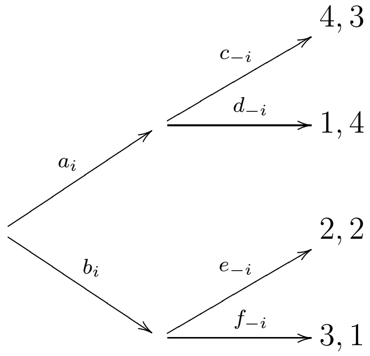
In any case, these are different causal mechanisms, and we interested here in the functional rather than the causal aspects. To see what is involved, let us look at learning (the *functional* aspects will also apply, with certain modifications, to the other items). For our current purposes, we can define *learning as the transformation of correlations into associations.* That is, when I repeatedly see that A is followed by B (a correlation), I make the association that whenever I see A, then B will follow. This is, for instance, the basic mechanism underlying Pavlovian conditioning.

All this is, of course, pretty simple, and so let us see how it fares when we move on to stage 2, that is

3 Playing against others

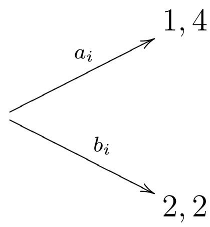
Even in the simple optimization situation just described, the player might have faced an opponent, but might not have been aware of that, and had simply assumed some pay-off as a consequence of an own action. That pay-off might have arisen from a reaction of the opponent, but when the opponent reacts in some fixed manner to each own action, the pay-off of that action is determined, and the player could simply proceed on that basis. In that sense, an equilibrium between two players can be reached without either of them being aware of this. When evolution or individual learning is involved on both sides, that equilibrium might only emerge in the course of the process during which the actions and reactions and therefore also the pay-offs change. At a Nash type equilibrium, however, no player can unilaterally deviate without a decrease in pay-off.

Here, we wish to analyze the case where a player can anticipate the opponents' reactions and select the own action accordingly. Since we are changing the perspective here, we now call our players i, referred to by the pronoun "she", and $-i$, referred to as "he". Pairs of numbers like $(5, 3)$ will indicate the pay-offs of i, 5 in this example, and $-i$, 3, for a given pair of action combinations. Let us analyze the possible reasoning structure at a simple example,



(1)

Here, i acts first and can choose either a_i or b_i. When i acts by a_i, $-i$ can respond by c_{-i} or d_{-i}, whereas to b_i, he can respond with e_{-i} or f_{-i}. The resulting pairs of pay-offs are indicated at the leaves of the tree. Now, when i can anticipate the reaction of $-i$, she can reason that when she chooses a_i, $-i$ will respond with d_{-i} because that gives the pay-off of 4 which is better than 3 which would result from c_{-i}. Similarly, he will respond to b_i with e_{-i}. Therefore, i can work with the reduced tree



(2)

and she will therefore play b_i as this results in a pay-off of 2 for her, whereas a_i would only give her 1. We note that the combination (a_i, c_{-i}) would have resulted in a better pay-off, $(4, 3)$, than the chosen combination (b_i, e_{-i}) which yields $(2, 2)$.

So far, from i's perspective, this does not require that she ascribes any intentionality or rational reasoning to $-i$. She simply anticipates his reactions to her own actions and then acts accordingly. Of course, the reduced tree (2) is as in an individual decision situation, but the point is that it has been arrived at here by a reasoning process taking into account the reactions of an opponent.

Through that reasoning, the original tree (1) is reduced to (2). First, i adopts $-i$'s perspective and eliminates his strongly dominated reactions from consideration. For instance, when she had played a_i, then c_{-i} would yield a lower pay-off for him than d_{-i}. Thus, c_{-i} is strongly dominated by d_{-i}. When she then has arrived at the reduced tree (2), she sees that her action a_i is strongly dominated by b_i, and therefore eliminated again. That is, what remains in the end is

$$\xrightarrow{b_i} 2 \quad . \tag{3}$$

Thus, in the end, she only has to play b_i. The iterated elimination of strongly dominated actions leads to a considerable reduction of the complexity of the situation. Rational reasoning can prune the tree, from (1) to (3).

Therefore, whenever she encounters the game (1), she simply has to play b_i. While the chain of reasoning reducing (1) to (3) was needed for the logical analysis, in the end, i no longer needs any of that. That is our conclusion for the moment, but later on, we shall return to this game (1) and discuss under which conditions the action pair (a_i, c_{-i}) can be realized which would yield a higher pay-off for both players.

4 Iterative reasoning

4.1 Equilibria in games

The above game (1) was easy in the sense that i acted first and then $-i$ reacted, and knowing $-i$'s pay-off, i could anticipate $-i$'s reaction and select her own action accordingly. The situation was unsymmetric between i and $-i$. But this is not the paradigm of game theory, nor the only possible scenario for encounters between agents, be they human or animal. We rather should look at symmetric situations where each player can react to the other. In game theory, such a symmetry is introduced when the players act simultaneously, knowing both their own options and those of the opponent. In that case, the preceding reasoning is inadequate because then not only i should think about the possible actions of $-i$, but $-i$ in turn should also contemplate i's action, and, this is the important point, i should include this fact in her considerations, and so should $-i$ by symmetry. And this should then be iterated. Of course, the analysis of this symmetry is the basic point of game theory and leads to the concept of a Nash equilibrium. Such an equilibrium is characterized by the condition that neither player can deviate from her equilibrium action without triggering a reaction of the opponent that is disadvantageous for her. This is, of course, well known stuff, but for our subsequent purposes, we are less interested in equilibria than in the conditions and mental processes leading to them. So, let us consider a concrete game,

	a_{-i}	b_{-i}	c_{-i}
a_i	3,1	1,2	4,5
b_i	2,3	2,5	5,4
c_i	1,1	3,2	2,1

. (4)

This game can easily be solved by the iterated elimination of strongly dominated strategies. First, whatever i does, a_{-i} always leads to a worse result than b_{-i} against the same action. So, $-i$ should not contemplate a_{-i}, and i knows that. Against the remaining actions of $-i$, then i will always do better by playing b_i instead of a_i, and so, a_i is eliminated as well. But then, against the remaining two actions for i, that is, b_i and c_i, $-i$ will always do better by playing b_{-i} than c_{-i}, and so, b_{-i} is the only remaining choice for $-i$. Against b_{-i}, however, i should best respond with c_i. Thus, (c_i, b_{-i}) is the only equilibrium, leading to the pay-off pair $(3, 2)$. This is worse for either player than what they would get at either of the non-equilibria (a_i, c_{-i}) or (b_i, c_{-i}). There are several points to make here:

1. The solution of this game requires more than simple anticipation of the opponent's reaction. For instance, i not only anticipates that $-i$ will not play a_{-i}, but also that $-i$ in turn knows that and therefore expects i not to play a_i, and so on. Here, the iteration terminates after a finite number of steps, but we shall see at the next example that this need not be so.

2. The above analysis assumes that the players are both rational in the sense that they apply logically consistent reasoning. A simpler type of reasoning could lead to a better result in this game. For instance, i could simply ask how $-i$ would react to each of her actions. Against a_i, he would react with c_{-i}, but against either b_i or c_i, he would react with b_{-i}. In the first case, i would receive the pay-off 4, whereas in the other cases she would only get 2 or 3. Thus, playing the action a_i that had been eliminated by the above dominance argument would actually get her a better pay-off than the equilibrium strategy c_i. The only reason why she would not play a_i is her own temptation to play b_i rather than a_i against c_{-i}, which, however, would induce $-i$ by the same reasoning to switch to b_{-i} etc. If $-i$ argued from the same perspective, that is, choosing that action of his for which the response of i would be most favorable for him, he would choose c_{-i} with i's response b_i.

3. Thus, the rationality of the players appears somewhat shortsighted, as they only consider how they could best react to an opponent's action, but not the entire chain of counterreactions. In that sense, the reasoning process just described is only consistent at the equilibrium, but not away from it.

4. The question then is by what kind of consistent, or at least self-confirming reasoning the players could achieve either (a_i, c_{-i}) or (b_i, c_{-i}), and if so, by what criterion to decide between these two possibilities, as the first one is better for $-i$, the second for i. Again, the problem is that when that player for whom such a combination is worse tries to get an improvement, then a chain of reactions will be triggered that will be worse for either of them.

We next consider the game

	a_{-i}	b_{-i}	c_{-i}
a_i	−10,1	10,0	10,10
b_i	−10,10	11,−10	10,0

. (5)

Here, for i, action a_i is weakly dominated by action b_i, that is for every action of $-i$, b_i yields either the same or a higher pay-off than a_i. Thus, i might discard a_i and only play b_i. This, however, seems to be a bad idea, because against b_i, $-i$ would play a_{-i} which would yield a negative pay-off, -10, for i. In fact, $-i$ would never play b_{-i} anyway, because it is dominated by both a_{-i} and c_{-i}. In all remaining cases, however, a_i is not worse than b_i, and therefore, it can no longer be eliminated by a dominance argument. Thus, when i, before her own elimination of strategies, had contemplated $-i$'s situation, they would end up at the equilibrium (a_i, c_{-i}) where each receives the pay-off 10. The point here is that the situation on which i's dominance argument is based, $-i$ playing b_{-i} will actually never occur. Reasoning should not be based on such purely hypothetical scenarios.

In fact, a similar phenomenon even emerges for the elimination of strongly dominated strategies, as the following slightly modified game indicates

	a_{-i}	b_{-i}	c_{-i}
a_i	−10,1	10,0	10,10
b_i	−9,20	11,−10	11,0

. (6)

Here, a_i is strongly dominated by b_i, but again eliminating a_i and playing b_i is not a good idea for i because in that case $-i$ would be all too happy to respond by a_{-i}. Had i correctly anticipated the reactions of $-i$, she would have played a_i with $-i$'s response c_{-i}, although that is not a Nash equilibrium. This example demonstrates a limitation of the concept of a Nash equilibrium to which we shall return below.

Our next example is

	a_{-i}	b_{-i}	c_{-i}
a_i	4,0	0,4	0,0
b_i	0,4	4,0	0,0
c_i	0,0	0,0	2,2

. (7)

Here, there are no dominated actions. Nevertheless, there are Nash equilibria. Obviously, (c_i, c_{-i}) is one. Furthermore, when both players play either a or b with probability .5, we have a mixed Nash equilibrium, in which the expected pay-off for either player is $(2, 2)$ which happens to be the same as the certain pay-off at (c_i, c_{-i}). We could then define an action d for each player in a reduced game that would correspond to playing that mixed strategy in the original game. The reduced game would thus be

	d_{-i}	c_{-i}
d_i	2,2	0,0
c_i	0,0	2,2

. (8)

This game would then have two pure Nash equilibria, (d_i, d_{-i}) and (c_i, c_{-i}), both with pay-offs $(2, 2)$, and in addition a mixed one, each playing d and c with probability .5, with a pay-off of $(1, 1)$ only.

For a mixed Nash equilibrium, when the game is actually played, each player in the end has to relinquish her randomness and choose a definite action. The mixed equilibrium concept does not tell her which. It is here that the concept of rationalizability enters which we now briefly describe as it is relevant for our analysis. In the game (7), i might play a_i because she believes that $-i$ played a_{-i} because he believed that i played b_i because she believed that $-i$ played b_{-i} because he believed that i played a_i. This could go on indefinitely, but in this example, at this level of reasoning, the period is completed, as we had started with i putatively playing a_i. So, the reasoning suggesting to i to play a_i seems consistent, or to use the right technical term, rationalizable. But so, in this example, would in fact any possible action of either player be. This, of course, leads to a state of complete indeterminacy. As any action of either player is rationalizable, any action combination could result, for instance (b_i, a_{-i}) which would justify $-i$'s chain of reasoning, or even (c_i, b_{-i}) which would disappoint either player. The equilibrium (c_i, c_{-i}) is different as this is constellation that would be simultaneously expected by either player when they chose to play it. That is, it represents a state of correct mutual anticipation of both players.

The indeterminacy can occur even in the simple game (8). Player i might play d_i because she expects $-i$ to play d_{-i} because she thinks that he expects her to play d_i because she thinks that he will play d_{-i} and so on. In turn, $-i$ might play c_{-i} on the basis of an analogous reasoning for the c actions. Then, both players will be surprised to get the pay-off 0 when the combination (d_i, c_{-i}) is played that had been expected by neither of them. The self-consistent iterative reasoning of the two players does not survive the reality check.

How can such a failure of coordination be avoided? Of course, the players could learn, either in the evolutionary sense or in the stricter sense of individual experience. But there still is a problem. For instance, when the game is iterated and in the first round (d_i, c_{-i}) is played, then i could reason that it would be better for her to switch to c_i whereas $-i$ for the same reason would decide to switch to d_{-i}. The resulting action combination (c_i, d_{-i}) then yields the unexpected pay-off of 0. In the next round, both players could switch again, or alternatively both decide that switching is no good and stay with their actions. In either case, no coordination will be achieved. This could go on forever. Of course, in this simple game, the solution is easy, to break the symmetry between the players, or to reason from a game theoretical perspective that the mixed Nash equilibrium in (8) is unstable. In particular, in an evolutionary setting, within a population a mutant could arise that matches the action of the opponents and then by its superior fitness spread in the population. Our concern, however, is more abstract and general. We want to investigate logical mechanisms that are making assumptions that when made by all players will thereby become true.

4.2 Looking further ahead

Before moving on to the more abstract aspects, we want to analyze the game theoretical aspects once more. In fact, in several of the games that we have analyzed, in (1), (4) and (6), we have seen that the Nash equilibrium yields pay-offs for the two players that are worse for both of them than in some other, non-equilibrium constellation. It is then a natural question by what kinds of mechanisms, or more precisely in our framework, logical reasonings, such a constellation can be stabilized.

Of course, for a single-shot game, the logic behind the Nash equilibrium is impeccable. The reason is that when one player deviates from a favorable non-equilibrium constellation in order to gain a whatever slight advantage, he does not have to worry about his opponent's reaction as the two players can only choose their action once, and then the game is over. Thus, there is no room for a logic like "If I did that, my opponent would respond thus, and since this would be bad for me, I'd better not do that in the first place". But such a reasoning would precisely stabilize the constellation (a_i, c_{-i}) in (4) or (6). That is also the reason why the tit-for-tat strategy in the iterated prisoner's dilemma game

	a_{-i}	b_{-i}
a_i	1,1	−2,2
b_i	2,−2	−1,−1

. (9)

turns out to be superior in large scale computer simulations to any other strategy, see [2]. Here, "tit-for-tat" means that in the first round, the player plays a ("cooperate") and than in each subsequent round repeats what the opponent played in the previous round. Therefore, whenever i plays b_i ("defect") against a tit-for-tat player $-i$, she knows that this will trigger the action b_{-i} in the next round. Anticipating that, she might rather decide to cooperate. The tit-for-tat strategy is usually interpreted in terms of retaliation and punishment. In contrast to that, we here emphasize the logical aspect that in the iterated prisoner's dilemma, a rational player should foresee the consequences of her actions in terms of the opponent's reaction.

An analogous strategy and reasoning could also be developed for (4) and (6).

5 Animal reasoning and signalling – a logical analysis

Before moving on to the next stages, that is, reflexive and cooperative reasoning, we want to discuss how some findings about animal reasoning and signalling fit into our logical scheme.

5.1 Alarm calls

Many social animals, like monkeys or fowl, have developed signalling behavior to alert some or all group members to the presence of a food source (see for instance [68]) or to indicate the presence of some danger. The best studied cases here are the systems of alarm calls that warn other group members when some predator is detected, see for instance [61, 63] for summaries and discussions of findings. These alarm calls can be predator specific [73, 60] and may also depend, for instance, on the distance of the predator [84, 85, 14]. One might then conclude that the animals in question possess some representation for each specific type of predator, and that the corresponding alarm call then evokes that representation. However, this need not be so, and alarm calls can also fulfill their purpose by simply triggering the escape behavior that is most appropriate under the circumstances. Let us therefore list different aspects of such signalling more systematically. An alarm call could express, or more precisely, be interpreted as

1. the intensity of a threat – louder alarms might signify a more imminent danger
2. a functional reference – suggest what to do, that is, how and where to seek cover, and distinguish between the different such actions required for different types of threats
3. a causal inference – for instance, from the alarm call of a different species, an animal could conclude that there is a specific danger
4. information about the presence of a predator, and more precisely, about its type and distance
5. the emotional state of the alarm caller,

and in fact, these different items do not necessarily exclude each other, but could operate simultaneously. For instance, 1. and 5. could well go together. The more imminent the danger, the higher the emotion of an animal detecting it. Less trivially, the emotional state might be a cognitive shortcut, in the sense that for an animal hearing the alarm call, the correlation between the danger and the intensity of the call expressing the emotional state of the caller has been turned into the association that an excited alarm caller indicates an imminent threat. In particular, even though the signaler may simply express its emotional state and may not "intend" to convey any information, it may still provide information to the listener, see [61]. On the other hand, there is the audience effect, that is, typically alarm calls are only emitted in the presence of kin, group members or potential mating partners. Roosters of the species *Gallus gallus* can emit alarm calls in order to protect hens – potential mating partners – from aerial predators, but also to direct the attention of a predator to another rooster – a potential rival [41]. Also, an alarm call can vary in different dimensions; for instance, its type could inform about the nature of the danger, and its loudness could indicate its degree. Alarm callers, however, do not seem to be able to assess the receptivity of a possible receiver. It seems that the presence of a conspecific or group member simply is a trigger that provokes an alarm call when a predator is approaching. Thus, if we were to interpret alarm calls as expressions of the caller's emotion, we would need to concede at least that that emotion is not only caused by a perceived danger for oneself, but in addition modulated by the presence of others or perhaps other contextual cues. .

However, there is a certain tension between 2. on one hand and 3. and 4. on the other hand. In fact, there has been a lot of debate in the ethological literature whether an alarm call tells what to do or informs about some state of the world.

The most parsimonious interpretation is based on the following simple observation. For, say a vervet monkey (*Cercopithecus aethiops*) [73, 60, 16], that is, a species preyed upon by leopards, seeing a leopard, hearing a leopard's growl, receiving the leopard alarm call of a conspecific or that of a member of a different species are all functionally equivalent, in the sense that they indicate some danger, from a leopard in this example, and suggest a specific strategy to avoid predation, for example climbing into the thin branches of some tree where the leopard cannot follow. The same applies to eagle or snake alarms. The corresponding signals are different, and different escape strategies are used by the monkeys. In fact, diana monkeys treat conspecific leopard alarms and the growl of a leopard as similar to each other, but different from groups of signals that indicate a different type of predator, like conspecific eagle alarms and the shriek of an eagle [83]. Thus, either of these indicators of the presence of a leopard is a valid and important sensory signal. And this can be learned by a simple association schemes, as described above. It does not matter for our purposes whether this takes place on an evolutionary or individual scale or by teaching of parents or other senior group members. And, in fact, the association need not even take place between the leopard and the alarm call. The young monkey could just notice that an alarm call is always followed by a particular escape strategy, and it could simply always participate in such a group escape, and then associate the alarm call with the escape behavior. Clearly, this is simpler and requires much fewer logical prerequisites than a Shannon [62] type interpretation of an animal decoding a message. We refer to [61, 63] for further discussions in this direction.

However, there do exist certain examples in the animal behavior literature that might require a somewhat more complex logical analysis than the preceding.

Diana monkeys (*Cercopithecus diana*) which are preyed upon by leopards (*Panthera pardus*) as well as chimpanzees (*Pan troglodytes*) and humans make an alarm call in the presence of a leopard but try to remain silent and cryptic in the presence of a chimpanzee or human. But since chimpanzees also emit leopard alarm calls, the monkeys have also learned to react to the latter appropriately. Moreoever, they are sympatric with guinea fowl that produce the same alarm call to leopards and humans. The Diana monkeys not only understand the alarm call of guinea fowl, but also use circumstantial evidence to decide how to react to it, that is, whether to produce an alarm call themselves or try to silently hide. That is, they make some kind of causal inference whether the guinea fowl alarm call which does not discriminate between the types of predators had been solicited by a leopard or a human. See also [57] for such an ability to distinguish between different alarm calls in birds.

5.2 Hidden food

We now move on to chimpanzees (*Pan troglodytes*). In contrast to monkeys, they may possess some, whatever rudimentary, understanding of the intentions and mental states of others. Here, we do not wish to enter that discussion, but simply treat one example (see e.g. [32] and references therein). An ape i knows where food is hidden, but there is another (dominant) ape or a human $-i$ present who does not see the food or knows its location. If i went for the food, however, then $-i$ would see this, and i would expect him to snatch the food from her and perhaps even beat her up. Thus, i needs to find some way to get the food without $-i$ noticing it. A game theoretical analysis might invoke the concept of unawareness (cf. [25], [34]), as $-i$ does not know that there might be food. i would then plan her action on that assumption of $-i$'s unawareness.

Of course, the behavior of i can be explained by anticipatory reasoning. She simply expects that when she openly approached the food, $-i$ would follow her and grab the food. Therefore, she rather chooses a hidden action, and if this should not be available, simply gives up the food. For that, she does not need to entertain any higher order beliefs about $-i$'s state of mind. i's reasoning is simply based on her observation that $-i$ neither sees the food nor approaches it. Whether the actual reasoning of chimpanzees is as simple as that, of course, we cannot really tell. But it provides at least the simplest explanation for some observed behavior.

5.3 Group hunting

We next discuss an example that involves some coordination within a group, the hunting of monkeys by wild chimpanzees in Côte d'Ivoire (cf. [8] and [9]). Occasionally, varying numbers of apes hunt (single) monkeys, and in such a hunt proceed in a coordinated manner in order to block the victim's different escape routes. During the hunt, different hunters perform different 'roles' according to the requirements of the circumstances. This interplay of roles might suggest that the apes have some kind of common understanding of a joint enterprise. This would then imply that the hunters entertain at least 2nd-order beliefs. Every participant would then believe that the others, just as himself, are aware that "they are hunting together as a group". It is however highly disputed whether group hunting of chimpanzees is based on such a deliberate cooperation or whether each hunter simply acts upon an individual decision about his best action. Not being primatologists, we cannot solve this issue here. We offer an analysis of the situation from the perspective of an individual ape which is based on a combination of simple optimization and some anticipation of the behavior of both the monkey and the fellow hunters.

In order to catch the monkey, the ape has to anticipate its fleeing behavior, i.e. in which direction it will most likely run. This is what individual and group hunting have in common.

In group hunting, however, the fleeing direction of the monkey is influenced by the positions and actions of the other hunters. Therefore, in order to anticipate what the monkey will do, an ape should also observe what the other hunters are doing and anticipate what they are about to do. Whenever an anticipating reasoner takes a position within the hunt, he maximizes his chances of catching the monkey given the anticipated actions of his fellows as well as the anticipated reactions of the monkey to those. As a result, the apes mutually take positions that together block the best escape routes of the monkey. When all hunters proceed in this manner, the sum of their individual behaviors comes to resemble an organized maneuver, i.e. a combination of complementary actions that result e.g. in an encirclement or an ambush. Thereby the chances increase that the group, or more precisely, some members of the group will catch the monkey (though they are not necessarily maximized, as a coordinated plan of actions between the hunters may be superior). Our point is that an ape benefits from anticipating the interactions between different 3rd parties, i.e. the monkey and the fellow hunters. This is more complex, but principally not different from the principles which an ape would employ in a lone hunt. This illustrates how reasoning on the fellow hunters combined with person reasoning on the monkey provides a chimpanzee with information that serves as premises for his optimization reasoning. This combination is admittedly complex, but it does not (have to) involve any higher order beliefs.

Perhaps, after all, this is not so different to what goes on in human group activities like soccer matches where also the team members need to cooperate to achieve their goal(s). Again, in a given situation, a player needs to seek his role by anticipating the trajectory of the ball, the actions of the opposing team and perhaps of his own team mates. While abstractly, one may of course reason about an underlying common belief here, this may not be so relevant after all when the game is actually played. What is important is to find the right action as quickly as possible, and for that purpose any shortcut of the reasoning chain is valuable. In particular, it is the purpose of systematic training to install certain routines in the players by which they can avoid more complex reasoning chains.

6 Reflexive and recursive reasoning

6.1 Rationality and reflexivity

First, we should clarify that for us, rationality means the ability for logically consistent reasoning, as distinct from the optimization of one's individual pay-off. As we have seen, logically consistent reasoning does not necessarily lead to pay-off maximization. On the other hand, in a game like (4), it could also be better to forego direct pay-off optimization and thereby ultimately achieve a higher pay-off. We shall first analyze the reflexive nature of rationality, corresponding to stage 3 of the introduction, and in Section 7, we shall then move on to the cooperative aspect, that is, stage 4.

First, one can only rationally reason about another player when that player is rational as well. The question to be answered by rational reasoning is "What would the other player rationally think and do in that particular situation?". Now, of course, the rational thinking about the rational opponent also needs to include his rational thinking about my rational reasoning. In short, *rational reasoning about the reasoning of others implies the assumption* *that those others are rational in the same manner.* In particular, the situation between rational reasoners is symmetric. That is, *it must be common belief among rational individuals reasoning about each other that everybody is rational.* The important concept of *common belief* emerging here will be discussed in the next section. But first, we need to clarify another aspect. When analyzing the games (4), (7), (8), we have utilized several solution concepts, that of a Nash equilibrium and that of rationalizability. In particular, we have found that a rationalizable solution need not be a Nash equilibrium. But what would happen if the players themselves adopted different rationality concepts? Then when each player is consistent as required by rationality and thereby attributes the same rationality concept to his opponent, he might reason differently about the opponent than that opponent might reason herself. Or to put it the other way around, reflexive reasoning implies the more specific *common assumption that all players reason with the same specific rationality concept*, whatever that concept might be, as long as it satisfies the rationality criterion. Such a common choice is therefore necessary for a social alignment of decision making.

6.2 Common belief

The reflexive reasoning as just analyzed is interdependent. It represents a collective or common perspective as viewed by an individual actor. This kind of reasoning was informally pointed out in [52] and is similar to Searle's ([59]) notion of 'collective intentionality', to the 'we-intention' described by Tuomela ([79]), and is a direct descendent of Lewis' ([43]) concept of 'common belief' underlying a 'convention'. In game theory 'common belief of rationality' is a recurrent topic that, in addition to the traditional Nash-equilibrium concept, is also linked to concepts such as 'rationalizability' ([7]). In e.g. institutional economics and in public goods theory, some problematic implications which arise from lack of a cooperative type of interdependent reasoning have been explored (e.g. [53], [54]). Along with cooperative social institutions, notions such as group or team reasoning have been analyzed with respect to their (in)consistency with rational utility maximization. Part of that literature combines Lewis' work with logical models and with new variants of game theoretical concepts (e.g. [72], [19]). Also, in theories of social interaction and social systems (e.g. [47]) concepts like 'expectations of expectations' have been developed which are supposed to describe alignment of social actors' reasoning processes, decisions, and actions.

The concept of common belief is now a standardized element of modal epistemic propositional logic, the logic of beliefs and knowledge (cf. [15], [21], [30]). Epistemic logic can include counterfactuals ([44], [30]), and it can be combined with dynamic logic (of actions, e.g. [20]). With a few additional concepts, e.g. preferences, an explicit formal representation of knowledge, beliefs, and actions of players in a game-like situation can be achieved (e.g. [71], [45], but there is no standard approach yet). Philosophical, economic and computer science approaches often parallel research in epistemic game theory ([1],[4]). In the literature, there are different suggestions as to how to formalize something like an institutional fact resp. the type of collectively aligned reasoning on which it is supposed to rest (e.g. [46], [72]). To our knowledge, there is however no standard concept so far.

We shall now formally connect the reflexive mode of reasoning that emerged in the last section with the logical concept of common belief established in a group of actors. We shall employ notations and concept from modal logic. $B_i p$ will stand for the assertion that the actor i believes that some proposition p is true and $B_E p$ means that all members of a given group E believe p, that is, $B_E p = \forall j : B_j p$. CBp then expresses the *common belief* among the group members that some proposition p is true. This means that not only every member i of E believes p ($B_E p$), but also that every i believes $B_E p$, i.e., $B_E B_E p$, that every i believes $B_E B_E p$, i.e., $B_E B_E B_E p$, ... etc. ad infinitum. That is,

$$CB_p = B_E p \wedge B_E B_E p \wedge B_E B_E B_E p \wedge \dots \tag{10}$$

This is more than a simple infinite iteration of ever higher order beliefs which, in itself, might give rise to both logical and psychological objection. Common belief is not just iterative, but *inherently recursive* (cf. e.g. [10] for recursion in logic). CBp is a proposition which makes the following logical equivalence true:

$$CBp \Leftrightarrow (\ B_E p \wedge B_E CBp\ ) \tag{11}$$

Thus, instead of the infinite iteration (10), (11) is a fixed point definition. More precisely, CBp is the greatest fixed point[4] of

$$\beta = B_E p \wedge B_E \beta. \tag{12}$$

An actor who has a concept of common belief may therefore be interpreted *to be capable of recursive thinking*. Understanding of the recurrent pattern amounts to understanding of the "etc." in (10), and thereby of the whole expression CBp. The infinitely many iterations in (10) then no longer need to be spelled out. Since the infinitely many iteration steps are all structurally identical, they can be condensed into the recursive form (12) which no longer involves an iteration.[5]

In line with the general structure of our argument, we do not propose any particular interpretation of how an individual might conceptualize common belief. The analysis in terms of modal logic that we develop need not be mirrored in actual human cognition. We only wish to isolate the logical structure underlying common belief at an abstract level.

(11) represents the outside perspective of an omniscient modeler. In order to develop the perspective of the actors involved, we define the reflexive belief of i as

$$RB_i p :\Leftrightarrow B_i p \wedge B_i(\forall j : RB_j p). \tag{13}$$

Then $\beta := \forall j : RB_j p$ satisfies the fixed point property (12). Thus, when the $RB_i p$ are the largest such expressions, $\forall j : RB_j p$ satisfies the definition of common belief CBp. (13) then becomes

$$RB_i p :\Leftrightarrow B_i p \wedge B_i(CBp). \tag{14}$$

This says that actor i does not know that something is common belief, but can only believe that it is. She believes that p holds and that p is common belief. Thus, the recursiveness of common belief leads to the reflexivity of the reasoning of i, that is, the assumption that her own reasoning is aligned with that of the others. Conversely, common belief results when

[4]For propositional μ-calculus, the theory in which the notion of the greatest fixed point is formalized and utilized, see [42].

[5]In epistemic game theory, this yields a basis for the theory of types, see for instance Chapter 11 in [49].

all members of the group E reason in that way. The reflexive reasoners thus know how reflexive reasoning works, what they could achieve through successful alignment, and what it takes to achieve alignment – and this is itself subject of their reflexive reasoning. Again, this alignment (13) of the beliefs of the members of E can only be believed by them, but not observed or verified in general. They could believe different things to be common belief, but then there would be no common belief. When there is common belief, everybody believes in it, but when everybody believes that there is common belief, then there need to be actual common belief, because everybody could believe that something different be common belief.

This also sheds light on the various game theoretical concepts that we have analyzed above as instances of rational reasoning, like rationalizability, Nash, etc. Only when a particular such concept is common belief among the players will there rational reasoning not only be internally consistent from the perspective of each individual player, but also consistent among the players.

7 Cooperation

We can now put the preceding considerations together to arrive at cooperation. According to our analysis, cooperation requires (at least) the following ingredients

1. The rational conclusion of each player that a cooperative solution is better for her or him than a non-cooperative one.

2. The rational insight of each player that the other players are as rational as her/himself.

3. A self-fulfilling coordination mechanism between the players.

Again, we emphasize that we are developing a logical analysis, which is distinct from any factual claim that cooperation as found, for instance, among infants, depended on conscious rational decisions. The fact is rather that cooperation when it occurs is more or less automatic. We ony wish to identify the structural substratum underlying cooperation. The subtleties of that substratum then might contribute to understanding why cooperation among animals is so rare and why a full-blown version seems to occur only among humans.

In a game like (4) or (13), it is clear that whenever the players are not engaging in an isolated single-shot game, a cooperative solution is better for each player than the Nash equilibrium. As analyzed, a cooperative solution can be stabilized by the insight that a deviation by either player would trigger a chain of reactions and counter-reactions that ultimately would leave everybody worse off. Whenever one player ceases to cooperate, rationality induces the opponent to switch to the non-cooperative default mode. And each player, so we assume, believes that of the opponent at every level of iteration, that is, there is common belief of rationality. As we have discussed, the switch from high-order or even infinite iteration to recursion, while logically more difficult, yields an enormous simplification of the structure.

When it comes to rationality, we might even speak here of common knowledge instead of common belief, as at least in the idealized situations discussed here, we assume that, at this stage, rationality is a prerequisite and not an outcome of the reasoning process.

We need somewhat more, however. Let us consider the following game

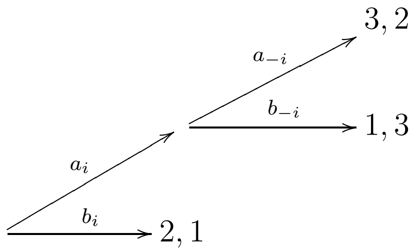
	a_{-i}	b_{-i}
a_i	6,4	0,0
b_i	0,0	3,5

. (15)

Here, there are two Nash equilibria, (a_i, a_{-i}) and (b_i, b_{-i}). The first is better for i, the second for $-i$, and so, the players need to agree on a selection criterion. An obvious choice would be to agree on (a_i, a_{-i}) because in that case, the sum of the two pay-offs is highest. When the players can agree on that, also the problem of rationalizable solutions would be avoided that comes up here as in game (8). Thus, a common belief that in cases of ambiguity, each player will go for the solution with the highest pay-off sum will coordinate the players on finding that solution and thereby will confirm itself.

Of course, the situation can be more complicated. In game (4), we had identified two cooperative solutions, (a_i, c_{-i}) and (b_i, c_{-i}), and it remains to be decided which of them is played. The first solution is better for $-i$, the second for i while the sum of the two pay-offs is the same. Here, a comparison with the Nash equilibrium might help. That Nash equilibrium yields the pay-off 3 for i, 2 for $-i$. One might therefore argue that (b_i, c_{-i}) with the pay-off structure $(5, 4)$ is fairer than (a_i, c_{-i}) with $(4, 5)$ because in the former each player gains the same amount, 2, when compared with Nash. Again, abstractly, a coordination criterion, in this case a notion of fairness, is needed, and when this is common belief among the players, the corresponding solution will be realized. Of course, in more complicated or less transparent scenarios, some metarule is required to select among possible coordination criteria. We do not explore this here, because, from a logical point of view, this is not principally different from coordination at lower levels.

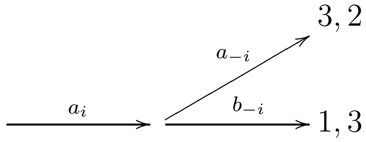
We now consider a game in sequential form,



(16)

Here, the cooperative solution would be (a_i, a_{-i}), but when i, the first mover, played a_i, $-i$ could respond with a_{-i} to secure himself the pay-off 3, leaving only 1 for i, and so in view of this, she might play b_i to achieve the pay-off 2, leaving only 1 for $-i$ in turn. This is the sandard backward induction. But how could the cooperative solution (a_i, a_{-i}) then be achieved? In the literature, see [5], this has been discussed under the label "guilt game". The reasoning here is that if i played a_i, thereby offering $-i$ generously a higher pay-off, $-i$ would feel guilty by responding to this generous gesture by b_{-i} and hence would rather play a_{-i}, the desired cooperative behavior. For our purposes, however, this would beg the solution because we want to identify the rational reasoning underlying such a notion of guilt. The same applies to, e.g., [79] where also the reasoning is based on the existence of what is called we-intentions.

Now, in our game, if i had not had the option b_i, but could only have played a_i, then $-i$ could simply decide to play b_{-i} to secure himself the higher pay-off,

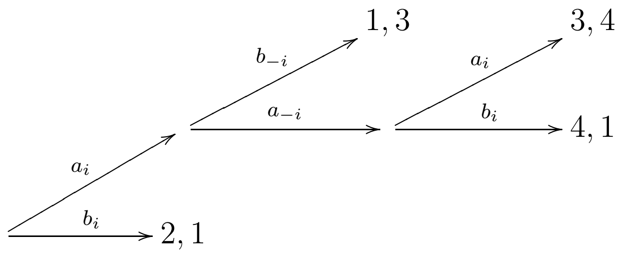


(17)

In (17), however, she did have that option, and if she did not realize it, but played a_i instead, then $-i$ should understand and interpret that as a rational gesture aiming at the pay-off pair $(3, 2)$ and inviting $-i$ to cooperate to achieve that. $-i$ can compare (a_i, a_{-i}) both with b_i and with (a_i, b_{-i}). Precisely that $-i$ understands that i did not exercise her option b_i in order to invite him to cooperate should then rationally induce that $-i$ cooperates, indeed. Similarly, in game (1), i might have played a_i instead of b_i, and $-i$ realizing that alternative, could then have responded with the cooperative c_{-i}, guaranteeing both a higher pay-off than at the Nash solution (b_i, e_{-i}).

This can be iterated, as in

(18)



Here, both players have an outside option, b_i or b_{-i} that at the corresponding stage of the game would secure them a higher pay-off. But when i did not play b_i, then for the same reason $-i$ should return that invitation for a cooperation and not play b_{-i} either, expecting that i, having offered a cooperation first, would then also cooperate at the next stage and play a_i once more. Understanding that the other player did not exercise an available non-cooperative action and at the same time accepting her/him as a rational actor should then induce each player to cooperate.

In any case, the crucial point here is that the existence of an outside option, playing b instead of a, that is not, but might have been realized, or could be exerted in future instances of the game, can induce cooperation. The action a thereby becomes a signal, since the opponent knows that an alternative action b would have been possible.

8 Final discussion

It is perhaps the most remarkable finding of modern experimental comparative research on human vs animal behaviour (see esp. [75]) that human infants are naturally cooperative whereas great apes are not. Human infants can make considerable efforts to provide others with useful information without necessarily expecting a benefit from such information sharing. They seem just to want to be helpful to others. Following a line of reasoning of Tomasello [75], this might be related to the ability to see themselves in others. It has been discussed extensively in the recent literature to what degree great apes possess a 'theory of mind', that is, the ability to understand that others are intentional beings like oneself and to interpret their actions from that perspective and also to infer their intentions and to predict their future actions. Tomasello goes on to argue "that human infants understand joint activity from a 'bird's-eye view' " ([75]: 179), that is, from an objective outside perspective seeing themselves on equal terms with the others involved in the activity. We note here the crucial difference between simply having common interests and the bird's eye perspective that sees the – possibly different – interests of all player. Such a view then, for instance, enables infants to switch roles in joint activities, which is something that apes cannot easily do, if at all. In this paper, we provide a formal analysis supporting such an interpretation.

The underlying reflexive reasoning that builds expectations about the behavior of others and even anticipates such mutual expectations and that then makes cooperation possible is not necessarily carried out consciously. As the findings for human infants show, it already emerges at some more elementary level.

The step from simple to cooperative reflexive reasoning seems small. In evolutionary terms, the ability for cooperative reasoning might have arisen as an unintended consequence of the reflexive reasoning ability. Here, the metaphorical expression "unintended" means that it had already evolved for a different function, i.e. for the non-cooperative reflexive reasoning whose evolutionary advantages have been described in Section 6, and not for the cooperative reasoning. In the terminology of Gould and Vrba [28], this would qualify as exaptation (see [27] for a systematic discussion of the relationship between structure and function from such a perspective).

Or it may be the other way around, that the scheme of rational reasoning – not necessarily consciously –, being of considerable evolutionary value in itself, may suppress direct pay-off maximizing behavior in cases where the two are in conflict, as in some of the cases analyzed in this paper. It may require too many cognitive resources, which would be better devoted to other goals, to always optimize the pay-off in every petty situation, instead of simply applying some more general device like rational cooperation without further thinking about it. This argument could be further developed in the direction of H.Simon's concept of bounded rationality [64, 65].

Cooperative behavior, once evolved, can then be stabilized and further enhanced by evolutionary mechanisms. Of course, groups with cooperators can gain a selective advantage over non-cooperative ones, but beyond this collective effect, cooperation can then also be stabilized within groups. This requires the ability to consistently identify individuals within a group, a cognitively very difficult,if not impossible, feat for almost all animals, and to consistently ascribe their cooperative or non-cooperative actions in past encounters. Thus, individuals can gain a reputation as cooperators, as argued in [38]. Being cooperative and thereby inducing others to cooperate can then become an advantage even within a group, and such a cooperator will then become a desirable mating partner and can thus pass on the inclination to cooperate to progeny. Or expressed more succinctly, a reputation for cooperating can be turned into a high social status within a group and thereby increase the mating chances, and by such mechanisms, cooperation will spread in the group. Whether the fact that such a group with many cooperators will then also successfully compete with others can then be seen as an unintended side effect which, however, in turn enhances the original effect.

We have also argued that the complexity of reasoning on one hand increases at each stage and enables individuals to cope with more and more complex situations, but that, on the other hand, each reasoning mode also reduces the complexity of the option configurations that have to be taken into account. Thus, a higher complexity can lead to more efficient reasoning and free cognitive resources and through this in turn enable the individual to cope with still more complex situations. This dialectical interplay between complexity reduction and generation, or between external and internal complexity is analyzed in [47, 37].

In the individual perspective of simple decision making, the complexity of the situation is reduced through the elimination of strongly dominated actions. In the already more complex perspective that includes anticipation of the reaction of others, such strongly dominated actions of others are eliminated as well. A logical consequence of such anticipatory reasoning is an iteration. The complexity of iteration, however, can be reduced through the device of recursion. Recursion is more abstract than iteration, but replaces a possibly infinite iteration by a single step. Recursion then naturally leads to reflexive reasoning. Unfolding the full potential of reflexivity by achieving a bird's eye view (in the words of Tomasello) then leads to cooperative reasoning. Again, this represents a higher level of complexity that makes short-cuts possible that achieve considerable simplifications. This may then finally trigger a process of social evolution where social norms and institutions are generated that more efficiently enforce cooperation and prevent falling back into a more primitive non-cooperative state. Once cooperation becomes quasi automatic, as seems to be the case in human infants, again complexity gets reduced as one can automatically assume that everybody involves cooperates. At each stage, we can witness the emergence of common beliefs that are self-fulfilling and thereby automatically stabilize themselves.

Appendix: Game theory with rational agents

In this appendix, we briefly recall the basic notions and results from game theory that will be referred to in this paper. There are many textbooks available where detailed treatments can be found, for instance [23, 49].

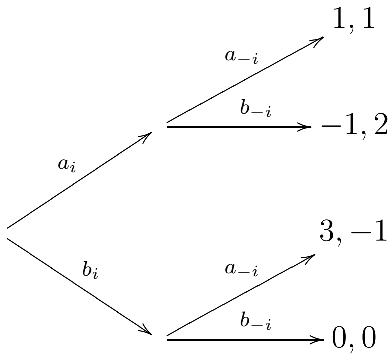
Game theory represents an optimization scheme for interacting agents. We restrict the presentation here to two agents only, i ("she") and $-i$ ("he"). They perform certain actions and receive a pay-off as a result of their own and their opponent's action. Each player wishes to maximize her pay-off, knowing that her opponent is trying to do the same. In particular, both players know each other's pay-offs.

We now introduce some basic formalism for finite games. Everything will be developed from the perspective of i, but since the situation is symmetric, everything will apply to $-i$ as well. Each player can perform a finite number of actions, or strategies as they are usually called. Thus, she has a finite set $S_i = \{1, \ldots, m_i\}$ of pure strategies. $S := S_1 \times S_2$ is called the set of pure-strategy profiles, or the pure-strategy space. We then have the pay-off function

$$\pi_i : S \to \mathbb{R} \tag{19}$$

that assigns to player i her pay-off $\pi_i(s)$ for $s = (s_1, s_2) \in S$, that is, when i plays s_i and

the opponent plays s_{-i}. We also consider $\pi = (\pi_1, \pi_2) : S \to \mathbb{R}^2$. Thus, a game between two players is given by the pair (S, π). The key point is that player can only determine her own strategy s_i while her pay-off also depends on the strategy s_{-i} of her opponent.
The game can either be played sequentially or simultaneously. In the first case, one player, say i, would move first, and the other, $-i$, would choose his action as a reaction to the action selected by i. A sequential game can be represented in the form of a tree. Let us consider an example,



(20)

Here, i can first choose between her actions a_i and b_i, and then $-i$ chhooses between a_{-i} and b_{-i}. At the leaves of the tree, we have written the resulting pay-offs, first i's and then that of $-i$. Thus, when i plays a_i and $-i$ responds with b_{-i}, i will receive -1, but $-i$ will get 2. Such a game can be easily solved by backward induction. It is clear that $-i$ should play b_{-i} in either case because regardless of i's action, b_{-i} in this game will secure him a higher pay-off than a_{-i}. Thus, i, being a rational pay-off maximizer and knowing that so is $-i$, will reason that when playing a_i, she would get -1, but when playing b_i, she would get 0. Since the latter is higher than the former, she will consequently play b_i. Thus, (b_i, b_{-i}) is the unique solution of this game. Other games may have several solutions, but we shall now switch to simultaneous games where similar results apply. The structure of such a game is represented by a pay-off table, as in this example

	a_{-i}	b_{-i}
a_i	4,2	0,0
b_i	0,0	1,3

. (21)

Here, for instance the entries in the lower left box tell us that when i playes b_i and $-i$ plays b_{-i}, then i will get 1 and $-i$ will get 3. Now, the reasoning goes as follows. As long as i plays a_i, $-i$ should play a_{-i}, because he will then get 2, whereas if in this situation, he played b_{-i}, he would get 0. In this game, the same applies to i. Thus, (a_i, a_{-i}) is an equilibrium, in the sense that unilateral deviation is disadvantageous for either player. However, by the same reasoning (b_i, b_{-i}) also is an equilibrium. In the example

	a_{-i}	b_{-i}
a_i	10,0	0,10
b_i	0,10	10,0

, (22)

however, there is no such equilibrium. Whenever $-i$ knows what i plays, he will act in such a way that he gets 10 and i gets 0, but the situation is symmetric between the players. Or putting it differently, for any action combination (s_i, s_{-i}), one of the players can do better by changing her/his action. Thus, in this game, it is best to be unpredictable for the opponent and play each action with a probability of 1/2. If both do that, the expected pay-off for each is 5. When, however, one of them, say i, changed her probabilities and played, say, a_i with probability .4 and b_i with .6, then $-i$ would best respond by playing only a_{-i} as this would increase his expected pay-off to 6 whereas that of i would decrease to 4. We now formalize these insights.

The preceding example tells us that when we want to secure the existence of an equilibrium, it may not suffice to consider pure strategies only, that is, elements of the finite set S_i for player i. We should also include *mixed* strategies, that is, for i, probability distributions on S_i. This means that i plays action α with probability $p_{i\alpha}$ and we have $\sum_\alpha p_{i\alpha} = 1$.

For the general discussion, we need to slightly switch our conventions. i now stands for either of the players. that is, we have players 1 and 2 (of undetermined gender) and i could be either 1 or 2. Before addressing the general question of the existence of equilibria, we first turn to the elimination of non-optimal strategies.

Definition. We say that the (mixed) strategy $q_i \in \Sigma_i$ of player i *strictly dominates* p_i if

$$\pi_i(q_i, r_{-i}) > \pi_i(p_i, r_{-i}) \text{ for any } r_{-i} \in \Sigma_{-i}. \tag{23}$$

We say that q_i *weakly dominates* p_i if

$$\pi_i(q_i, r_{-i}) \geq \pi_i(p_i, r_{-i}) \text{ for any } r_{-i} \in \Sigma_{-i}, \tag{24}$$

with strict equality for at least one r_{-i}.
p_i is said to be *undominated* if it is not weakly dominated by any other strategy.

Strongly dominated pure strategies can be eliminated from consideration, until arriving at a reduced version of the game without strongly dominated pure strategies. We should note that this reduction makes crucial use of the rationality assumption, that is, no player will choose strategies that leave her unconditionally worse off than others, and that each player knows that her opponent is behaving that way. The question whether also weakly dominated strategies can or should be eliminated is more subtle and discussed in the main text.
We now come to the concept of a Nash equilibrium.

Definition. A pure-strategy profile $s^* = (s_1^*, s_2^*)$ is called a *Nash equilibrium* if for both $i = 1, 2$

$$\pi_i(s^*) \geq \pi_i(s_i, s_{-i}^*) \text{ for all } s_i \in S_i. \tag{25}$$

This means that no player can do better by changing her strategy when the other one keeps her strategy. In other words, no player can gain from a unilateral move. Thus, unless the players would or could make some coordinated move – which, however, is not allowed by the rules of the game –, they should stick to their strategy. In this sense, this represents an

equilibrium.
Such a Nash equilibrium in pure strategies need not exist, as we have seen in Example (22). Therefore, the preceding Definition is generalized to

Definition. A mixed-strategy profile $\sigma^* = (\sigma_1^*, \sigma_2^*)$ is called a *Nash equilibrium* if for both $i = 1, 2$

$$\pi_i(\sigma^*) \geq \pi_i(\sigma_i, \sigma_{-i}^*) \text{ for all } \sigma_i \in \Sigma_i. \tag{26}$$

The fundamental theorem of Nash [51] tells us that such an equilibrium does exist, indeed.

We finally discuss the concept of a *rationalizable strategy* introduced in [7, 55]. Again, we restrict ourselves to two players, i and $-i$. Rationalizability is an iterative concept. A strategy s_i^0 of i is rationalizable if it is a best response against a strategy $s_{-i}^!$ of $-i$ that in turn is a best response against some strategy s_i^2 of i which then has to be a best response against some strategy s_{-i}^3 of $-i$ etc ad infinitum. Rationalizability eliminates strongly dominated strategies. In (22) any strategy of i is rationalizable. She might play a_i because she believes that $-i$ would play a_{-i} because he believed that i would play b_{-i} because she believed him to play b_{-i} which he would play in the belief that i played a_i at which stage the reasoning will repeat itself. Likewise, in (21) either strategy is rationalizable.

Acknowledgement

We are grateful to Nils Bertschinger, Eckehard Olbrich and Michael Tomasello for several insightful discussions.

References

[1] R. Aumann and A. Brandenburger, Epistemic Conditions for Nash Equilibrium, Econometrica, 63 (5), 1161-1180, 1995

[2] R.Axelrod, The evolution of strategies in the iterated prisoner's dilemma, in: C.Bicchieri, R.Jeffrey, B.Skyrms (eds.), The dynamics of norms, Cambridge Univ.Press 1997, pp.1-16

[3] P. Battigalli, On rationalizability in extensive games, Journal of Economic Theory 74, 40-61, 1997

[4] P. Battigalli, M. Siniscalchi, Strong Belief and Forward Induction Reasoning, Journal of Economic Theory, 106 (2), 356-391, 2002

[5] P. Battigalli, M.Dufwenberg, Guilt in games, Amer.Econ.Review 97 (2), 170-176, 2007

[6] T. Behne, U. Liszkowski, M. Carpenter, M. Tomasello. Twelve-month-olds comprehension and production of pointing, British Journal of Developmental Psychology 30, 359-375, 2012.

[7] B. Bernheim, Rationalizable Strategic Behavior, Econometrica, 52(4), 1007-1028, Jul. 1984.

[8] C. Boesch, H. Boesch, Hunting Behavior of Wild chimpanzees in the Tai National Park, American Journal of Physical Anthropology, 78, 547-573

[9] C. Boesch, Cooperative hunting in wild chimpanzees, Animal Behaviour, 48, 653-667, 1994

[10] J. Bradfield, C. Stirling, Modal mu-calculi, Studies in Logic and Practical Reasoning, 3, 72156, 2007

[11] J. Bräuer, J. Call, and M. Tomasello, All Great Ape species follow gaze to distant locations and around barriers, Journal of Comparative Psychology, 119-2, 145-154, 2005

[12] L. Camaioni, P. Perucchini, F. Bellagamba, and C. Colonnesi, The role of declarative pointing in developing a theory of mind. Infancy, 5-3, 291-308, 2004

[13] M. Carpenter, M. Tomasello, and T. Striano, Role reversal imitation in 12 and 18 month olds and children with autism. Infancy 8, 253-278, 2005

[14] Cristiane Cäsar, Klaus Zuberbühler, Robert J. Young, and Richard W. Byrne, Titi monkey call sequences vary with predator location and type, Biol. Lett. 9, no.5, 2013

[15] B. Chellas, Modal Logic: An Introduction, Cambridge Univ. Press, 1984

[16] D. Cheney and R. Seyfarth, How monkeys see the world: Inside the mind of another species. University of Chicago Press, 1990

[17] H. Cochet and J. Vauclair, Pointing gesture in young children: Hand preference and language development, Gesture 10, 2-3, 129-149, 2010

[18] H. Cochet and J. Vauclair, Pointing gestures produced by toddlers from 15 to 30 months: Different functions, hand shapes and laterality patterns, Infant Behavior and Development 33, 431-441, 2010

[19] R. Cubitt, R. Sugden, Common Knowledge, Salience and Convention: A Reconstruction of David Lewis' Game Theory, Economics and Philosophy 19, 2, 175-210, 2003

[20] H. Ditmarsch, W. Van der Hoek, B. Kooi, Dynamic Epistemic Logic, Springer, 2007

[21] R. Fagin, J. Halpern, Y. Moses, M. Vardi, Reasoning about Knowledge, MIT Press, 1995.

[22] F. Franco and G. Butterworth, Pointing and social awareness: declaring and requesting in the second year, Journal of Child Language 23, 307-336, 1996

[23] D. Fudenberg and J. Tirole, Game Theory, MIT Press, 1991

[24] S. Gould, The Structure of Evolutionary Theory. Harvard Univ.Press, 2002

[25] R. Fagin, J. Halpern, Belief, awareness and limited reason- ing, Artificial Intelligence 34, 39-76, 1988

[26] H.Gintis, S.Bowles, R.Boyd, E.Fehr (eds.), Moral sentiments and material interests, MIT Press, 2005

[27] S.J.Gould, The structure of evolutionary theory, Harvard Univ.Press, 2002

[28] S.J.Gould and S.Vrba, Exaptation – a missing term in the science of form, Paleobiology 8, 4-15, 1982

[29] H.P. Grice, Meaning, Philosophical Review 66, 377-88, 1957

[30] J. Halpern, Reasoning about Uncertainty, MIT Press, 2003

[31] W. Hamilton, The genetical evolution of social behaviour, I, J.Theor.Biol.7 , 1–16, 1964

[32] B. Hare, J. Call, M.Tomasello, Chimpanzees deceive a human competitor by hiding. Cognition 101, 495-514, 2006.

[33] M. Hauser, The Evolution of Communication, MIT Press, 1996

[34] A. Heifetz, M. Meier, B. Schipper, Interactive unawareness, Journal of Economic Theory 130, 78-94, 2006

[35] K. Jensen, J. Call, M. Tomasello, Chimpanzees are rational maximizers in an ultimatum game, Science 318, 107-9, Oct 2007

[36] P. Johnstone, Sexual selection, honest advertisement and the handicap principle: reviewing the evidence, Biol. Rev.70, 1–65, 1995

[37] J. Jost, External and internal complexity of complex adaptive systems, Theory Biosc. 123, 69-88, 2004

[38] J. Jost, Formal Aspects of the Emergence of Institutions, Structure and Dynamics: eJournal of Anthropological and Related Sciences 1, Article 6, 2005

[39] J. Jost, Temporal correlation based learning in neuron models, Theory Biosc. 125 (2006), 37–53

[40] J. Kirchhof and K. Hammerschmidt, Functionally referential alarm calls in Tamarins. (*Saguinis fusicollis* and *Saguinis mystax*) – Evidence from playback experiments, Ethology 112, 346-354, 2007

[41] A.Kokolakis, C.Smith, C.Evans, Aerial alarm calling by male fowl (Gallus gallus) reveals subtle new mechanisms of risk management, Animal Behaviour 79 (6), 1373-1380, 2010

[42] D.Kozen, Results on the propositional μ-calculus, Theor.Computer Science 27, 333-354, 1983

[43] D. Lewis, Convention: A Philosophical Study, Harvard Univ Press, 1969

[44] D. Lewis, Counterfactuals, Wiley-Blackwell, 2001

[45] E. Lorini, D. Longin, B. Gaudou, A. Herzig, The logic of acceptance: grounding institutions on agents' attitudes, Journal of Logic and Computation, Oxford University Press, 19(6):901-940, 2009

[46] E. Lorini and F. Schwarzentruber, A Modal Logic of Epistemic Games, Games 1, 478-526, 2010.

[47] N. Luhmann, Soziale Systeme, Suhrkamp, 1987

[48] J.M. Macedonia, What is communicated in the antipredator calls of lemurs: evidence from antipredator call playbacks to Ringtailed and Ruffled Lemurs, Ethology 86, 177-190, 1990

[49] M. Maschler, E. Solan, S. Zamir, Game theory, Cambridge University Press, 2013

[50] J. Maynard Smith, Evolution and the Theory of Games, Cambridge University Press, 1982

[51] J. Nash, Non-Cooperative Games, Annals of Mathematics 54(2), 286-295, 1951

[52] J. v. Neumann, O. Morgenstern, Theory of Games and Economic Behavior, Princeton: Princeton University Press, 1944

[53] M. Olson, The Logic of Collective Action, Harvard Univ. Press, 1965

[54] E. Ostrom, Governing the Commons, Cambridge Univ. Press, 1990

[55] D. Pearce, Rationalizable strategic behavior and the problem of perfection, Econometrica 52, 1029-1050, 1984

[56] D.J. Povinelli, J.E. Reaux, D.T. Bierschwale, A.D. Allain and B.B. Simon, Exploitation of pointing as a referential gesture in young children, but not adolescent chimpanzees, Cognitive Development 12, 423-461, 1997

[57] H.J. Rainey, K. Zuberbühler, and P.J.B. Slater, Hornbills can distinguish between primate alarm calls, Proceedings of the Royal Society of London B 271, 205-207, 2004

[58] L.J. Savage, The Foundations of Statistics, New York: Wiley, 1954

[59] J. Searle, Rationality in Action, MIT Press, 2001

[60] R.Seyfarth, D. Cheney, P. Marler,. Vervet monkey alarm calls: semantic communication in a free-ranging primate. Animal Behaviour 28, 10701094, 1980

[61] R.Seyfarth, D.Cheney, Signalers and receivers in animal communication, Annu. Rev. Psychol. 54, 145173, 2003

[62] C.Shannon, The mathematical theory of communication, 1949, reproduced in: C.Shannon, W.Weaver, The mathematical theory of communication, ed. R. Blahut, B.Hajek, pp.29-125, Univ. Illinois Press, 1998

[63] Shettleworth, Sara. ^{2}2010. Cognition, Evolution, and Behavior. Oxford: Oxford University Press.

[64] H.A. Simon, A behavioral model of rational choice. Quart. J. Econom. 69(1), 99118, 1955

[65] H.A. Simon, Rational choice and the structure of environments. Psych. Rev. 63, 129138,1956

[66] B. Skyrms, Signals. Evolution, Learning, and Information. Oxford University Press, 2010

[67] C.N. Slobodchikoff, J. Kiriazis, C. Fischer and E. Creef, Semantic information distinguishing individual predators in the alarm calls of Gunnison's Prairie Dogs, Animal Behaviour 42, 713-719, 1991

[68] C.Smith, A.Taylor, C.Evans, Tactical multimodal signalling in birds: facultative variation in signal modality reveals sensitivity to social costs, Animal Behavior 82, 521-527, 2011

[69] D. Sperber and D. Wilson, Relevance. Communication and Cognition, Blackwell, ^{2}1995

[70] V. Southgate, C. van Maanen and G. Csibra, Infant pointing: communication to cooperate or communication to learn? Child Development 78-3, 735-740, 2007

[71] R. Stalnaker, Knowledge, Belief and Counterfactual Reasoning in Games, Economics and Philosophy, 12, 133-163, 1996

[72] R. Sugden, The Logic of Team Reasoning, Philosophical Explorations: An International Journal for the Philosophy of Mind and Action., 6 (3), 2003

[73] T.Struhsaker, Auditory communication among vervet monkeys (Cercopithecus aethiops). In: Social Communication Among Primates (Ed. by S. A. Altmann), pp. 281324. Chicago: University of Chicago Press, 1967

[74] M. Tomasello, Why don't apes point? in N.J. Enfield and S.C. Levinson (eds.), Roots of Human Sociality: Culture, cognition and interaction. Oxford/ New York: Berg, 506-524, 2006

[75] M. Tomasello, Origins of human communication, Cambridge: MIT Press, 2008

[76] M. Tomasello, Human culture in evolutionary perspective. in M. Gelfand (ed.), Advances in Culture and Psychology. Oxford University Press, 5-51, 2011

[77] M. Tomasello, B. Hare, and B. Agnetta, Chimpanzees follow gaze direction geometrically, Animal Behaviour 58, 769-777, 1999

[78] M. Tomasello and K. Zuberbuehler, Primate vocal and gestural communication, in M. Bekoff, C. Allen, and G. Burghardt (eds.), The cognitive animal: empirical and theoretical perspectives on animal cognition, Cambridge: MIT Press, 2002

[79] R. Tuomela, We-Intentions Revisited, Philosophical Studies 125, 327-369, 2005

[80] L.S. Vygotsky, Thought and language. Cambridge: MIT Press, 1962

[81] L.Workman and W.Reader, Evolutionary Psychology. Cambridge: Cambridge Univ.Press, 2004

[82] A. Zahavi, Reliability in communication systems and the evolution of altruism, in: B.Stonehouse, C.Perrins (eds.), Evolutionary Ecology, London: Mac Millan, 1977, pp. 253–259

[83] K. Zuberbuehler, D.Cheney, R.Seyfarth, Conceptual semantics in a nonhuman primate. J. Comp. Psychol. 113, 3342, 1999

[84] K. Zuberbuehler, Referential Labeling in Diana Monkeys, Animal Behavior 59, 917-927, 2000

[85] K. Zuberbuehler, Predator-specific alarm calls in Campbell's monkeys, *Cercopithecus Campbelli*, Behavioral Ecology and Sociobiology 50, 414-422, 2001