Social calculus as the origin of syntax

Myrthe Dijkman (0715212)

June 29, 2012

Bachelor Thesis

Title: Social calculus as the origin of syntax

Author: Myrthe Dijkman

Student number: 0715212

Supervisors: Dr. W.F.G. Haselager and Dr. I.G. Sprinkhuizen-Kuyper

Date: June 29, 2012

Radboud University Nijmegen

Contents

Abstract

When examining the evolution of syntax, the question why we would start using syntax in the first place attracts attention. Bickerton (2002) proposed that the syntactic categories of agent, theme and goal were not developed for use in language, but existed as part of a much older phenomenon: social calculus. I have implemented this theory by creating and comparing two groups of agents in a simulated world. One group has a full social calculus, while the other is lacking the ability to categorize.

The simulation shows that the two groups differ in both grouping and sharing behaviour. The agents with social calculus are more selective in their social interactions and also more robust as a population. These behavioural differences lead to differences in the ages that agents reach, but those differences are very situation dependent.

1 Introduction

"Look, more fruit!"
"You're hurting me!"
"Mommy, I'm hungry!"
"A bear, run!"

I think we can all see why communication is extremely useful to any social species. Your survivability increases greatly when you can let each other know where good food sources are, what you need at the moment and, most importantly, can warn everyone immediately when someone has spotted danger. Many animals have evolved this kind of communication in more or less sophisticated forms. We all know the intricate dances bees use to tell the rest where honey can be found, or the frantic flight of a group of deer when one of them calls out a warning.

Humans have taken their communication one step further. Our language capacity is a unique ability, and also a very complex phenomenon. The question "How did language evolve?" is far from solved, though it has received a lot of attention in the past decades. This thesis focuses on the structural side of language, in other words: syntax. More specifically, it is based on a theory by Bickerton (2002), which states that the foundations of syntax are based on the concept of social calculus.

First I will explain a bit more about the nature of language and the assumptions that are the basis of the field of language evolution. I will also briefly describe some of the more important theories and experiments in this field. After that I will explain Bickerton's theory in more depth, introduce the research questions and sketch the outlines of the experiment.

1.1 The evolution of language

Most researchers assume that the call system that many animals use forms the basis for human language evolution. However, these basic communication systems lack two essential properties that are needed before we can call it language. Those properties are symbolic representation and structure.

<u>Symbolic representation:</u> One of the essential features of human language is that the symbols we use to refer to things, namely words and phrases, bear no similarity or direct connection to the concept or object they are referring to. It is unclear to what extent animals have the ability to use symbols. For example apes and primates have in captivity successfully learned to use symbols for communication. A lot of work is done in this area, but this is not the focus of this thesis.

<u>Structure:</u> The other essential and defining feature of human language is its structure, in the form of syntax. This is not merely a formality; structure actively contributes to the meaning of each sentence. Today we cannot imagine our language to be devoid of any form of structure.

However, when we go back in time and imagine a primitive proto human group, using calls to communicate about important situations, then the question arises: "Why would they start using structure in those calls?"

1.2 The evolution of syntax

Both the evolution of language in general and the evolution of syntax in particular are still very poorly understood. Many theories exist and researchers from many different fields have attempted to shed some light on it (see Christiansen & Kirby (2003) for an overview). One reason for this plethora of theories is that human language is a highly complex phenomenon. At least three very different but interacting systems play a role: individual learning, cultural transmission and biological evolution. On top of that, the debate of what language actually *is*, what its defining features are, is still very much alive.

Another reason is that the development of language may very well have been one of the key factors that set the early humans aside from the other ape species. This raises a number of fundamental questions like “What properties made that proto humans could evolve language, and other primates did not?”. Answering this and other questions may have implications that go far beyond the field of language evolution.

One example of a cultural take on syntax evolution is an experiment done by Kirby (2000). This simulation shows that a syntax can evolve over many generations, starting with random utterings, as long as each generation of agents is actively trying to discover a structure in these utterings. The resulting syntax can discriminate between nouns and verbs. However, the crucial condition here is that agents are actively looking for structure before there is an actual structure to find. Bickerton (2002) suggests a theory that reverses this order. This theory forms the basis for the rest of this thesis and is explained in the next section.

1.3 Bickerton's theory

Bickerton argues that the basic elements that make up sentence structure were not developed when man started speaking, but had already existed as concepts a long time. They had developed as part of what Bickerton calls a 'social calculus', which developed when reciprocal altruism did. Reciprocal altruism is a concept that has received evidence from observations of apes and some monkey species (Seyfarth & Cheney, 2003), as well as theoretical evidence (Trivers, 1971). The basic idea is 'If I do something good for you now, you'll do something good for me later'. The 'later' here could be weeks or even months later. This mechanism is applied to a variety of social interactions within a group of animals, including grooming, sharing food and aiding each other in disputes. Of course these kinds of relationships are subject to 'cheaters', individuals that give back less than they receive.

With this social structure in mind, a strong selective pressure would exist towards individuals that are able to detect cheaters reliably. This is accomplished by a mechanism which is called social calculus. Social calculus is a mental 'score sheet' that keeps track of hundreds of interactions within the group an individual lives in. The purpose of this score sheet is to detect cheaters in the group, allowing the individual to avoid those cheaters in the future, and build ties with trustworthy members of the group instead.

This social calculus probably does not consist of a simple enumeration of all actions. It is more likely that different types of actions are tabulated separately. Each action has fixed categories of participants. For instance, the action 'grooming' has two categories, the one that grooms and the one that is groomed. Note that each individual can perform each of these roles at a certain time. This calls for a certain level of abstraction in the brain, because these categories cannot be tied to given sets of individuals.

We can identify three categories that can together cover all possible actions.

- Agent: the one that is performing the action,
- Theme: the one, or the object, that undergoes the action
- Goal: the recipient

1.3.1 Agent, theme and goal

The concepts of agent, theme and goal originate from Noam Chomsky's Government and Binding Theory (1981). The Government and Binding Theory uses theta-roles to represent the syntactic argument structure required by a given verb. For example, the verb “eat” requires two arguments, something that eats and something that is eaten, so it is said to “assign” two theta-roles. The theta-roles that a verb assigns must be filled, otherwise the sentence is ungrammatical.

Agent, theme and goal are the three theta-roles that a verb can assign. The verb “give” assigns all three. Here the agent is the one that is doing the giving, the theme is the object that is given and the goal is the receiver of the object. Many verbs assign less than three theta-roles. To repeat the earlier example of the verb “eat”, this verb only requires two theta-roles, an agent and a theme. The agent is the one that does the eating, while the theme is the object that is eaten. A verb like “sleep” only assigns one theta-role, the theme. This may seem counter-intuitive, since one might expect an agent to be assigned instead of a theme. However, a verb like “sleep” denotes an action that you cannot actively perform. In essence, “you sleep” means something more along the lines of “you are being slept”, so this kind of verb assigns a theme and not an agent.

While theta-theory is a syntactic theory in origin, it has strong connections to the semantic side of language. Not only are the theta-roles named after their most prominent thematic relation to the verb, the number of theta-roles that a verb assigns is not determined syntactically, but semantically. Syntactic structure can differ between languages, but there can be no language where the verb with the meaning “give” assigns only one theta-role, or the verb with the meaning “sleep” assigns two. These theta-roles are fixed because of the meaning of each verb, which is a semantic property, not because of some structural or syntactic property of the language.

Because of this duality, theta-theory is not usually considered a central property of syntax. When describing syntax, linguists value the autonomy of syntax as very important, which means that any mechanism that has semantic as well as syntactic implications is automatically disqualified from any major role. However, this same duality makes it an excellent mechanism for introducing syntax into a protolanguage that had until that point consisted only of semantic elements.

If we assume that the beginnings of syntax were built upon concepts that were developed in primates' minds for social calculus, the first question that needs answering is “Would primates have developed these concepts of agent, theme and goal for their social interactions?”. Evolutionary theory dictates that any system is only viable if it provides the population with some kind of advantage. So using the concepts of agent, theme and goal must increase the survivability of primates in some way, else it becomes unlikely that they would have developed at all.

1.3.2 Research questions

I have constructed a simulation that implements this theory. The aim of this implementation is twofold. First, the attempt to implement part of the theory may reveal gaps or flaws in the theory, and will show whether the theory is clear and complete enough to be modeled into a simulation. Second, the implementation can provide some support for the theory.

In this experiment we will compare two groups of agents: calculus agents and associative agents. The calculus agents use social calculus for their interactions with others. In order to use their social calculus they need to possess the ability to recognize the roles of agent, theme and goal when they see other agents interacting.

The calculus agents will be competing against so-called 'associative agents'. These agents have no social calculus and do not use any categories, but they simply associate agents with food if they see them interacting with food. The associative way of learning is often suggested as an alternative for “true” pattern learning or relationship learning in various animals, including monkeys (see for example Zentall, Wassermann, Lazareva, Thompson & Rattermann, 2008), which makes it a logical strategy to test against social calculus.

If agents with social calculus perform better than those without social calculus, then the assumption that early humans possessed a social calculus becomes more likely. For this part there are two research questions:

1. Does possessing social calculus increase performance of agents?
2. Does possessing social calculus alter the behaviour of agents?

Performance increase will be measured primarily in terms of the mean age that the agents reach in each population, and additionally in terms of the spread of those ages. Behavioural changes can be measured in terms of the amount of sharing that is done, and can also be determined by visually observing the simulation.

2 Method

2.1 The world

The world consist of a large open space, 150 'units' large, surrounded by walls. There are no further obstacles in the world. Agents cannot move through the walls, when they collide with one of them they will “bounce” and move on in the new direction. Food is placed randomly in the world, there are no places where food is more or less likely to spawn. Time in the world is measured in terms of 'turns'. Each turn allows all agents to move and/or perform other actions.

2.1.1 Food distribution

There are multiple possibilities to handle the food distribution. The way that is perhaps most obvious is to generate food randomly as time goes by, irrespective of how much food exists in the world at that time. This also seems to be the most natural way, since this is how food would appear for a foraging species. They will find food at random locations, and eating the food at one location does not influence the appearance of food at another location in any way.

However, within a controlled simulation and a limited space for the agents in which they can move, this creates the effect that a lot of agents will die quickly, leaving only as many alive as the world can sustain for a certain rate of food generation. These 'survivors' may live very long lives, and have therefore a disproportionally large impact on the mean age and other statistics. There is even a chance that they will live forever, if enough competition has been eliminated and the agent is lucky enough.

For this simulation this is not a desired effect. The differences between the two groups of agents only become visible once the agents have been living in the world and interacting with each other for quite a while. It is important to keep as many agents as possible alive for a decent amount of time, to give them enough time to fill their score sheets with data from their observations and start acting on it. The only difference between the two groups is in how they witness and handle social interactions, so this is also why having two or three agents that live a very long life does not say much about the difference between the two groups. When there are so little agents that hardly any interactions take place, the two groups become virtually identical.

There are at least two ways to prevent this, which leaves us with three different options to handle the food distribution:

1. Set a generation rate without a minimum of maximum
2. Set a fixed maximum amount of food
3. Make the maximum amount of food dependent on the number of live agents.

When using the second method, to set a fixed maximum, this maximum would depend on the size of the world. Food is generated whenever the amount of food that is currently in the simulation falls below the maximum amount. This way guarantees a somewhat even spread of food across the world, not dependent on how many agents are alive and how much food is eaten. This method counters the effect described above partly, but not completely. While food is generated faster when more food is eaten, there is still more competition for each piece of food when more agents are alive. After the population has thinned out, survival becomes easier for the remaining agents, though not as dramatically as with the first method.

The third way is to make the maximum amount of food dependent on the number of live agents, so that less food gets generated as agents die. This ensures that agents will never profit from the deaths of other agents, since any extra food that they would have access to is also gone. However, this method may achieve that goal a little too well, because the world does not get smaller when less agents are alive, so the remaining agents have to travel larger distances on average to get to a piece of food. This means the remaining agents are often at a disadvantage once most agents have died, which causes them to die relatively quickly.

I have chosen to use the second method as the basic food generation mechanism; to set a maximum amount of food for a certain world size and keep the amount of food stable. While not the most natural method, I believe it has the least effect on the experiment and causes no unwanted interference. The first method creates the effect that most agents die early while a few live very long lives, and this is not an effect that allows a good investigation of the research questions. Method 3 creates the opposite: when more agents die, surviving becomes increasingly difficult for the remaining agents. While this effect may not be as detrimental to the experiment as the effect of the first method, it still influences the outcome. The second method seems the most 'neutral' one and therefore the best.

2.1.2 Seasons

Within one simulation run the amount of food available is varied, to simulate a 'summer period', in which plenty of food is available, and a 'winter period' in which far less food can be found. Each season lasts 3000 turns. The exact amounts of food differ, since a few different settings were used. An overview of the food amounts used in different experiments can be found in table 1.

	Summer	Winter
Experiment 1	150	60
Experiment 2	90	30

Table 1: maximum food amounts for different experiments

2.2 The agents

2.2.1 Agents and energy

The primary goal of each agent is to keep its energy supply up. Agents start the simulation with 1000 energy units and they lose 1 energy unit in each turn of the simulation. This is a fixed amount, it does not depend on the activity of the agent. Since agents cannot rest to lose energy less quickly, all agents are constantly moving. They can regain energy by eating food, either self-found food or food that was received from another agent. When food is shared between agents, the energy that is gained from the food is shared as well. An agent dies when its energy drops below zero. An overview of the variables related to energy can be found in table 2.

Name	Description	Value
Energy	The amount of energy that an agent has.	1000 at the start of the simulation, varies after that. When an agent has 0 energy left, it dies.
EatingThreshold	When an agent has more energy than this, it is fully satisfied and will not eat any food.	2000
Starving	When the agent has less energy than this, it is considered starving	200
FoodEnergy	The amount of energy that is gained from eating one piece of food	300
SharingRatio	The amount of food, out of the total, that is given away when sharing food. 0,6 means 60% is given away and 40% is eaten by the agent itself.	0,5

Table 2: variables related to energy consumption and management

2.2.2 Agent behaviour

All agents have the same behavioural algorithm. The only difference between the two groups is in how they perceive interactions between agents, and as a consequence of that, what information they store about other agents. Since the two groups store different information about other agents, they also have different ways to determine whether they consider a given agent an ally or an enemy, and how 'good' the ally is. I will describe this process in more detail in the respective agent sections.

Basically, all agents have two immediate goals: finding food, and staying close to an agent they consider an ally. Food will always have the first priority for any agent, if they can see food they will immediately go towards it and pick it up. They have a vision range that determines from what distance they can detect food. This vision range is equal for all agents and is relatively small (see Table 3 for precise values of this and related variables). They also have a vision range for detecting other agents, this range is 4 times as large as the food vision range.

When an agent has found food, it needs to decide what to do with it. There are three options: eat it all, share it with another agent, give it away, or carry the food with it. When an agent is “starving”, which means it has very little energy left, it will always eat the entire unit. All agents have an “eating threshold”, which means that if they have more than a certain amount of energy, they are considered fully saturated and cannot eat food until their energy gets below that threshold again. If the agent is completely saturated, it will give its food away if there is an ally nearby that isn't saturated. If it has some more energy to spare, but is not saturated, then it may decide to share the food with another agent, or it can eat it by itself. If it decides to share, then it will look for the 'best' ally nearby, and if that ally has less energy the agents share the food. Otherwise it will still eat the entire unit. If the agent cannot eat the food and has no ally nearby to give it to, it will carry the food with it. It will keep carrying the food until there is an ally nearby and/or it becomes possible for the agent to eat the food itself.

Each agent, from both types, has a random 'social score'. This social score determines how likely they are to share food with another agent. This creates mixed populations, from very social agents that will share food freely, to 'misers' that will never give food away, and everything in between. Essentially the agents with a very low social score act as cheaters, because they will receive food every now and then, but they will never give any back.

If they come near another agent, they may decide to follow that agent around, if that agent is an ally. They will never follow an enemy, and if an agent becomes an enemy they will immediately cease following that agent. They will try to stay near the other agent, but never closer than their own food vision range. This way, they can always search different parts of the world at any given time, making the search for food more efficient. When another agent comes near, they can switch which agent they are following if the other agent is a 'better' ally. They can also stop following anyone and start wandering on their own again. This can happen at any time and depends on a fixed chance that is the same for all agents.

2.2.3 Associative agents

These are the most simple agents, meaning that they have no full social calculus. They also do not have the capability to use the categories of agent, theme and goal. However, they do witness the social interactions that take place around them, and are able to recognize individual agents. They also keep a score sheet in which they can store some of the information they gather from both interacting and watching others interact. The main difference between these agents and the ones with a full social calculus, is that these agents cannot identify the roles of the participants of an interaction. That means that, while they can witness an interaction in which one agent shares food with another agent, they cannot identify which agent is the one that is giving the food, and which one is the one that is receiving the food. All they can identify is that these two agents are interacting, and that food is involved. This leads to an associative form of score keeping; these agents simply keep track of how often others are seen interacting with food.

These agents usually consider everyone a potential ally. The 'worth' of each ally depends on the number of times it has been associated with food. The more associations, the better the ally becomes. However, a “creep factor” is introduced, which causes agents to label an agent that is associated with food too often as an enemy. When an agents association-score is higher than all other agents' scores combined (in other words, when one agents score is more than half of the total), this agent becomes a “creep” and others will avoid him until the score is down to less than half of the total. This is done to avoid a self reinforcing pattern: an agent is associated with food a few times, so some agents decide to give him food, and with each agent that gives him food the association becomes stronger which causes more agents to give their food to him.

2.2.4 Calculus agents

These agents keep a score sheet that stores which agents they owe something, and how much. They also store the amount other agents owe them. When they witness a social interaction, they can identify the participants and for each agent store whether it is receiving food or giving it away. So they keep two kinds of score lists: one for their personal interactions with others, and a separate one for interactions they witness.

These agents will consider the agents they owe something their best allies, since those agents have apparently been generous towards them. They will also ally with those who they have witnessed giving food to others, or those agents that owe them only 1 or 2 food units if no better allies are available. Agents that owe them three or more food units are considered enemies, as they are apparently not repaying their debts very well.

Name	Description	Value
EatingThreshold	When an agent has more energy than this, it is fully satisfied and will not eat any food.	2000
Starving	When the agent has less energy than this, it is considered starving	200
VisionRange	The maximum distance at which an agent can see other agents	30
FoodRange	The maximum distance at which an agent can see food	7,5
WorldLength	The length of the world (which is always square-shaped)	150
MaxFood	The maximum amount of food units that an agent can carry.	3
SocialScore	The chance that an agent decides to share its food instead of eating it all.	A value between 0 and 1, different for each agent

Table 3: all agent-related variables

2.3 Encouraging social behaviour

The first test runs showed that the current setup appeared to favour the agents with a low social score. In these test runs, the chance that an agent would start following another agent was also dependent on their social score. The agents with a low social score generally lived longer than the agents that shared food and formed groups. I have identified two main reasons for this:

1. Groups do not move very much, since all agents have random movement but move back to the others if they move too far away, so any food in the area is quickly eaten
2. Ultimately the amount of food an agent get is the sole factor that determines its age. Not giving any food away (and occasionally receiving some food from others that have not yet discovered that you won't pay back) maximizes your amount of food.

However, this experiment is not about social agents versus non-social agents. We want to test two different strategies to approach social interaction. The non-social agents of both types behave very similar, so if their strategy is the best one, any differences between the groups may become invisibly small. This is why social behaviour needs to be at least a valid strategy, if not the better one.

2.3.1 Lévy walks

The first attempt to encourage movement over larger distances was to implement Lévy walks as the standard algorithm for movement. Until this point, agents moved each turn in a semi-randomly chosen direction (the previous direction influenced the new direction). The idea of Lévy walks is that agents move for a certain time frame in one direction. These time frames are usually short, but have a small chance to be very long. The distribution from which the time frames are sampled behaves as $\frac{1}{x^2}$. I have used a minimum value of 5 (their vision range for food) and a maximum value of 5000 (approximately the time the agents need to cross the entire world).

Lévy walks using this distribution show a behaviour that looks like a random walk for most of the time, when the algorithm returns small time frame values. However, once in a while, the agent will travel a long distance in one direction. The mean distance traveled increases linearly with the number of steps taken and the chance to visit a place that is already visited is smaller than with random (Brownian) movement.

It has been shown that Lévy walks are a good approximation for many movement patterns seen in many different animals (see for example Viswanathan et al., 1996). It has also been shown that Lévy walks are more efficient for foraging than random (Brownian) movement (Viswanathan et al., 2000).

2.3.2 Following behaviour

Implementing the Lévy walks increased the efficiency of the foraging quite a bit, and also seemed to make groups a bit more mobile. However, the difference in age between the social and non-social agents wasn't completely eliminated. The next step was to remove the connection between the social score and the following behaviour. Instead of letting the chance to follow depend on the social score of the agents, all agents now have an equal chance to start following another agent if they like the other agent enough.

This simulation is designed with social animals like monkeys in mind. However, the environment and options for interactions are quite limited; only food plays a part. In reality the chance to find more food is only a small part of why animals would be living in groups. Protection from predators and increased survival of their offspring are very valid reasons to live in groups, and there are many more. So this is why the less social individuals are 'forced' to seek out others just as much as the more social individuals.

With this change, the non-social agents lose their advantage of having all the food in their area for themselves. All that remains is that they are not sharing their food as often as the more social agents.

2.4 Parameters for the experiment

There are two parameters that are varied between conditions:

- Spread of social scores
- Type of agents

2.4.1 Spread of social scores

There are two conditions for this. In the first they are randomly generated numbers between 0 and 1, distributed evenly with the mean approximately at 0.5. This condition is referred to as the “standard social scores” condition. In the second condition there are relatively more agents with a higher social score, to simulate a population that consists mainly of social agents and some exceptions. In this condition 10% of the population has a social score between 0 and 0.3, 20% has a social score between 0.3 and 0.7 and the remaining 70% of the population has a social score between 0.7 and 1. This condition is referred to as the “high social scores” condition.

2.4.2 Type of agents

In one condition there are only calculus agents in the world, while in the other condition there are only associative agents in the world. The two types of agents are never mixed.

2.5 Experiments

There are two experiments conducted. The first experiment is the basic experiment. The second experiment is identical to the first except that the “following” behaviour of agents has been removed.

The experiments have a 2x2 design. The two parameters that are varied are the social scores (standard vs high) and the type of agents (calculus vs associative). Section 2.4 describes these in more detail. Each experiment consists of 3 simulation runs per condition, which amounts to 12 simulation runs total.

Each simulation run is started with 60 agents in the world. These agents will “live” in this world for a while, behaving as specified in the previous chapter, until they die. There are no additional agents generated in any way. Once the number of agents is down to 10% or less of the initial population, the simulation is cut off. This is done because these last agents can potentially live for a long time, but with so few agents in the world there is very little opportunity for any social interaction. As explained before, without social interaction the difference between the two groups is meaningless, so this stage of the simulation runs is cut off.

For each simulation run there are two types of data collected.

1. The age of each agent is recorded as it dies. From these individual dying ages the mean dying age is calculated for each simulation run.
2. The number of sharing actions that took place over the entire simulation is recorded. From this number a sharing ratio can be computed that gives us an idea of the frequency at which the agents have shared food in that simulation. The sharing ratio is defined as the number of sharing actions divided by the total age of all agents in the simulation.

3 Results

3.1 Experiment 1

This is the basic experiment, with all behaviour and parameters as described in the previous chapter. The maximum food amounts used were 150 in summer and 60 in winter.

3.1.1 Ages reached

The difference in the mean age that the agents reach in each run between the two groups of agents is very clear. Table 4 shows the mean age from each run.

Mean ages in Experiment 1	Calculus	Associative
Standard social scores	23146	7166
	17558.5	3861
	21371	3718.5
High social scores	20716	5128.5
	33911	5121
	17593.5	2883.5

Table 4: The mean age reached for each simulation run. The numbers are not very close together, especially in the calculus group, but it is clear that the calculus groups consistently reaches higher ages than the associative group.

This data is visualized in Figure 1 and Figure 2. Figure 1 shows the mean ages reached in the high social score condition. Figure 2 shows the mean ages reached in the standard social score condition. These figures show a very clear pattern for both social score conditions: calculus agents live much longer lives than associative agents. A Mann-Whitney U-test showed that this difference is significant ($p = 0,002$).

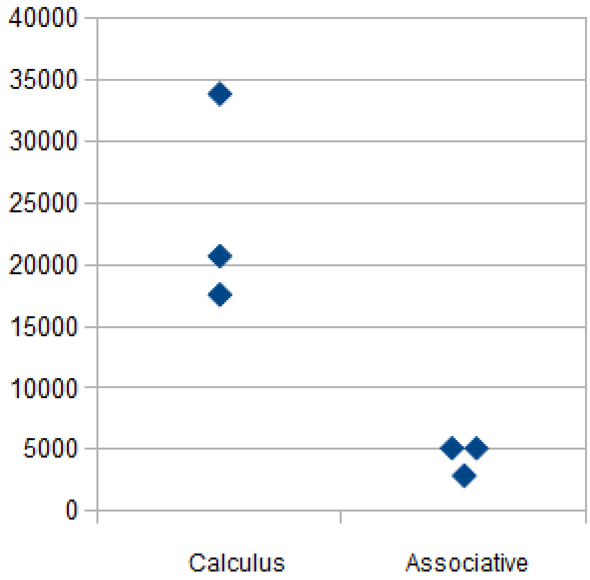

Figure 1: Mean ages for the high social score condition. The calculus agents reach much higher ages than the associative agents, the differences between the runs are also bigger.

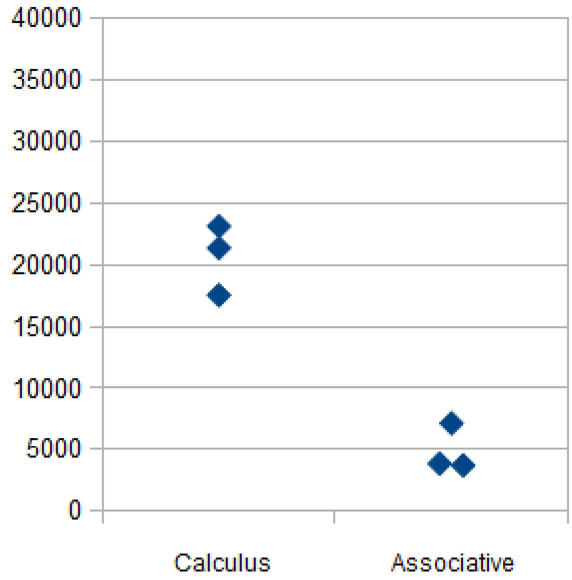

Figure 2: Mean ages for the standard social score condition. Again the calculus agents reach much higher ages than the associative agents.

There is no clear trend visible from the effect of the social score condition on the age that the agents reach. This is confirmed by the Mann-Whitney U-tests: there is no significant effect from social score condition on age reached, neither in the overall population, nor within one of the agent groups ($p = 1.000$ in all three cases).

3.1.2 Number of sharing actions

Table 5 shows the average number of sharing actions per agent turn. These sharing ratios also show a clear pattern. As Figure 3 shows, calculus agents share less food than associative agents. This difference is significant ($p = 0{,}002$). Agents from both groups also share more in the high social score condition, but that “effect” is meaningless since their chance to share depends directly on their social score.

Sharing actions in Experiment 1	Calculus	Associative
Standard social scores	0.0002/turn 0.0003/turn 0.0002/turn	0.0009/turn 0.0010/turn 0.0011/turn
High social scores	0.0004/turn 0.0003/turn 0.0004/turn	0.0015/turn 0.0014/turn 0.0014/turn

Table 5: The number of sharing actions per agent turn that took place in each of the simulation runs. Associative agents shared more than calculus agents in both social score conditions. Both groups also shared more in the high social score condition than in the standard social score condition, but that is to be expected since their chance to share is higher in that condition.

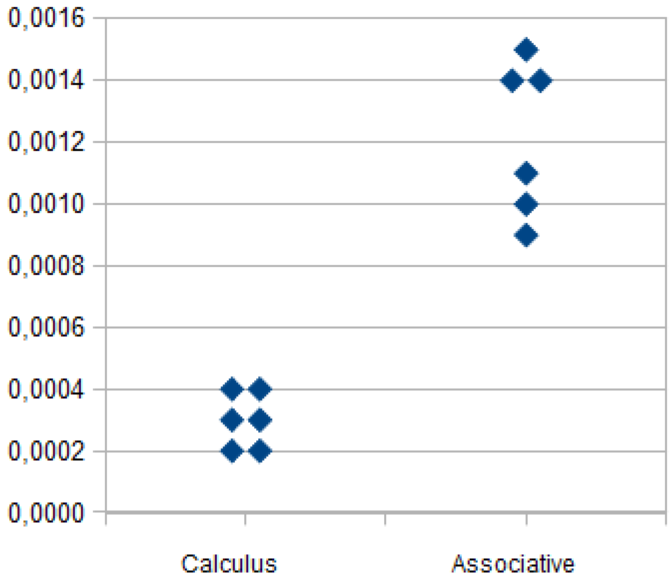

Figure 3: The sharing ratios for both types of agents in Experiment 1.

3.1.3 Behavioural observations

There is a clear difference between the groups in the grouping behaviour visible when watching the simulation run. The calculus agents usually form about five groups that are each about ten agents large. Figure 5 shows a screenshot of a situation like that. These groups are very loose, agents are constantly shifting between groups or abandoning groups. The associative agents on the other hand are slowly but inevitably drawn into one big group that usually settles more or less in the center of the world. Figure 4 shows an example of a situation like that.

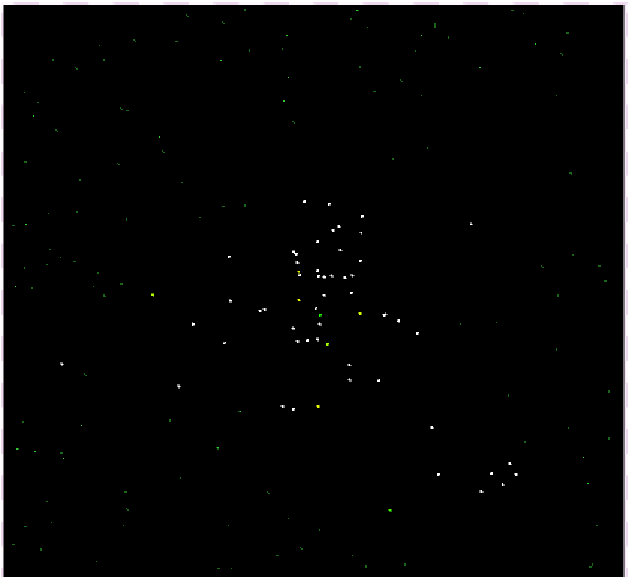

Figure 4: Screenshot of the associative agents. The bigger white, yellow and green dots are all agents, the tiny green dots are food. All agents are gathering in the center of the screen.

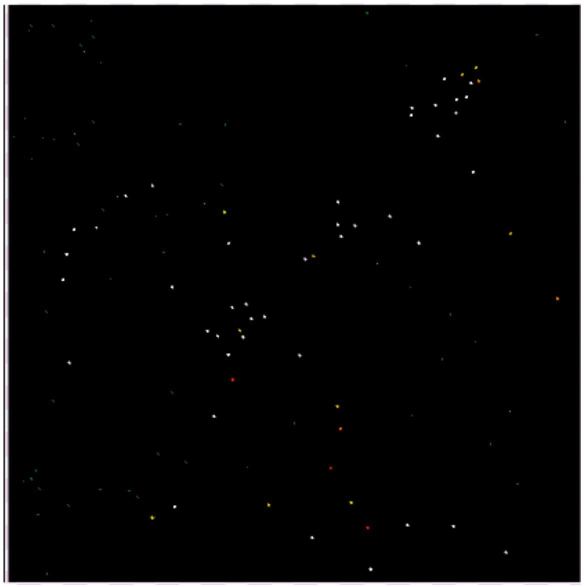

Figure 5: Screenshot of the calculus agents. The bigger white, yellow and red dots are all agents, the tiny green dots are food. You can see a group in the upper right corner and two less clearly defined ones below that. The string of agents on the left is also a group, they are following each other.

3.2 Experiment 2

Based on the visual inspection and the data of experiment 1 one might suggest that the grouping behaviour has a rather large influence on the ages that the two types of agents reach. Therefore, a second experiment was run in which two things were changed compared to Experiment 1.

1. The "following" algorithm was removed.
2. Only 90 maximum food in summer and 30 maximum food in winter.

The food amounts were reduced because all agents appeared to be much more efficient in finding food when they are not following others around. Other than these two points, the two experiments are identical. Just as in Experiment 1, the ages of all agents are recorded, as well as the amount of sharing actions. There is no behavioural analysis since the agents cannot show any group behaviour without the "following" algorithm.

3.2.1 Ages reached

The mean ages reached are shown in Table 6. They display a trend that is the opposite of Experiment 1. The associative agents reach much higher ages than the calculus agents, as can clearly be seen in Figure 6 and Figure 7. This difference is significant ($p = 0{,}002$). These figures also show that the mean ages for the calculus agents are much closer to each other than the means for the associative agents, where one run can easily be twice as long as another.

The social score again has no significant effect on the ages reached, neither overall ($p = 0.589$) nor within the agent groups ($p = 0.7$ in the calculus group and $p = 0.4$ in the associative group).

Mean ages in Experiment 2	calculus	associative
Standard social scores	14233,5 8363,4 12060,9	104058,5 42521 55461
High social scores	13058,5 13746 12231	44101 136653,5 127858,5

Table 6: The mean ages reached in Experiment 2

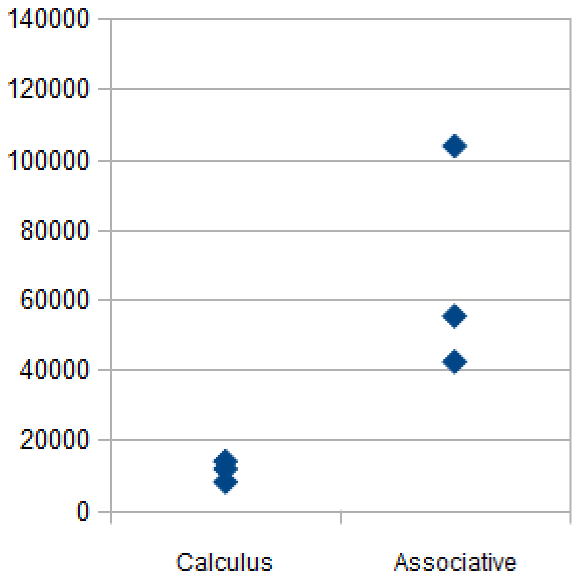

Figure 6: Mean ages reached for the standard social score condition in Experiment 2

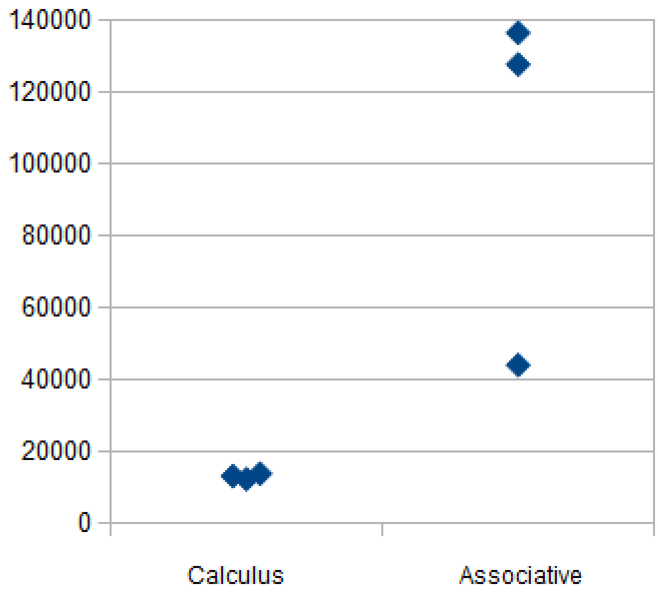

Figure 7: Mean ages reached for the high social score condition in Experiment 2

3.2.2 Number of sharing actions

Table 7 shows the average number of sharing actions per agent per turn for experiment 2. The same trends are visible that were also there in experiment 1: associative agents share more than calculus agents, and both groups share more in the high social score condition. The difference in sharing actions between the two groups of agents is significant ($p = 0,002$). The fact that both groups share more in the high social score condition is due to the fact that the agents' chance to share depends directly on their social factor.

Sharing actions in experiment 2	Calculus	Associative
Standard social scores	0,0004	0,0012
	0,0006	0,0013
	0,0006	0,0013
High social scores	0,0005	0,0019
	0,0006	0,0017
	0,0006	0,0017

Table 7: The number of sharing actions for Experiment 2

3.2.3 Age difference within the population

From observing the simulations, as well as Figure 8 and Figure 9, it seems that while the associative agents live longer, they also die faster once the first agents start dying. Table 8 shows the number of turns that passes from the moment that the first agent dies until the moment the last agent dies. For the associative agents there were two runs in which one or two agents died early, but the rest kept going for a long time after that. In these cases I excluded these agents and started counting from the moment that a larger part of the population died. The calculus agents did not have cases like that, when one died more would soon follow.

Age differences in experiment 2	Calculus	Associative
Standard social scores	20850	11100
	16800	11700
	14700	7050
High social scores	18150	6000
	19800	7350
	19350	5400

Table 8: The age differences between the first agent that died and the last agent that died for each simulation run in experiment 2.

As shown in table 8, associative agents die faster than calculus agents once they start dying. This effect is significant ($p = 0,002$). From the raw data, it seems like the social score might have an effect on this within the associative agents group, but this effect is not significant ($p = 0,2$).

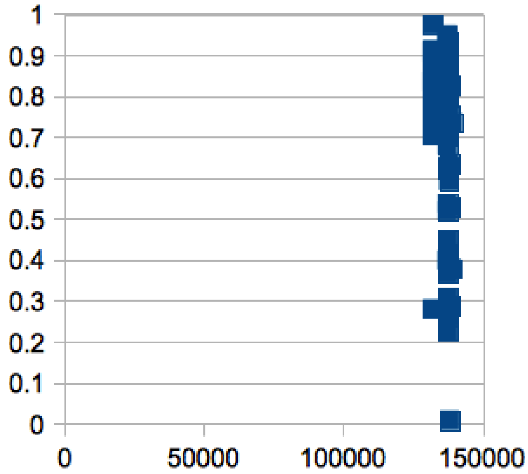

Figure 8: A typical dying pattern of the associative agents in experiment 2. The individual dying ages are set out against the social score of each agent. All agents live very long, and then suddenly the whole population dies.

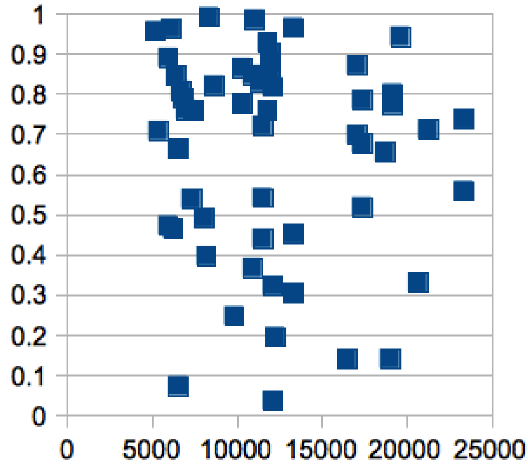

Figure 9: A typical dying pattern for the calculus agents in experiment two. The individual dying ages of the agents are set out against their social score. The population dies much more gradually than the associative agents.

4 Discussion

While the results themselves seem clear and the differences are very large, analyzing them proves a little less simple. The basic picture is that calculus agents get much older in the basic experiment, while associative agents get much older in the non-following experiment. In both cases it does not matter whether the social scores are standard or high. Associative agents always share more often than calculus agents.

When examining the difference between calculus agents and associative agents further, two properties come to mind: selectiveness and robustness. These will be discussed in detail in the following two sections.

4.1 Selectiveness

The first result that attracts attention is the difference in grouping behaviour between the calculus agents and the associative agents. Calculus agents form small, very loose groups that change a lot, while associative agents form one big group that doesn't move or split once it is formed.

Where does this big difference in grouping behaviour come from? All agents have the exact same following algorithm, their only difference is how they judge other agents based on what they observe. And here the calculus agents can be much more selective than the associative agents, because their observations allow them to blacklist anyone that doesn't seem to be paying back their dues. When they have blacklisted another agent, they will also stop following that agent. The associative agents have the possibility to exclude one other agent (the “creep” that was mentioned earlier), but all others are potential “following material”. Thus the calculus agents create a much more dynamic and shifting interaction pattern, while the associative agents form one big group that almost never excludes anyone.

In the second experiment, the calculus agents die much sooner than the associative agents. A possible explanation for this difference can be found in the amount of sharing actions that has taken place in each simulation. The associative agents have shared about 4 times more often than the calculus agents throughout their simulation. In this case, being more selective in their sharing seems to have a negative outcome for the calculus agents. All agents only share food with an agent if that other agents energy level is lower than its own energy level. Taking this into account, the strategy of more sharing with everyone means that the whole population is trying to keep each other alive, and succeeding pretty well. The calculus agents only try to keep those agents alive that are paying back frequently, and apparently that is not enough. This is very visible in the pattern of the individual dying ages of the agents, as shown in graph 1 and graph 2 in the result section.

The ability to be selective about your allies is an inherent property of social calculus. This system is designed just as much to determine who you should avoid as to determine who your friends are. The associative system lacks this property. It can tell you who you may want to favor, but, aside from the “creep”, there is no clear cut-off point after which you should avoid an agent. An agent's score can go from low (meaning neutral, not particularly favored) to high (a wanted ally), but there is no equivalent of a negative score, and thus no intuitive way to designate enemies. Having no way designate enemies means that they can not be selective in their choice of friends; anyone that is near qualifies.

4.2 Robustness

Another observation is that the associative agents seem to form a more vulnerable society. In experiment 2, the associative agents all die at almost the same time, while the dying ages of the calculus agents are much more gradual and spread out. The associative agents depend heavily on each other to survive, to the point where the whole population dies almost instantly if a part of the population (say, around 25%) dies. The calculus agents are able to fend for themselves a little longer when part of the population is gone.

In the first experiment this effect is present as well, but in a slightly different form. Here the calculus agents live much longer than the associative agents, but again the dying ages from the calculus agents are more spread out than those of the associative agents. The associative agents in this experiment form one big group, supporting each other until the whole group dies from lack of food in the area. The calculus agents do share food and form groups, but these groups are smaller and shifting often, so when part of the population dies the others can survive for much longer.

4.3 Bickerton

The next question is “What do these results mean for Bickerton's theory?”. In order to get a clearer picture of this, I will first return to the two research questions as formulated in the introduction.

4.3.1 Does social calculus increase performance?

This question is hard to answer from the results of this experiment. If we look only at the ages reached, then the conclusion seems to be that social calculus increases performance greatly in experiment 1, with the “following” behaviour, but social calculus decreases performance just as much in experiment 2, without the “following” behaviour. However, as explained in the previous sub chapters, the numbers in experiment 1 are influenced heavily by the different grouping behaviour of the two groups of agents. The numbers in experiment

2 are also a direct consequence of the difference in selectiveness between the two groups. In other words, in these experiments performance cannot be fully separated from behaviour. While keeping this in mind, the two traits that we have identified on which the groups differ can be related to performance:

Vulnerability is a trait that is usually not considered to be a good one. If a population cannot withstand some negative events and is unable to recover from them, it would go extinct very soon in this world. As discussed above, associative agents form a more vulnerable society than calculus agents in both experiments. Hence calculus agents do perform better.

Selectiveness may be less obvious in its implications. In this simulation, being more selective worked out to be a positive trait in experiment 1, but a negative trait in experiment 2.

Intuitively, it seems like a bad idea to keep sharing your food with others if they don't give anything back. A certain amount of distrust seems a very healthy and necessary trait to survive in a society that contains selfish agents. In experiment 2 however, the trusting strategy of “share with everyone, hope something comes back” that the associative agents adopted seems to work rather well for quite some time. An interesting point here is that the social score has no significant effect at all. A lower average social score means there are more “cheaters”, agents that won't share their food even if others have shared with them. However, the associative agents did not get significantly older in the high social factor condition. Apparently the number of cheaters in the world is not that important in this type of world and simulation.

It is important to note that effects like this one depend greatly on how the simulation is set up and what elements are present. The ability to detect cheaters is an important for any social species that has some kind of sharing system. Wilkinson (1988) even names it as one of the five requirements for any reciprocal altruism system. This is why we must be careful to generalize this success of the associative agents without further testing with different simulation setups.

4.3.2 Does social calculus affect behaviour?

This question has been answered in the previous section. The summary is that yes, social calculus does affect the behaviour of agents in the simulation. It makes them more selective in their choice of allies, both for grouping behaviour and for sharing food. It also makes the population a bit less vulnerable so that parts of the population can survive longer after some agents have died.

4.3.3 How does this translate to Bickerton's theory?

This experiment does not say anything about the use of agent, theme and goal for syntax, nor does it claim that monkeys must use these exact categories in order to survive. The first aim of this experiment was to see whether the theory can be implemented and what gaps need explaining. This goal was reached, meaning that the theory is concrete enough and contains no obvious gaps in the part that was implemented.

The second aim was to provide support for the theory. This is of course a very vague and broad term. The first step is that the results of this experiment do not disqualify the theory. If the calculus agents would perform worse across the board, or be no different from the associative agents except for an extra layer of abstraction (which is expensive energy-wise), then the assumption that social calculus evolved at all in early humans or primates becomes less likely. This was not the case in this experiment, so there is no reason to disqualify the theory. Then we can go one step beyond that, and ask whether this experiment provided any support for the theory. The answer to that question lies in the answers to the two research questions as discussed in the previous sections. Those answers are mostly situation dependent: in some situations social calculus provides an advantage, and in some it doesn't. All in all, this experiment provides some limited support, and further research will have to shed more light on the conditions that are needed for social calculus to provide an advantage.

4.4 Further research suggestions

Much of the effects that were found in this experiment depend heavily on the setup of the simulation. There are countless aspects that could be changed in order to gain new insights. A selection of the options that come to mind:

1. What happens when predators are introduced into the world? This probably asks for a more complex actions system, so that agents can also gain positive points for warning others when there are predators nearby, which would make it even more vital for agents to be close to others that are friendly enough to warn you.
2. What happens to the observed vulnerability of the associative agents in a population that has the ability to generate children and in which agents can die from old age? Does the effect persist in a more realistic, continuous population?
3. What is the effect of the population size? It seems that social calculus agents are better equipped to form subgroups, while associative agents try to interact with the entire population. The 60 agents that were used in this experiment is not a small group, but it is doable to have ties with 60 agents. It has been estimated that the maximum group size for a stable group in which each individual knows the position of other individuals and their relationship ties lies around 150 for humans, also known as “Dunbar's number” (Hill & Dunbar, 2003). For primates this number is lower (Dunbar, 1992). When the population grows beyond these numbers, calculus agents might be able to cope better.
4. What happens when a hierarchical social system is introduced? In this simulation agents were all equal, but most ape and monkey species that live in groups have a hierarchical system that dictates many of the options for social interactions. For example, the dominant individual is usually entitled to any food that he or she wants, meaning the lower placed individual has to give it up. On the other hand, these are also usually the strongest individuals, meaning they can provide protection against others. This complicates social interactions a lot, and the effects of social calculus may be different from what they were in a non-hierarchical population.

5 Conclusion

The experiments showed two differences between calculus agents and associative agents. The calculus agents are more selective in their choice of allies, and they are more robust as a population. These traits can be either an advantage or a disadvantage for the agents, depending on the exact circumstances. This means that the experiment provides partial support for Bickerton's theory. Further research will have to show whether these differences persist under different circumstances, and under what other circumstances agents will benefit from social calculus.

References

Bickerton, D. (2002). How protolanguage became language. In C. Knight, M. Studdert Kennedy and J. R. Hurford (eds), *The Evolutionary Emergence of Language*. Cambridge: Cambridge University Press, pp. 264-284.

Chomsky, N. (1981). Lectures on Government and Binding. Dordrecht: Foris Publications.

Christiansen, M. H., & Kirby, S. (2003). Language evolution: consensus and controversies. *Trends in cognitive science,* vol 7. Elsevier, pp. 300-307.

Dunbar, R. I. M. (1992). Neocortex Size as a Constraint on Group Size in Primates. *Journal of Human Evolution 22.* pp. 681-735

Hill, R. A., & Dunbar, R. I. M. (2003). Social network size in humans. *Human Nature, vol 14, no 1.* pp. 53-72

Kirby, S. (2002). Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In C. Knight, M. Studdert Kennedy and J. R. Hurford (eds), *The Evolutionary Emergence of Language*. Cambridge: Cambridge University Press, pp. 303-323.

Seyfarth, R. M., & Cheney, D. L. (2003). The structure of social knowledge in monkeys. In F. B. M. de Waal & P. L. Tyack (eds), *Animal social complexity.* Harvard University Press, pp. 207-229.

Trivers, R. L. (1971). The evolution of reciprocal altruism. *The quarterly review of biology, vol 46, no 1.* The University of Chicago Press, pp. 35-57.

Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Havlin, S., da Luz, M. G. E., Raposo, E. P., & Stanley, H. E. (2000). Lévy flights in random searches. *Physica A: Statistical Mechanics and its Applications, 282,* 1-12.

Viswanathan, G. M., Afanasyev, V., Buldyrev, S. V., Murphy, E. J., Prince, P. A., & Stanley, H. E. (1996). Lévy flight search patterns of wandering albatrosses. *Nature, 381,* 413-415.

Wilkinson, G. S. (1988). Reciprocal altruism in bats and other mammals. *Ethology and sociobiology, 9,* 85-100.

Zentall, T. R., Wassermann, E. A., Lazareva, O. F., Thompson, R. K. R., & Rattermann, M. J. (2008). Concept Learning in animals. *Comparative Cognition & Behaviour Reviews, vol 3,* 13-45.