

An Honest Invasion: A Study of Cooperative Behaviour

Inge Solheim Storakaas

Graduate Thesis

Trondheim, 2015

Supervisor: Robert Biegler

Norwegian University of Science and Technology
Faculty of Social Sciences and Technology Management
Department of Psychology

AN HONEST INVASION: A STUDY OF COOPERATION

Abstract

This thesis has investigated whether species other than humans would feel emotionally committed to honesty, and if they would feel this commitment towards other species. We have used data from a wide variety of sources, including comparative, behavioural, and neural studies, computational analyses, and mathematical and theoretical works. We have found that our neural circuitry predispose animals towards discounting temporally distant benefits, thus favouring cheating in public goods problems. In humans, this has been fixed with an emotional reaction against dishonest behaviour. We discuss several mechanisms which may have promoted cooperation in similarly advanced species, and also which factors could influence cooperation between different species. We conclude that cooperation is likely to have evolved in species advanced enough for civilisation, but that interspecies cooperation is likely to face significant challenges.

Preface

I would like to thank my wife Maren for breaking down the thesis in manageable parts, cleaning up the language, and providing invaluable support and solutions when Time seemed not my friend. She also provided some much needed social input, in short supply after long hours at the keyboard.

I would also like to thank my supervisor, Robert Biegler, for his endless contributions and clarifications, and seemingly inexhaustible supply of fresh research. His guidance was crucial in developing the thesis questions.

My dog, Garm, has made typing more enjoyable, if not more efficient, by routinely falling asleep upon my wrist. The Great Internet is also deserved of my thanks, for providing inspiration and distraction in (not always) equal measure.

Thanks also to my friends and family, for bearing with my convoluted, and often unsuccessful, attempts at describing what I was trying to work out. You see, it's a bit complicated...

Table of Contents

Abstract.....	i
Preface.....	iii
Table of Contents	v
List of Figures	vii
Introduction.....	1
Comparison	3
Between species	3
Between cultures	4
Game Theory	7
Equilibrium	7
Prisoner’s dilemma	7
Tit-for-Tat	9
Reputation	12
Spectators	12
Image scoring	13
Golden opportunities.....	14
Temporal discounting.....	15
Collection-risk.....	16
Opportunity-cost	17
Short-term rewards.....	17
Neural Coding	21
Ratio comparison	21

Neural representation	22
Generalisability across modalities	25
Logarithmic or linear coding?.....	27
Alternatives	29
The verdict	31
Commitment?.....	31
Rationality and cooperation.....	33
Generous TFT	33
Pavlov	33
Zero-determinant.....	34
Memory-one.....	35
Bounded rationality.....	37
Cooperation-inducing mechanisms	39
Kinship.....	39
Group augmentation.....	39
Punishment.....	39
Morality.....	40
Cooperation between species	43
Mutualism	43
Indirect reciprocity.....	43
Asymmetry.....	44
Communication barrier	45
Empathy	45
Cooperation?	46
Discussion.....	49

What if we landed on their planet?	49
Possible sources of error	51
Implications.....	52
Relevance	53
Conclusion	55
References	57

List of Figures

Table 1: Gain matrix for prisoner's dilemma	9
Table 2: Gain matrix for different strategies.....	10
Figure 1: TFT versus constant defectors over 50 rounds of interactions.....	11
Figure 2: TFT versus constant defectors over 10 rounds of interactions.....	11
Figure 3: Hyperbolic versus exponential discounting	15
Table 3: Grid of interactions between species	43

Introduction

Cooperation is an integral part of society, whether among humans, ants, or anything in between (Doebeli & Hauert, 2005). However, with cooperation arises the possibility of cheating, of not upholding one's end of the bargain. Cheating occurs in microorganisms (Crespi, 2001; Strassman, Zhu, & Queller, 2000), insects and plants (Dicke & Sabelis, 1992), jay birds (Stephens, McLinn, & Stevens, 2002), simple software programs (Yamada & Sakama, 2013), and, of course, humans (e.g. Frank, 1988), to name a few. Taken at a glance, cheating often seems to be more profitable for the individual in terms of short-term fitness, but at a cost to cooperative, and thus societal, function. How can cheating exist in, but not pervade, a species?

To investigate cheating and its place in nature, we ask the following questions, put somewhat to the extreme: If aliens stepped out of a saucer tomorrow, could we expect them to have an emotional commitment to honesty? And if yes, how confident could we be that they would feel this commitment towards us? In other words, how general is commitment to honesty in sufficiently advanced species, and could we expect this to apply between species and cultures?

To answer this, we first need to figure out how we can make any predictions about the behaviour of alien life forms. Countless movies, novels, and games have tried, but more often than not, it boils down to variations over human psychology, often used as a commentary on society. While we cannot fault writers their portrayal, we can try to answer the question without resorting to anthropomorphising. One approach that springs to mind would be to examine behaviour across species and cultures.

Comparison

Between species

A good place to start comparing behaviour is the closest relatives of humans, the simians. Brosnan et al. (2011) found that economic behaviour, highly relevant for cooperation, was surprisingly similar across humans, chimpanzees, and capuchin monkeys. They found that the differences between the groups were due mostly to the proportion of individuals who used similar decision-making processes. This means that there were rational capuchin monkeys, but fewer than rational chimpanzees, and even fewer compared to humans. This means that economic/rational behaviour exists on a continuum in and between species, and not just in humans.

Wrangham and Peterson's 1996 book, *Demonic Males*, traced violence in humans all the way back to the great apes (orangutans, gorillas, chimpanzees and bonobos), and argued that humans are violent because evolution has selected for violence. However, using the relatively peaceful Bonobo as evidence, the authors show that our shared genetic heritage did not necessarily doom us to violence. While their findings discuss violence, one would assume the same holds for economic conduct as well, that related species does not necessarily share behaviour. Similarly, Stevens, Hallinan, and Hauser (2005) found a striking differences in the ability to discount future payoff for a smaller, immediate payoff differed in two closely related primates (common marmoset and cotton-top tamarin). The authors argued that this is because of the animals' different environments.

For our argument, one of the lessons to take from *Demonic Males* and Stevens, Hallinan, and Hauser's paper is that while ancestral genetic makeup undoubtedly affects the psyche, significant variations occur in closely related species, and that the environment is a powerful factor in even fundamental cognition. This warns us that comparison across species

might be an inconclusive approach. This is not to say that the approach is useless, but rather that additional methodologies are needed to provide supplementary evidence. How about comparison within species? Let us look at culture.

Between cultures

Before the British colonisation, the Nayar of Kerala, India, led unconventional family lives. The men traditionally worked as mercenaries, which meant that they would be away for long periods at a time, if they returned at all. While the men usually married, it was both common and accepted for the wives to take lovers, making it difficult, if not impossible, to tell if their wife's children were their own. In a very pragmatic tradition, the men would invest in their sisters' offspring rather than their wives', that is, investing in the highest average genetic relatedness. When the British unified India, regional wars abated, and the need for mercenaries declined. Within a generation, the mating system changed to more conventional norms (Gough, 1961).

The Nayar were an exception to the rule of patriliney, inheritance via the male line. In their case, this was because of paternal uncertainty. There are other examples of matriliney, or female line inheritance, although they are vanishing with increasing economic development (Holden, Sear, & Mace, 2002). However, these examples do not rely solely on paternal uncertainty, but rather on whether the resources benefit sons or daughters the most. In horticultural societies, where matrilineality is prevalent, raiding is not a great concern. This means that strong fighters unencumbered by children are not in great demand, and sons and daughters are valued the same. Since the paternal linear is uncertain compared to the maternal, these societies tend to practice matrilineal inheritance. When societies turn to cattle farming, both raiding and defence from raiding becomes an issue, so sons in general becomes more

highly valued. This offsets the paternal uncertainty, and patrilineal or mixed decent becomes the norm (Holden & Mace, 2003).

While not directly related to cooperation, the traditional Nayar mating pattern shows us that, when compared to the traditional western or modern Nayar pattern, significant differences in behaviour can occur within a species. The same is true for tribes practicing matrilineal inheritance. Although this variability does not bode well for generalisation, it is interesting that these patterns can be predicted using game theory. If cultural difference can be predicted mathematically, then we would have a much more robust base upon which to build our case. Let us apply game theory to cooperation.

Game Theory

Game theory is the mathematical exploration of rational decision-making. It was originally developed by John von Neumann and Oskar Morgenstern in 1944 as a means to predict trends and behaviours in economics. Game theorists looked at specific situations, called games, where agents (known as players) would interact to gain some sort of reward. The agents can play one or several different strategies, that is, behave in one way or another. Game theorists analyse the situation to find the optimal strategy in each game (Stanford Encyclopedia of Philosophy).

Evolutionary Game Theory (EGT) was developed somewhat later. Maynard Smith wrote *Evolution and the theory of games* in 1982, which is considered the seminal work of EGT. EGT applies game theory in a biological setting, but with several unique qualifiers. First, the agents are genes, not individual animals. Second, a strategy isn't chosen by the agent, it evolves by trial and error over many generations (Stanford Encyclopedia of Philosophy)

Equilibrium

In any game, there are different outcomes. Some outcomes favour one player over the other, others distribute payoff equally. Some outcomes are only possible if both players play optimally, but deciding optimal play is not just looking at the best possible outcome. A Nash Equilibrium is when neither player in a game have an incentive to change the strategy they are currently employing, given their opponent's choice of strategy. This means that if you think your opponent is playing the best response to your strategy, then you should play the best response against his strategy.

Prisoner's dilemma

Common (or public) goods problems are a category of games where all players, if only interested in maximising their economic benefit, reaps the greatest profit by defecting, i.e.

refusing to contribute to the common good, taking more than their share, or outright theft. However, if players attempt to maximise profit, they end up with less than if they had cooperated. Suppose there are ten families that share enough land to sustain ten cows. If one family adds another cow, all eleven cows will be slightly undernourished, yet the family with two cows does better by having two-elevenths, compared to the other families' one-elevenths. If the other families follow suit by adding their own extra cow, then they are all worse off than if they only had one cow each. Still, any family that goes back to keeping one cow will be even worse off. Other real world examples of this includes overfishing or generating power in a way that hurts the planet. (R. Biegler, personal communication, April 30, 2015). The prisoner's dilemma is a widely researched two player version of this game. Frank (1988) has explained the scenario like this:

Smith and Jones run a restaurant together, Smith being the cook and Jones being the manager. Since Jones can't cook, Smith can easily use second rate ingredients and pocket the difference. Likewise, Jones can siphon off money as "management expenditures", with Smith being none the wiser. If they both cooperate, the restaurant goes well, and they both get 3 monetary units. If one of them cheats, the cheater gets 5 units and the honest one gets zero. If they both cheat, they both get 1 units, because the restaurant isn't going so well. We assume they only agreed to work together for one season (one round of the game).

Smith benefits from cheating, no matter whether Jones cheats or cooperates. So if Jones cooperates, Smith gets 5 for cheating, but only 3 for cooperating. If Jones cheats, Smith gets 1 for cheating, but only 0 for cooperating. No matter what Jones does, Smith gets more for cheating. The same is true for Jones, so both should cheat if they want to maximise profit, yet then they end up getting 1 each instead of 3 each. This makes cheating the Nash equilibrium if the game is played only once.

		Smith	
		Defect	Cooperate
Jones	Defect	Both: 1	Jones: 5 Smith: 0
	Cooperate	Smith: 5 Jones: 0	Both: 3

Table 1: Gain matrix for prisoner's dilemma.

Prisoner's dilemma can exist either as a one-shot game, where the players only interact once, or as an iterated game, where the players interact repeatedly. The iterated version allows us to look at long-term gains of strategies. Additionally, iteration offers us the possibility of cooperation as an adaptive strategy, compared to the depressive mutual defection that is the Nash equilibrium of one-shot game.

Tit-for-Tat

There have been suggested several different strategies for optimal play in iterated prisoner's dilemma. The arguably most famous strategy, Tit-for-Tat, was presented by Anatol Rapoport in an iterated prisoner's dilemma tournament held by Robert Axelrod (Axelrod, 1980). The contestants wrote a computer program, which was paired against the other contestants programs. TFT won the tournament by getting the most points, but it was also remarkable for having very few and simple instructions compared to the other contestants. The strategy goes as follows: Start by cooperating, then copy the opponent's choice for each consecutive round. In other words, if I defect on you, you defect on me until I start cooperating. This ensures that both players get approximately the same score.

Let us run some simulations. We use the same payoff matrix as earlier (Table 1) and set the number of interactions to 100, then compare outcomes for the different strategies.

		Player 2		
		TFT	D	C
Player 1	TFT	1: 300 2: 300	-	-
	D	1: 104 2: 99	1: 100 2: 100	-
	C	1: 300 2: 300	1: 0 2: 500	1: 300 2: 300

Table 2: Gain matrix for different strategies.

TFT = Tit-for-Tat, D = Always Defect, C = Always Cooperate

We see here that while TFT loses by a slight amount when playing with a constant defector (because of the first round), the benefits of using TFT is evident when playing either against TFT or a constant cooperator. As in the one-shot game, defectors have a solid advantage over cooperators.

It seems like the iterated game is very similar to the one-shot, in that defectors and cooperators receive the same payoffs as in one-shot games. What makes iterated games special is that they allow for memory, which is crucial to TFT. TFT only needs memory for what another player did when encountered the last time, because TFT simply copies the other's behaviour from last time but ignores anything in the more distant past.

To simulate the invasion of a strategy into a population, we can graph the percentage of TFT and constant defectors in a population. If we also specify how many interactions they would have with each other, an important feature of TFT appears.

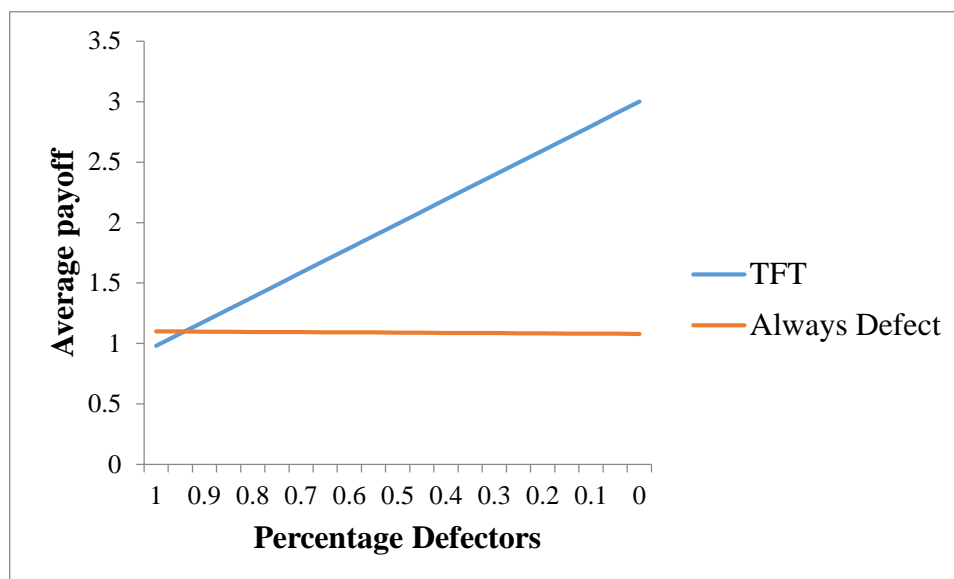


Figure 1: TFT versus constant defectors over 50 rounds of interactions, and with different ratios of defectors in the population.

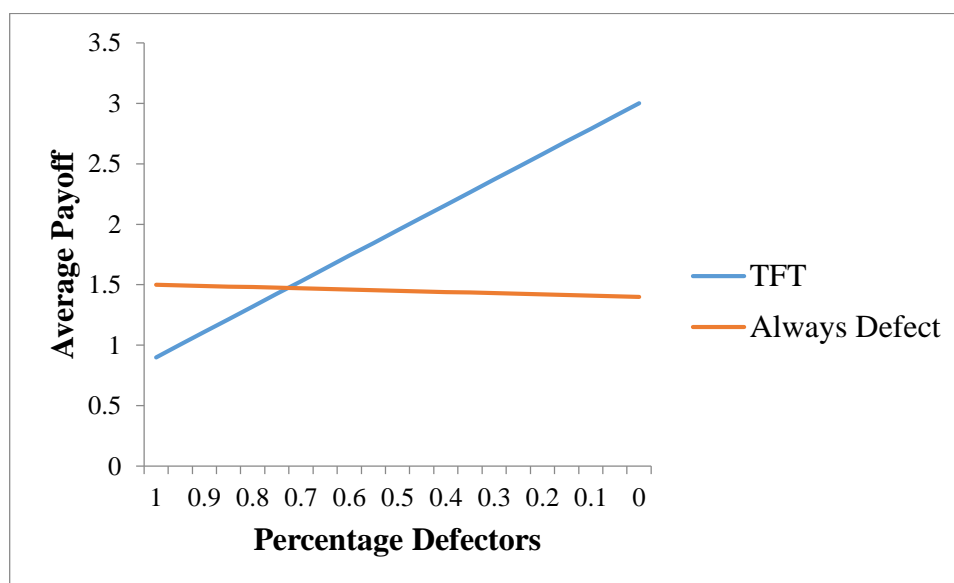


Figure 2: Same as figure 1, but with 10 rounds of interactions.

As we can see, the fewer rounds of interactions there are, the harder it is for TFT to invade a population of defectors. TFT loses out in the first round, and fewer interactions before switching means more first rounds. If players using TFT could know beforehand if they were

up against a defector, then this would make TFT a much stronger strategy. We need a way to recognise individuals and remember their actions. We need reputation.

Reputation

We have seen that TFT only solidly beats constant defectors when there are numerous interactions before a switch. Another way in which TFT could gain the upper hand, is if a player's reputation precedes them. Of course, TFT with reputation is not simple TFT anymore, since we are expanding the game into more complex territories. Reputation gives defectors a penalty, in that it increases the chance that future interactions will result in low payoffs. If they are known defectors, then players should defect automatically. If we also add the option of avoiding interaction with unsavoury characters altogether, and assuming avoidance does not come at the cost of losing out on an interaction, then that would make TFT even more outstanding. Reputation gives TFT its edge, in that not only will they always play fair, they will also punish defection by reciprocating the move.

Spectators

In order for reputation to favour cooperation, however, it is required that other people see how nice or naughty you have been. This is usually not a problem, as your co-player presumably would spread the word to their group, or there would be spectators of some kind. There is evidence that spectators influence actions among less advanced species. Dzieweczynski and colleagues found that Siamese fighting fish will alter their behaviour dependent on the audience (Dzieweczynski, Earley, Green, & Rowland, 2005). Baltz and Clark (1997) found that male budgerigars, a socially monogamous bird, would refrain from cheating on their partners if there were flock-mates present, but do so when there was no-one to watch the infidelity.

An interesting discovery was made by Bateson, Nettle and Roberts (2006), who found that pictures of eyes would increase honesty compared to pictures of other things. They set up the experiment at an “honesty box” in a university coffee room, in essence an unmanned hot drinks stations with a price list and donation box, and no-one to watch whether the staff and PhD students actually paid for the drinks. The researchers found that when pictures of eyes that looked directly at the participants were displayed, they paid almost three times as much as when pictures of flowers were displayed. Such results indicate that reputation does matter.

Image scoring

Related to reputation is image scoring, a type of indirect reciprocity. This is a proposed mechanism that allows organisms to cooperate in one shot games without being worse off than if it had defected (Nowak and Sigmund, 1998). The idea is that you have an image, or a reputation, and that not cheating increases the value of your image. Someone who uses image scoring will cooperate with another player if that player’s image score, the proportion of past interactions in which that player was known to cooperate with others, exceeds a threshold value. As a consequence, even if the image scorer is exploited, those who observed the interaction will increase their image score for that person and be more likely to help in future.

The evolutionary value of image scoring is greatly increased by language, which facilitates transmission of your image beyond eye-witnesses. While Nowak and Sigmund only performed computer simulations, Wedekind and Milinski (2000) found the same effect in live humans. It also appears that image scoring, or something similar, can exist in more primitive species. Pinto, Oates, Grutter, and Bshary (2011) have observed that cleaner fish cheat less in the presence of potential clients, and that clients shun cleaner fish that cheats.

Implementing reputation seems like a magic bullet for defection. Why cheat at all, when the long-term costs are so high? First of all, while TFT only requires an exceedingly simple

behavioural blueprint in theory, this may not be the case for the mechanisms of reputation. Reputation requires memory, individual recognition, and either spectators or communication. Of these, individual recognition is presumably the latest development, as it requires both communication and memory to function. Individual recognition, while fairly common in vertebrates, is rare among invertebrate species (Caldwell, 1986), suggesting it is a relatively advanced feature (Tibbetts & Dale, 2007), or at least more rare than cooperation, which occurs in microorganisms (e.g. Crespi, 2001). So, while invertebrates could still find cheating a viable option, reputation may even the balance in evolutionary younger species.

Golden opportunities

Humans are a curious species with regard to cheating. Frank (1988) brought up the puzzling fact that most of us will not cheat even when there is no-one to watch us. His example is picking up an abandoned wallet full of cash in a deserted park. To use our already established terms, taking it counts as defection, giving it to the police is cooperation, receiving a reward is co-player cooperation, and leaving empty-handed is co-player defection. Game theory predicts that the only rational choice in this situation is to defect, after all, the potential reward for turning in the wallet is neither immediate nor as good as taking the contents. All the same, the wallet is usually ignored or returned (Frank, 1988). Indirect reciprocity cannot account for this choice, as there are no witnesses, and so no incentive for cooperation.

Emotional commitment to honesty. Frank (1988) have proposed a mechanism for why this happens. He called it emotional commitment to honesty, and it is basically an extra layer of protection against cheating. When the temptation of a large, immediate payoff rears its ugly head, an emotional reaction, a feeling of shame or guilt, quickly balances the scales, helping us focus on the long-term benefit. To understand why we need this mechanism, however, we need to look at temporal discounting.

Temporal discounting

It is a curious fact of animal cognition that short-term rewards are much more potent than equal, or even greater, long-term rewards. This preference for immediate rewards compared to temporally distant ones is called temporal discounting, and lack of preference follows a steep hyperbolic curve (Green, Myerson & McFadden, 1997). This means that as the reward's delay increases, the value of the reward steeply drops, until it gradually levels off as the delay gets longer. This is similar, but not identical to an exponential curve.

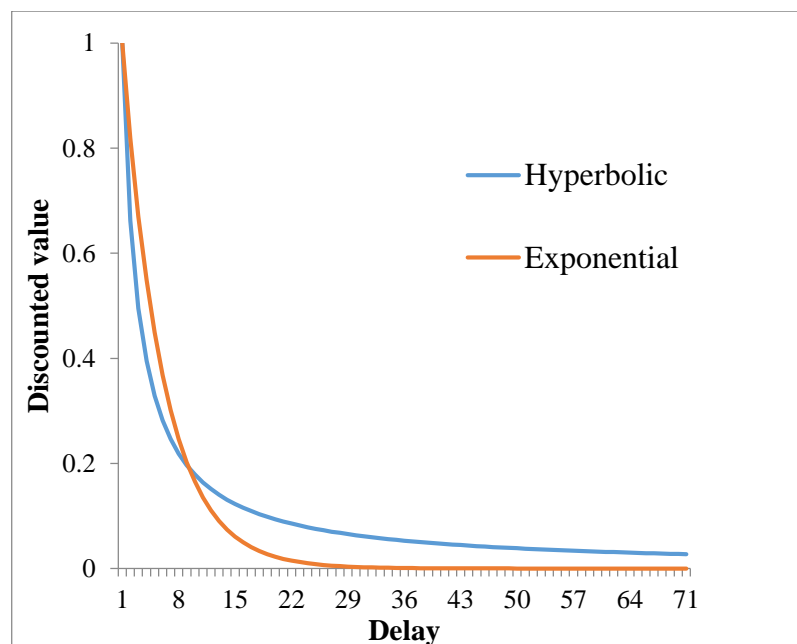


Figure 3: *Hyperbolic versus exponential discounting.*

We see plenty of examples of temporal discounting in nature, even among relatively advanced species. Dufour, Pelé, Sterck and Thierry (2007) found that chimpanzees could postpone a reward for 4 minutes if the delayed reward item was 2 to 8 times larger than the immediate reward. They could also postpone the reward for 8 minutes if it was 40 times larger. Similar experiments have been done to capuchin monkeys (Ramseyer, Dufour, Pelé, Chauvin

& Thierry, 2005) and preschool human children (Mischel, Shoda & Rodriguez, 1989), with human children doing comparably well, about 5 minutes maximum reward delay, and the capuchins doing as much as 24 times worse than the chimpanzees. Rachlin, Raineri and Cross (1991) found that human adults, on the other hand, can delay rewards for months or even years.

Multiple explanations for temporal discounting have been presented. These explanations can be related to this thesis through game theory. However, it is important to note that temporal discounting is not exclusive to this field. Temporal discounting is relevant for the behavioural sciences as a whole, and it is not required that there are multiple agents, like in games. We will be using game theoretic examples, since this is more relevant to the thesis as a whole.

As mentioned, there are several different explanations for temporal discounting. We will present these based on Stephens' (2002) two categories: Collection-risk discounting and opportunity-cost discounting.

Collection-risk

This first class of explanation holds that delayed rewards are less valuable because delay reduces the likelihood that the reward is obtained. This can further be divided into two subtypes, if one supposes a sequence of reward collections. The subtypes are identical if one only considers a single collection event. Termination risk is where something might occur that terminates the sequence before the delayed reward can be collected, and the sequence cannot be resumed. In cooperative game terms, this could entail something happening that prevents collecting the benefits of cooperation, making cheating, i.e. getting the immediate reward, a safer option.

Interruptions risk is closely related, but here the game sequence is not terminated, so you only have the risk of losing one reward. There is still the opportunity of getting the next

reward in the sequence. It differs from termination risk in that the games continue even if one player cheats, there is no punishment other than maybe reducing the cheated player's incentive to play nice. Collection-risk could mediate how much consecutive games are valued in scenarios with or without punishment or reputation.

Opportunity-cost

This class deals with scenarios where waiting has a cost associated with loss of opportunity elsewhere. One type, called rate cost of delay, asserts that since you cannot utilise two sources of benefit at once, then you lose or gain the difference in benefit between the sources. This can be exemplified by a predator waiting in ambush at a certain spot. This prevents the predator from trying another ambush spot if the first one is unproductive, since it cannot wait in two places at the same time. Likewise, playing with a defector means you cannot be playing with a cooperator, and thus lose the difference.

The second type embodies investment, where waiting for a better opportunity prevents you from utilising and increasing resources now. This can be exemplified by an animal conserving energy by forgoing smaller prey for something larger. Hunting the smaller prey might tire you to the point where the larger prey can escape, but waiting for the larger prey will also drain energy, albeit at a slower pace. From a cooperative game perspective, this would entail turning down potential co-players, i.e. known defectors, in favour of the right type of player, i.e. known cooperators. Not playing is worse than playing, as you can always guarantee a small payoff.

Short-term rewards

Temporal discounting will often lead to smaller rewards being chosen, even when the long-term average payoff of choosing the larger reward is greater. In cooperation situations, cheating is this smaller, immediate reward, while cooperating is the larger, but distant, reward.

We can exemplify this with reputation. In populations where reputation is present, cheating gives you less chance of interacting with someone after you have cheated them, thus less potential long-term payoff. However, the initial round gives you a large payoff if your partner cooperates. If you discount the future payoff from repeated interactions, and likewise discount the trouble cheating is likely to cause you in the future, then cheating becomes an overwhelmingly tempting option. This way, temporal discounting makes cooperation difficult, as the long-term rewards of cooperation, and costs of cheating, are reduced to negligible amounts. If temporal discounting rules, cheating should prevail.

Stephens, McLinn, and Stevens (2002) tried to reduce temporal discounting in cooperation games with blue jays. They found that, compared to conditions with normal discounting, reduced discounting promoted cooperation to a high degree. In fact, cooperation did not match the rate prescribed by conventional strategies, but exceeded it. The birds forgave defection close to 50 % of the time, compared to the 10 % of Generous TFT, a very good “forgiving” strategy in simulations (Nowak & Sigmund, 1992).

Stephens (2002) have suggested that animals’ tendency to choose short-term rewards over long-term rewards might not be a bad choice after all. Long-term rewards are more obscured by random events than are short-term rewards, making it harder to be sure of the actual size of the reward. Put differently, the animal can be surer of actually getting the short-term reward, since a lot of things might happen between now and the future. However, in a cooperation game, and resource gathering in general, this myopic focus may not be as beneficial. In fact, Stephens et al. (2002) found no experimental evidence of animal cooperation in prisoner’s dilemma up to date.

Fixing the problem. Despite some positive features, temporal discounting is not perfect. If it was perfectly tuned to the environment in such a way that the best long-term payoff

were chosen, then we would have no need for an emotional commitment to honesty (Frank, 1988). Yet we do need it, and so temporal discounting must be flawed in some way. If the only function of emotional commitment to honesty is to enable cooperation, then we would term it a patch, or quick fix. There is precedent for evolutionary quick fixes in nature, as, for example, the selection for sickle cells in response to malaria can attest to (Wade, 2006). On the other hand, why would evolution not fix the problem by removing temporal discounting instead? One answer to this could be that temporal discounting is a manifestation of a more fundamental mechanism.

Neural Coding

Ratio comparison

Stephens (2002) have argued that temporal discounting reflects the way the brain compares and discriminates magnitudes. Ratio comparison appears to be preferred in neural systems (Stephens, 2002; Gibbon, Malapani, Dale, & Gallistel, 1997). The alternative to ratio comparison would be difference comparison, but Stephens (2002) have shown mathematically that, unlike ratio comparison, difference comparison is insensitive to time. We can explain it like this:

If animals compared different sources of food by subtractively comparing the time it would take to obtain it (e.g. 12 seconds is 2 seconds longer than 10 seconds), then the amount of time compared would not matter. The difference between 10 and 12 seconds would be the same as the difference between 100 and 102 seconds, and discrimination would be just as easy in both cases. However, if the animal uses the ratio between time intervals to discriminate, then the difference between 10 and 12 would be 1.2, and 1.02 between 100 and 102. We see experimentally that if the magnitude of the time intervals compared is large, then animals cannot accurately judge the difference between them, unlike the same interval of smaller magnitudes.

Stephens (2002) proposed that the tendency towards ratio comparison might either reflect a mechanistic restraint, or an unknown adaptive advantage. The two most influential theories on neural coding, logarithmic and linear coding, offer different insights on this. Logarithmic coding holds that ratio comparison happens because stimuli is coded logarithmically, which makes values at larger magnitudes less distinct. Linear coding, on the other hand, assumes the differentiation problems happens as a result of noise in the system. This is important, as the difference between logarithms is a ratio, making ratio comparison a

direct consequence of the neural coding scheme. Linear coding, on the other hand, can be subtractively compared, making ratio comparison not the only option. Since temporal discounting appears to be based on ratio comparison, then an evolutionary workaround could be to use difference comparison to avoid discounting future rewards.

Neural representation

The brain encounters large dynamic ranges every day. It needs to represent millimetres and kilometres, milliseconds and hours, tones ranging from 10 to 20 000 Hertz, and so on. One way to solve this problem could be by having a single neuron represent each distinguishable stimulus, much in the same way that the heavily criticised “grandmother cell” theories have proposed (e.g. Bower, 2009). There are, however, alternative approaches to neural representation of stimuli.

The Weber-Fechner law. In the middle of the 19th century, Ernst Weber discovered that the ability to differentiate between two magnitudes depended on their intensity. If the magnitudes were small, e.g. two unequal feathers, he found that subjects could easily distinguish between their weights. If the magnitudes were larger, however, greater difference between the objects were needed to make differentiating possible. His student, Gustav Fechner, later developed this relationship into a logarithmic equation, $P = k \ln(s)$, where P is perceived intensity, k is a constant (the Weber fraction), and s is the stimulus intensity (Dehaene, 2003). This equation is known today as the Weber-Fechner law.

A century later, Stevens (1961) would propose that the perceived intensity was a power function of the stimulus, rather than the logarithm. However, MacKay (1963) would quickly claim that Steven’s law and the Weber-Fechner law is indistinguishable without examining the underlying architecture. In the same year, Mountcastle and colleagues would argue that psychophysical laws are linear (Mountcastle, Poggio & Werner, 1963). This set the scene for

a debate that rages to this day, whether neural representation of stimulus is linear, or logarithmic. There are variations over these themes, but we will for the present purpose examine them in a more general way.

Logarithmic. Logarithmic representation means that each neuron fires in response to a specific range of numbers (or frequencies, temperatures, pressure, etc.). For example, in a base 10-system, neuron 1 fires when the stimulus is in the range of 1-9, neuron 2 fires during 10-99, The higher the number, the larger the range of the firing neuron, which equals less precision. This would make the Weber-Fechner law a feature, as a limited range of neurons can code for a large range of objective stimuli (Gallistel, 2011).

Linear. In a linear coding scheme, neurons encode stimuli as linearly proportional to the subjective representation. The imprecision embodied by the Weber-Fechner law is explained as noise introduced at a later stage (e.g. in memory encoding or retrieval), thus making the law an unintended consequence, a bug. A linear system would need some clever mechanism to encode large dynamic ranges, as the amount of dedicated neurons needed would be enormous (one per discrete quantity) (Gallistel, 2011).

Logarithmic coding does not have a specific, unifying model associated with it. Dehaene and Changeux (1993) have proposed a model, and Sun, Wang, Goyal, and Varshney (2012) have made a computational analysis, but on the whole, proponents of log coding appear to have mostly focused on whether data fit a logarithmic or linear plot, rather than theory-crafting. Linear coding, on the other hand, is primarily represented by Scalar Expectancy Theory.

Scalar expectancy theory. A prominent theory in linear coding is Scalar Expectancy Theory (SET). Proposed by Gibbon (1977), SET postulates an internal clock mechanism, containing a pace- maker which generates pulses, an accumulator which counts these pulses,

and long-term memory, which saves the counts when the interval has ended (e.g. reward received). There is also a decision mechanism, which compares current time with remembered time, and identifies an appropriate response outcome (Gibbon et al., 1997). SET has scalar properties, meaning that the standard deviation of time, or noise, grow as a constant fraction of the mean.

Accumulator models such as SET function like a stopwatch, keeping time from the start of an interval to the reward at the end. A problem with this is how the animal knows when the interval starts. It would require an absurdly large number of “stopwatches” to record the beginning of every potentially interesting interval (Kheifets, 2014).

Time-left. The time-left procedure has been an important experiment for proponents of SET (e.g. Wearden and Jones, 2007; Gibbon and Church, 1981)). The procedure consists of two levers, one which gives rewards every X (e.g. 60) seconds, and one every Y (e.g. 30) seconds. The lever with the most frequent rewards (Y) are only available after a certain time, T, which is varied by the experimenter. The experiment aims to discover when the animal thinks that both levers are equally likely to produce the reward when pressed, which is expected to measure the animal’s subjective time measurement.

Gibbon and Church (1981) used the time-left procedure to find evidence of temporal subtraction in rats and pigeons, and used these findings to argue that subjective time is represented linearly, not logarithmically. However, the experimental setup used, the time-left procedure, has been criticised by Cerutti and Staddon (2004) for being too complex to settle the question. They proposed that the temporal control hypothesis (a variant of temporal discounting) provides a simpler explanation of the results. Castro and Machado (2010) used an improved version of the time-left procedure and found that the results are compatible with logarithmic representation, but also with linear representation with bias. It would seem the

time-left procedure is unable to provide conclusive evidence for either coding scheme, and so cannot be used to decide which of them is more likely.

Source of variability. Logarithmic coding has variability built in at the time when the external input is received and translated into neural code. The variability is inherent in the coding from the very beginning. SET and other scalar models (Gibbon, Church & Meck, 1984; Gallistel & Gelman, 2000; Brannon et al., 2001) place the source of variability in memory, rather than stimulus encoding. However, Allan (2002) and Rodríguez-Gironés and Kacelnik (2001) claimed that memory is an unlikely candidate for variability.

Generalisability across modalities

The research on linear versus logarithmic representation is spread out across many sensory modalities. If these modalities could be shown to have a collective underlying mechanism, then we would have a large amount of data to compare, and perhaps finally have the means to put the debate to rest.

Space and time are closely related. Physicists treat time as a fourth dimension to the three spatial dimensions we normally perceive. Quantities are non-abstract numbers that are given to sets that exist in space (and time). Stevens (1960) coined the term “prothetic” to mean dimensions that can be experienced as “more than” or “less than”. There are, as we will see, grounds to assume that these quantities are closely linked in the mind as well, and that they may share a common mechanism.

Gallistel (2011) have argued that time, space, and number are interrelated. Representation of rate, for example, is numerosity divided by duration, and spatial density is numerosity divided by area. Thus, it is plausible that the system used to represent these magnitudes share a common mechanism.

Walsh (2003) has proposed a unifying theory of magnitudes (ATOM), where perception of time, numbers, and space are based on a common system. Walsh's model has space, time, and quantity all be computed by a single metric. ATOM does not assume either clock or non-clock explanations for underlying mechanisms (Buetti & Walsh, 2009), and is therefore valid for both logarithmic and linear representation schemes.

Staddon and Higa (1999) have proposed a model, multiple-time-scale (MTS), which explains the close relation between timing and counting. Staddon and Higa (2006) argued that the MTS model does not postulate any specialised timing process, thus having the advantage of fewer assumptions.

Brannon, Suanda, and Libertus (2007) have found experimental evidence that argues for a parallel development of time and number representation in children. This makes it likely that they share an underlying mechanism, or at least are closely related. Research on developmental dyscalculia also support this connection, but there appears to be a degree of independence as well (Cappelletti, Freeman & Butterworth, 2011). Skagerlund & Träff (2014) argued that their data on dyscalculia goes well with ATOM (Walsh, 2003).

Mackevicius, Best, Saal, and Bensmaia (2012) have argued that tactile, vibrissal, and auditory systems all rely on temporal information and are all similar in function.

This is not to say that all aspects of the world are encoded similarly. For example, much research has been done on grid cells and their role in encoding position (Moser et al., 2014), but grid cells might be working on a higher level of abstraction than time, space, and number perception. Buetti and Walsh (2009) have made it clear that the ATOM model does not support all aspects of time and space, and cited navigation as one of the exceptions.

Nieder and Miller (2003) have presented data that supports the analog coding hypothesis, the view that both higher-level cognitive functions and lower-level sensory functions have fundamentally similar representations, and should exhibit similar attributes. Nieder and Miller's results also point to both cognitive and sensory systems having the same underlying architecture and coding system. This is plausible, as evolution would seek to reuse designs (Nieder & Miller, 2003).

While not conclusive, these data suggest rather than confirm a common underlying mechanism to at least the basic dimensions of time, space, and number. This gives us a broader foundation upon which to build our case.

Logarithmic or linear coding?

Numerosity. Neural representation of numbers have received much attention in the logarithmic versus linear debate. A very interesting pattern emerges when looking at representation of numbers in different populations. In young children, uneducated adults, and non-human animals, numbers appear to be represented logarithmically (Nieder & Miller, 2003; Dehaene, 2003; Dehaene, Izard, Spelke, & Pica, 2008; Roberts, 2006). In educated adults and older children, however, representation appears to be linear (Sielger & Booth, 2004; Dehaene et al., 2008). Linear representation of numbers become gradually more prominent as children age, and for a while, both logarithmic and linear representation are present (Lasik & Siegler, 2007). Note that by "educated" we mean "educated according to western standards". This change in number representation has been attributed to culture (Dehaene et al., 2008; Ansari, 2008).

While linear representation of numbers are found among humans, it appears to be a later addition via invention or culture. Evolution appears to favour logarithmic coding of this aspect of space-time.

Time. Roberts (2006) have found that pigeons' time scale support a logarithmic framework. Yi (2009) criticised Roberts' conclusion on the grounds that they made unnecessary assumptions, but supported logarithmic coding in rats, using signal detection theory. Signal detection theory assumes only that the strength of sensory and cognitive events are continuously variable (Macmillan, 2002).

Wearden and Jones (2007) have found time perception to be linear in educated adults. They used the time-left procedure, which has been criticised earlier. They claim, however, that the criticism do not apply to their version of the procedure. Yi (2009) also criticised Wearden and Jones, and claimed that their averaging process made their results less valid. Cordes, King, and Gallistel (2007) also used the time-left procedure when they found evidence of linear representation in mice. This may or may not discount the results entirely, but does at the very least cast some doubt as to their validity.

Van Rijn and Taatgen (2008) argued for a logarithmic representation of time, with a single time-keeping mechanism. This is interesting because other pace-maker models usually support linear time (e.g. Mech and Church, 1983). Van Rijn and Taatgen (2008) also found that strategic use of arithmetic negated the need for more clocks. However, while Crystal (2001) seem to agree with the logarithmic representation, he argued that the underlying mechanism must be based on oscillators, not pace-makers. Crystal (2006) have presented an oscillator model for discrimination of temporal intervals, ranging from milliseconds to days.

While no consensus has been reach regarding the underlying mechanism for timing, be it clocks or oscillators, it seems that logarithmic representation is a viable coding scheme regardless of mechanism, and that the majority of authors support this.

Other modalities. Johnson, Hsiao, and Yoshioka (2002) have used data from texture perception to argue that linearity, rather than logarithmicity, is the basic law of psychophysics.

However, Shirado and Maeno (2005) found that human texture perception has a logarithmic relationship with objective stimuli. Nieder and Miller (2003) did not see the conclusion of Johnson et al. (2002) as incompatible with logarithmic coding. After the initial logarithmic encoding, signals could be communicated in a linear fashion throughout the brain.

Sun et al. (2012) have shown that sound perception follows a logarithmic scale. This matches how we denote sound intensity, as the decibel scale is logarithmic (Encyclopædia Britannica). While decibel notation might be logarithmic primarily out of convenience and clarity, it is a potentially interesting parallel.

Varshney and Sun (2013) claim that, while the majority of sensory representation appear to be logarithmic, there are exceptions, night vision being one of them. Research into deviations from the norm could prove to give us a broader understanding of the strengths and weaknesses of different sensory coding schemes.

Computational analysis. In addition to examining how the empirical data fit either representation, some authors have approached the problem with mathematical modelling. Portugal and Svaiter (2011) have formally shown that the logarithmic scale provides the minimal worst-case relative error. They assume that relative error is more important in nature than absolute error, and thus that the system that minimises relative error is biologically most feasible. Sun et al. (2012) argue that Portugal and Svaiter's (2011) focus on worst case error is flawed, and that expected relative error is more biologically plausible. However, while disagreeing with the method, Sun et al. also found that logarithmic and related representations are optimal for evolution.

Alternatives

We have primarily focused on logarithmic versus linear coding thus far. There are other proposed mechanisms, but since existing research mainly discuss linear or logarithmic systems.

It is beyond the scope of this paper to compare all the contending mechanisms at the level of scrutiny required to reach a definitive answer, but we will mention these newcomers in the hopes that more thorough investigations can be instigated.

Autoscaling. Gallistel (2011) has proposed an alternative scheme for the representation of magnitudes. Autoscaling is a mechanism where sensitivity is automatically adjusted based on the strength of the input, so that weak input is more accurately measured than stronger input. This is a feature we find in many measuring instruments. An advantage of this is that the output signal strength is independent of input strength, thus avoiding sensory overload. This mechanism would make the Weber-Fechner law a feature, not a bug.

Modulo coding. Modulo representation is a number representation system where numbers “wrap around” when reaching a certain value. An example would be a 12 hour clock, which, upon reaching 12, starts again at 0. Adding 5 to 9 in this system would result in 2, not 14. This clock is modulo 12.

Fiete, Burak, and Brookings (2008) compared different representation schemes in terms of capacity, ability to perform algorithmic operation without a look-up table, and more. They found modulo coding to be theoretically superior to the alternatives. However, the system lacks the ability to directly compare different magnitudes. A possible solution presented by Fiete et al. (2008) involves a mechanism that converts modulo phase to metric representation. The authors have constructed a neural network scheme that converts modulo code into linear code, but were troubled by the number of neurons required for the linear output. They speculated whether a similar conversion scheme could produce a more efficient output, like logarithmic code, but they did not present one. Since ratio comparison appears to be a fundamental process in the brain, this conversion might be necessary. This does add a step that other coding schemes avoid, giving some doubt to its value. However, modulo coding has other valuable features,

and there is no reason to believe that the brain is one hundred percent optimised in every aspect. More research is required on this promising framework.

The verdict

The literature on neural coding is immense, and going into the mathematical details of proofs would take more time and effort than can be spared for this thesis. Therefore, a literature review has been attempted instead. While logarithmic representation emerges as a plausible candidate for sensory coding from the literature reviewed, it is also an old framework with much work behind. Newer frameworks have yet to receive as much attention, skewing the review somewhat. This also highlights that there might be other frameworks, not yet thought of, that provide an even better fit. However, based on the data and theory available, logarithmic coding prevails.

If the fundamental coding of sensory input is logarithmic, and differences between logarithms is a ratio, then this might be the mechanistic restraint Stephens (2002) proposed. This means that the brain cannot help but use ratio comparison, and further cannot help discount temporally. This makes the emotional commitment to honesty an easier fix than changing sensory encoding at the neural level.

Commitment?

Now, how does this help us answer whether aliens would feel emotional commitment to honesty? So far, we know that logarithmic encoding of stimuli is a viable system for advanced lifeforms. We know that, according to Sun et al. (2012) and Portugal and Svaiteer (2011), it should be favoured by evolution. We know that because of this coding scheme, evolution ran into the problem of temporal discounting. This is bad for cooperation and long-term benefits, as immediate rewards get prioritised. We know that in our case, evolution fixed

this problem by implementing a quick fix, the emotional commitment to honesty, since it could not fix the underlying architecture.

So, according to our line of argumentation, humanity's ability to cooperate with each other is in part due to an evolutionary patch. Thus, our odds of meeting cooperative aliens would depend on how strong the selection pressure for logarithmic coding is, constraints on their neural substrate, their evolutionary history, and whether there are alternative patches for temporal discounting. However, we have only discussed the emotional drive towards cooperation, not the rational drive towards maximum gain and future benefits. This emotional commitment to honesty may only have needed to help us past short-sightedness and into the stage where rationality prevails. While researchers have found that human economic rationality often leaves much to be desired (e.g. De Martino, Kumaran, Seymour & Dolan, 2006), we cannot discount hyper-rationality in spacefaring civilisations. Whether this bodes good or bad for us, is another matter.

Rationality and cooperation

If aliens were hyper-rational, i.e. that they always choose the strategy that gives the maximum reward and least risk, which strategy would they chose? Earlier, we examined Tit-for-Tat and found that it can reliably beat constant defectors. However, much research has been done on the subject of optimal strategies, and the superiority of unmodified TFT has been challenged.

Generous TFT

A slight modification of TFT is Generous TFT, where opponents are forgiven 10 % of the time. Nowak and Sigmund (1992) have shown that Generous TFT gives slightly better results than pure TFT. The authors have also stressed that TFT needs to be in use by a small fraction of the population for it to take hold. This will in turn pave way for the generous version of TFT.

Dugatkin and Alfieri (1991) have shown that Guppies exhibit TFT-like behaviour. Moreover, they also appear to remember defectors and cooperators over at least 4 hours, giving some credence to a reputation system. The authors have also pointed out that this preferential treatment given to cooperators may help TFT gain the necessary foothold in a population.

Pavlov

Another strategy was presented by Nowak and Sigmund (1993). Named Pavlov, the strategy reminds one of reflex-like behaviour: If you win, you keep playing your last move, if you lose, you change behaviour. Win-stay, lose-shift. Winning is defined by cooperating with a cooperator, or defection against a cooperator. Conversely, losing is cooperating with a defector, or defecting against a defector. The authors have argued that Pavlov is better than TFT because it can correct occasional mistakes, and also exploit unconditional cooperators. This last feature prevents the unconditional cooperation strategy that invite defectors into the

population. However, Pavlov will be heavily exploited by unconditional defectors, as Pavlov will cooperate every other turn.

The Pavlov strategy is superior to Generous TFT only if both players act at the same time. Sequential acting may be more realistic in a natural environment, in which case Generous TFT is better. Pavlov can only be efficient in sequential games if longer memory is utilised (Nowak & Sigmund, 1994). Wedekin and Milinski (1996) tested students on simultaneous and sequential prisoner's dilemma, and found that Generous TFT did better in the sequential games, while a Pavlov-like strategy did best in simultaneous play, as predicted. While the games were organised so that only the last round was displayed, thus ideally making it a memory-one game, it is possible that the students used longer memory when playing. Press and Dyson (2012) have shown that memory greater than one shouldn't provide an advantage in simultaneous play, but they make no such claim for sequential play.

Since Pavlov is claimed to be a manifestation of win-stay, lose-shift, which is widely seen among animals, one would expect to see it in cooperation games as well. This appears not to be the case. One reason why might be that while Pavlov classifies both cooperating with a cooperator and defection against a cooperator as wins, animals might consider only temporally close rewards as wins, as per temporal discounting. This may lead to classifying game theoretic losses as wins, because the rewards were immediate, and likewise theoretical wins as losses (Clements and Stephens, 1995). In other words, we don't see Pavlov in nature because our "win" conditions may not be the animals' "win" conditions, since temporal discounting makes immediate wins more desirable.

Zero-determinant

A new class of strategies was recently discovered by Press and Dyson (2012), called Zero-determinant (ZD). They have the ability to force a linear relationship between the player's

own score and that of the opponent, making the player able to control the long-term payoffs for both players. This is achieved by assigning probabilities to the various responses, so that, for instance, cooperation is met by defection with a certain probability. There are many different strategies in this class, including the much older strategy TFT, which forces the opponents score to be equal to the player's own. Another interesting ZD strategy is extortion, where the more you let yourself be taken advantage of, the better you will score. If you wish to punish the extorter, you have to do so at a cost to yourself.

Of particular note is a subset of strategies called generous ZD (Stewart & Plotkin, 2013; Akin, 2012). These strategies forgive opponents' defection, but not too often. Stewart and Plotkin (2013) have shown that generous ZD strategies are able to outplay most other strategies, even earlier champions like TFT or Pavlov, in addition to less generous ZD strategies.

However, for the generous ZD strategies to be successful, they require that the population isn't too small. Extortion strategies dominate the smaller populations (between two and ten agents, depending on parameters). Adami and Hintze (2013) have argued against the evolutionary stability of ZD strategies, and have claimed that mutations of ZD strategies are unlikely to be ZD themselves. However, a 2014 corrigendum clarifies that they referred to a coercive subset of strategies, like those described by Press and Dyson (2012). Adami and Hintze has conceded that the generous ZD strategies described in Stewart and Plotkin (2013) are evolutionary stable, able to exist in a population without being outperformed by invading strategies. Zero-determinant strategies are still a fresh find, and so animal models are scarce, if not non-existent.

Memory-one

All these strategies for simultaneous iterated prisoner's dilemma have in common that they only use information gathered in the previous round. Press and Dyson (2012) have proved

that longer memory does not proffer any advantages, in that a player with a good memory-one strategy effectively sets the rules of the game. However, this does not mean that longer memory is irrelevant in cooperation, only that theoretically effective strategies need not bother with it. Longer memory is interesting as reputation, i.e. using information from earlier games, might be one way to ensure cooperation. It is important to note that Press and Dyson's proof is only valid for two-player simultaneous play, while Nowak and Sigmund (1994) was discussing sequential play when arguing that longer memory was required for the success of Pavlov.

It is worth noting that the research so far has been analysing two-player iterate prisoner's dilemma. Work has been done on n-player games, where n is greater than 2, and where each player's payoff depends on the payoff of n-1 other players. Needless to say, this quickly gets complicated, so this thesis have only focused on two players at a time. An interesting point about n-player games is that a longer memory appears to be important for winning (Li & Kendall, 2014). While complicated in regards to model-making, n-player games may be prevalent in nature. If memory equals reputation, then Li and Kendall's finding emphasises the importance of reputation.

Since longer memory isn't required, simpler organisms may have evolved these strategies more easily, thus increasing the probability of finding ZD strategies in nature. TFT and Pavlov have exceedingly simple instructions, further giving credence to their universality. ZD strategies contain probabilities of performing certain actions, which may or may not be as easily encoded. However, Webb and Dorris (2013) have proposed a neural model where probabilities can be implemented.

If these strategies are available to simple beings, then they are surely available to our hyper-rational alien friends. This is generally good for us, as these strategies typically goes for mutual cooperation, and will even forgive transgressions now and then. While there are some

less virtuous strategies in the ZD-class, the generous strategies have been proven to outperform these (Stewart and Plotkin, 2013). Constant defection is also not among the optimal strategies, giving further reassurance of rational beings' cooperativeness.

Bounded rationality

Between relying on hardwired neural preferences and hyper-rationality in deciding how to act, our aliens might behave closer to our standards of rationality, called bounded rationality. This model assumes humans act rational within the confines of available information, cognitive capacity, and time (Simon, 1955). This is different from hyper-rationality in that factors like stress (Van den Bos, Harteveldt & Stoop, 2009), glycogen stores (Gailliot & Baumeister, 2007), assorted biases (e.g. Tversky & Kahneman, 1974), etc., influence the choice.

A form of bounded rationality is a likely if the aliens have evolved from very simple organisms, since evolution would presumably not evolve a perfectly rational being given the various stages and adaptations that the being would pass through reach the end result. An exception is that the aliens could have been artificially designed into perfect rationality by another, advanced species, but that discussion is beyond the scope of this thesis.

If we assume that cooperation is necessary for an advanced society, then either they have evolved into fundamentally cooperative beings, or, if defection is likelier because of the neural circuitry, then this deficiency must be patched in some way or other for society to function. There is also the alternative of cooperation-inducing mechanisms.

Cooperation-inducing mechanisms

Kinship

The first line of defence against defection is shared genes. When asked if he'd save a brother from drowning, evolutionary biologist J.B.S. Haldane jokingly said, "No, but I'd jump in the river for two brothers or eight cousins". This is the essence of kin selection, the notion that one's relatives also carries your genes, so that investing in them equals investing in your own genes' proliferation (Hamilton, 1964a; Hamilton, 1964b). We see this in the case of the Nayar mercenaries mentioned earlier, where investing in their nephews and nieces was a viable solution to the illegitimacy of their wives' children.

Group augmentation

After relatives, the next most important thing is the group. Cooperation within the group strengthens the group, further improving the animals' chance of survival. Kokko, Johnstone, and Clutton-Brock (2001) have proposed that cooperative breeding, i.e. helping raise the children of unrelated others in your group, may improve fitness by augmenting the group. However, Clutton-Brock (2002) later argued that there appears to be no singular explanation for cooperative breeding, and that cooperative animal societies may show closer parallels to human society than previously thought. Depending on how large one considers one's group, group augmentation would likely boost cooperation in species. A prevailing presumption in social psychology holds that in-group cohesion might also boost out-group hostility (Brewer, 1999). However, Brewer (1999) have argued that in-group cohesion does not increase hostility towards other groups, it only increases the relative difference in attitude towards either.

Punishment

Other cooperation-inducing mechanisms are more generally applicable. If you can convince your partner that cheating will be paid for, then the treat alone might deter cheating. However, vengeance itself is expensive in terms of risking harm to oneself. Thus, having a

reputation as vengeful is helpful, but actually doing the avenging is not. In humans, enforcing costly punishment has been shown to both increase and decrease cooperation in prisoner's dilemma (Wu et al. 2009). The cited research was done on both US and Chinese citizens, and the authors explain the discrepancy with cultural differences. Punishment of cheaters is found both among humans and other animals (e.g. Rohwer, 1977; Møller, 1987). Zhang, Li, De Silva, Bednarik and Sigmund (2014) have shown how a social contract to not cheat can emerge spontaneously.

Morality

In advanced enough species, and humans are the only species we know of that fits this description, individuals may have internal rules that govern prosocial behaviour. We call these rules morality, and they are special in that there is no inherent external punishment for not following them. Of course, laws are often built upon morality, obfuscating whether we chose a certain behaviour out of fear of punishment or because our morality demanded it. In choices of less consequence, however, morality do play a significant, albeit variable part, as humans are as wont to follow their morality as not (Lahti & Weinstein, 2005). A hypothesis put forward by Lahti and Weinstein (2005) holds that variability in moral adherence is due to the need to balance group cohesion and in-group competition. This makes it difficult to predict behaviour based on morality.

Another problem with morality is its fluidity across cultures (e.g. Trompenaars, 2003) and social positions (Piff, Stancato, Côté, Mendoza-Denton, & Keltner, 2012). One culture's moral may be to play nice and keep the peace, another's could be to eradicate all unbelievers. This makes prediction behaviour from morality precariously close to guesswork.

We have seen that, while cooperation in a species is not a given fact, there is ample reason to expect it in a species advanced enough to make contact across the stars. However,

could we make any assumptions about how they would view us? To answer this, we need to look at other species.

Cooperation between species

Mutualism

Interspecies interaction can take many forms. A useful way to categorise these interactions is to make a grid like the one in table 3.

		Species 2		
		+	-	0
Species 1	+	Mutualism	-	-
	-	Predation/ Parasitism	Competition	-
	0	Commensalism	Amensalism	Neutralism

Table 3: Grid of interactions between species (Bronstein, 1994).

We are mostly interested in mutualism, i.e. mutually beneficial interaction between different species. Note that while mutualism is sometimes used to mean intraspecies interaction (cooperation within a species), as well, this is not technically correct, as West, Griffin, and Gardner (2007) have pointed out. The most widely studied kind of mutualism is that between animals and plants (Bronstein, 1994), as pollination and seed dispersal are among the most readily observable mutualisms in nature. Others examples are the relationship between intestinal bacteria and humans (Bäckhed, Ley, Sonnenburg, Peterson, & Gordon, 2006), cleaner fish and their clients (Pinto et al., 2011), and the evolution of the eukaryotic cell, when free living bacteria entered a symbiotic relationship with prokaryotes (Margulis, 1981).

Indirect reciprocity

Kin selection, which likely is the root of much cooperation in the animal kingdom (Clutton-Brock, 2009), cannot possibly apply between different species. Indirect reciprocity, however, is possible. Eckardt and Zuberbühler (2004) found cooperation between two forest monkey species. It was hypothesised that their overlapping habitat and resource-use should

lead to competition, not cooperation, but apparently the weaker species was tolerated because it provided defence from eagles. Interspecies cooperation was most prevalent when food was abundant.

Asymmetry

A problem with interspecies cooperation is asymmetry, which generally hinders cooperation (Dawkins, 2010). There are several ways in which species can be asymmetric, including power and perspective.

Power. Power is relatively straightforward, and differences in size, lethality, and tool-use can easily make the species asymmetric. If this difference is large enough, it would presumably make any interaction highly one-sided. Gingins, Werminghausen, Johnstone, Grutter, and Bshary (2013) have used cleaner fish to show that power asymmetry can be offset if the weaker species is able to terminate cooperation at will, and if the stronger species has a low (but still existing) incentive to cheat.

Perspective. By perspective we mean that the species world-view differs to a significant degree. Lifespan is a good example of perspective asymmetry. Long-term and short-term are relative terms, so long-term for us might be short-term for a near-immortal species. While relative immortality could be construed as a difference in power, it would also likely entail different viewpoints. In fiction, immortal characters often portrayed as having a disregard mortal life, having no qualms about sacrificing lesser beings for the greater good. In Douglas Adams' *The Hitchhiker's Guide to the Galaxy*, Earth is destroyed to make a hyperspatial express route through the solar system (Adams, 1979). While originally satirising bureaucracy, the situation also depicts differences in perspective.

One problem of finding comparable interspecies cooperation with respect to longevity is that other factors also tend to be asymmetric. Take, for instance, man's relation to insects.

While the difference in lifespan is just what we are looking for, other parameters are too dissimilar for meaningful comparison.

Communication barrier

Tim Clutton-Brock (2009) has posited that the human ability to maintain cooperation with non-kin (with delayed rewards) and lack of cheating could be credited to language, more specifically the ability to make agreements regarding future payoffs and sanctions. Language also facilitates reputation, as it is no longer necessary to personally observe the making or breaking of a potential co-players reputation.

Language is typically not shared across species; it is not even usually shared across cultures. This creates a potential stumbling block for cooperation. However, different civilisations have traded through the ages, and humans have trained animals to work for a reward of food later. Perhaps only a rudimentary signalling system is needed for cooperation to take place, at least if the cooperation is short-term. Long-term cooperation, on the other hand, likely requires more sophisticated communication.

Empathy

Whether we feel empathy or antipathy towards someone depends in part on how fair we perceive them (Singer et al., 2006). In other words, this means empathy is conditional and not automatic; receivers of our empathy need to earn it first. Since emotional commitment to honesty hinges on empathy (Frank, 1988), then this is another challenge that interspecies cooperation needs to overcome. No small task without communication or shared culture.

Oxytocin has been found to regulate social functions in many vertebrates (Miller, 2010). It is, among other functions, known to strengthen empathic bonds between people, including their pets (Beetz, Uvnäs-Moberg, Julius, & Kotrschal, 2012). However, it also seems to strengthen out-group antipathy (De Dreu et al., 2010), and decreases adherence to fairness

norms (Radke & de Bruijn, 2012). We would surely be the out-group to alien visitors, so if oxytocin or a similar hormone works the same for other advanced species, then their emotional commitment to honesty towards us seems uncertain at best.

Cooperation?

There is precedent for interspecies cooperation in nature, but there are a number of factors that makes it unlikely to happen on a regular basis. In the case of aliens and us, both power and perspective might differ to a large extent. First off, they have interstellar flight, suggesting powerful technology. This would presumably give them less incentive to cooperate with us, as the power balance tips in their favour.

Secondly, there is the matter of perspective. Since we have already established their technological superiority, it is not a great leap of intuition to assume they could be in possession of great longevity, or even relative immortality. While no research on the topic can be found, understandably enough, we could make a case for our long-term goals being short-term for them. Why value a human life, spanning mere decades, compared to their millennia?

There is also the issue of empathy. Humans only show empathy towards others deserving of it, making first impressions crucial. Even the hormone that causes parents to fall in love with their screaming new-borns have the additional effect of enhancing xenophobia. Since these functions are evolutionary beneficial, they are not unlikely to appear in similarly advanced species.

Last, and perhaps least, is the issue of communication. There is a chance we would be able to communicate in a rudimentary fashion, after all, early traders managed it without any prior knowledge of distant languages. On the other hand, humans share fundamental uniformities that can be referenced when communicating. Pointing, facial expressions, signalling hunger by gesturing at the mouth or stomach, peace by displaying empty hands, to

name a few, are likely universally understood. On the other hand, there are many gestures that show no overlap between cultures (Poortinga, Schoots & Koppel, 1993). Likewise, there no guarantee that our aliens would be humanoid enough for even basic gestures to translate well.

Discussion

This thesis has explored a wide range of mechanisms in the hopes of being able to predict behaviour in a theoretical way. This is useful when the luxury of observation is unavailable. There are many unknowns in the argument, and many presumptions, as it is impossible to be sure of even the most basic foundations of non-terrestrial life. However, based on what we do know is possible, we have constructed a scenario where some speculation can be attempted, independent of ecological or morphological variables.

There are evidently many factors that could cause contact, and cooperation, between advanced species to not proceed optimally. Are there any factors that predict any other outcome? We have previously investigated emotional commitment to honesty and its role in guiding humanity towards long-term payoffs. However, we do not assume that this patch moderates behaviour against other species. While not examined in this thesis, and not featuring too prominently in our history, humans also appear to have an emotional commitment against cruelty and injustice. This is relevant as it applies between at least some species, exemplified by our resistance to animal cruelty. As mentioned, empathy is at the core of emotional commitment to honesty, and it is likely that the same holds for cruelty. We are very capable of performing atrocities, but we have to assume that performing them hurt initially, before getting accustomed to them. We can only hope this applies to similarly advanced intelligences as well, but there seems to be many barriers between mutual empathy.

What if we landed on their planet?

What if we reversed the scenario, that instead of the aliens landing on earth, humans landed on the alien homeworld? In that case, the power asymmetry might well be flipped, and we have the reverse scenario. The difference is that we know how humans generally act, so there is less speculation, and more a retrospective look on human history.

There is no shortage of evidence for what happens when civilisations meet cultures they don't understand. The Spaniards decimated the Aztecs, the Americans forced the Native Americans into reservations and stole their lands, and in modern times, the Chinese occupied Tibet and killed as much as a fifth of the indigenous population (White, 2002), to name but a few.

It could be argued that a large portion of indigenous genocide and repression was committed through ignorance. The Spaniards, for example, unwittingly carried with them diseases which effectively defeated the Aztecs for them (Crosby, 1967). It could also be argued that pre-modern humans did not understand that indigenous people were the same as them, or that they were capable of governing themselves. The white man's burden, the imperative to bring culture and civilised values to peoples less fortunate than the white man, might conceivably have been founded on compassion, but the results were often eradication of culture, freedom, and lives. Perhaps rationalisation is a better explanation of white man's burden, as there always seemed to be plenty of land and resources to be had in the invaded countries. Desmond Tutu once joked that "When the missionaries came to Africa they had the Bible and we had the land. They said "Let us pray." We closed our eyes. When we opened them, we had the Bible and they had the land" (Gish, 2004).

So, while we have shown that humans possess emotional, in addition to rational, incentives to cooperate with each other, there is little precedence to suggest this would carry over to other, and in our eyes, inferior, creatures. We can always hope that rationality prevails, but while human evolution has been relatively swift (Wrangham & Peterson, 1996), we have not yet evolved away from the conquistadors.

Possible sources of error

Validity of prisoner's dilemma. Prisoner's dilemma and other public goods games are necessarily simplified versions of the games played in real life. The question is not how realistic the games are, but if they have ecological validity. Clements and Stephens (1995) have argued that there is no empirical justification for using prisoner's dilemma as a basic paradigm for studying non-kin cooperation in animals. Clutton-Brock (2009) seems to echo this view. Dawkins (2010) have argued that the assumption of symmetry in prisoner's dilemma is unrealistic in a natural setting. Moreover, since asymmetry appear to hinder cooperation, there is reason to doubt the validity of prisoner's dilemma in studies of cooperation.

Neuronal coding. While there is ample evidence for a logarithmic coding scheme, we found several alternative schemes have not yet been thoroughly examined. Reasons for this presumably range from their relative recent proposal, to the fact that it is difficult for new ideas to take root in academic soil (Campanario, 2002). Like MacKay (1963) posited in regard to whether coding were logarithmic or power-based, we can only get so far without examining the actual neural substrates.

A potential problem with our literary review of neural coding is that we have been forced to take the authors' words for why their evidence fits the particular model they argued for. We do not have the necessary mathematical foundation to do our own calculations, which is particularly troubling as some studies cited the same data, but with different conclusions. Moreover, it was beyond the scope of the thesis to do an exhaustive systematic literature review. However, we do maintain that the accumulated evidence points toward logarithmic rather than linear coding, but that more rigorous and exhaustive research is needed for a definitive answer.

We would like to address a curious pattern that emerged when examining the research on neural coding schemes. Although the published opinions favouring logarithmic coding of stimuli appear to be largely independent, we find that those in favour of linear coding all seem to have collaborated with John Gibbons, the originator of scalar expectancy theory, at some point. This might not mean much, but adding that to Wearden and Jones' (2007) opinion that opposition to linear representation might be due to opposition to scalar expectancy theory, then we get the feeling that the debate has gotten personal. While this is largely speculation from our side, it does not make a strong case for linear coding's objectivity.

Frank's golden opportunity. As Frank (1988) himself mentions, the existence of surveillance and "candid camera" makes even seemingly golden opportunities to cheat not worth the risk. There is simply no guarantee that you will not get caught. We do see, however, that birds will use these opportunities to cheat (Baltz & Clark, 1997), so something is evidently different between our species. It is possible that birds cannot conceive of an unseen watcher the way we can, yet there is no denying the emotional reaction opportunities to cheat elicit in us. Even if there is no such thing as a golden opportunity, there is such a thing as emotional commitment to honesty.

Implications

We have seen that the literature on neural coding have predominantly focused on linear and logarithmic coding schemes. We hope that this thesis have emphasised clearly that, while logarithmic coding is a viable scheme for neural representation of stimuli, and based on what research has been done to date, a probable one, it is not the only alternative. Coding based on modulo (Fiete et al. 2008), for example, seems a promising candidate, and Gallistel's (2011) autoscaling function deserves more than a theoretical examination.

A missed opportunity of this thesis is the mathematical aspect. It is possible that a mathematical analysis of the arguments we have presented could have given more confident answers, rather than informed speculation. However, given the author's qualifications, this proved beyond the scope of this thesis. A mathematical exploration of the foundations laid here could prove an interesting starting point for predictive models of advanced non-human intelligences.

Relevance

Psychology has always concerned itself with animal behaviour, human or otherwise. Knowing why something happens is the first step to predicting when it will happen, and how we can ensure that it does. This thesis has concerned itself with why cooperation happens in a universe where survival of the fittest is the rule everybody plays by. This should allow us to predict when cooperation breaks down, and which countermeasures can be taken to ensure it does not happen.

Neural coding is important for our understanding of the brain, as well as artificial intelligence. Nature is a great place to borrow designs from, but the designs aren't perfect. This is evident from the emotional commitment to honesty patch that Frank (1988) argued for. When designing an intelligence from the ground up, it would be useful to avoid quick-and-dirty fixes and ad hoc solutions.

With regard to psychopathology, when things go wrong, it is presumably less likely to be a flaw in the fundamental workings of the brain, but rather a fault in higher functioning. The honesty patch seems a likelier target than the neural coding scheme. Impulsivity, which is fundamental in disorders like ADHD, Borderline, and substance-abuse, is closely linked with temporal discounting (Wittmann & Paulus, 2008), which again is linked with logarithmic

coding. The more precise our understanding of how we tick, the more precise our interventions can be.

Lastly, while the focus of this thesis may come across as frivolous, it is none the less a scenario that, depending on who you ask, is likely to happen at some point or other. The universe is, to quote Douglas Adams, vastly hugely mind-bogglingly big, and already NASA have identified thousands of planets theoretically suitable for life (National Aeronautics and Space Administration, 2015). The main obstacle is the large distances involved, as crossing to the nearest galaxy, Alpha Centauri, would take almost five years travelling at the speed of light, and we are nowhere close to that speed yet. However, our hypothetical aliens might be.

Conclusion

We asked if aliens would feel this emotional commitment to honesty. What we have seen is that, while cooperation is integral to human society, cooperation appears to have become prevalent in part due to an evolutionary patch, the emotional commitment to honesty. This was necessary, as the underlying neural coding scheme likely favours discounting future benefits, making long-term cooperation difficult. Other factors, like reputation and punishment, likewise incentivise cooperation, but strong temporal discounting keeps the balance in check. Since our tendency to cooperate is partly based on a quick fix, generalisation becomes difficult. However, budding civilisations would likely need some form of cooperation to achieve progress, so it is reasonable to assume that cooperation exists. It might be based on rationality rather than emotion, but the effect would likely be similar. There are also a handful of cooperation-inducing mechanisms that are likely to be in effect, including punishment and group enhancement.

We also asked whether they would feel this commitment towards us, which is another matter entirely. There are many reasons for why cooperation should fall through, including communication barriers, asymmetric relationship between species, and lack of empathy. History also provides abundant examples of human non-cooperation with other species or cultures. In the end, it likely comes down to the cost-benefit ratio of partnership versus conquest.

References

- Adami, C., & Hintze, A. (2013). Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature communications*, 4.
- Adami, C., & Hintze, A. (2014). Corrigendum: Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature communications*, 5.
- Adams, D. (1979). *The Hitchhiker's Guide to the Galaxy*. London: Pan Books.
- Akin, E. (2012). Stable Cooperative Solutions for the Iterated Prisoner's Dilemma. *arXiv:1211.0969*.
- Alexander, J. McKenzie, (July 19, 2009). Evolutionary Game Theory. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*. Retrieved from <http://plato.stanford.edu/archives/fall2009/entries/game-evolutionary/>
- Allan, L. G. (2002). Are the referents remembered in temporal bisection? *Learning and Motivation*, 33(1), 10-31.
- Ansari, D. (2008). Effects of development and enculturation on number representation in the brain. *Nature Reviews Neuroscience*, 9(4), 278-291.
- Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of conflict resolution*, 24(1), 3-25.
- Bäckhed, F., Ley, R. E., Sonnenburg, J. L., Peterson, D. A., & Gordon, J. I. (2005). Host-bacterial mutualism in the human intestine. *Science*, 307(5717), 1915-1920.
- Baltz, A. P., & Clark, A. B. (1997). Extra-pair courtship behaviour of male budgerigars and the effect of an audience. *Animal behaviour*, 53(5), 1017-1024.

- Bateson, M., Nettle, D., & Roberts, G. (2006). Cues of being watched enhance cooperation in a real-world setting. *Biology letters*, 2(3), 412-414.
- Beetz, A., Uvnäs-Moberg, K., Julius, H., & Kotrschal, K. (2011). Psychosocial and psychophysiological effects of human-animal interactions: the possible role of oxytocin. *Frontiers in psychology*, 3, 234-234.
- Bowers, J. S. (2009). On the biological plausibility of grandmother cells: implications for neural network theories in psychology and neuroscience. *Psychological review*, 116(1), 220.
- Brannon, E. M., Suanda, S., & Libertus, K. (2007). Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Developmental science*, 10(6), 770-777.
- Brannon, E. M., Wusthoff, C. J., Gallistel, C. R., & Gibbon, J. (2001). Numerical subtraction in the pigeon: Evidence for a linear subjective number scale. *Psychological Science*, 12(3), 238-243.
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, 55(3), 429-444.
- Bronstein, J. L. (1994). Our current understanding of mutualism. *Quarterly Review of Biology*, 31-51.
- Brosnan, S. F., Parrish, A., Beran, M. J., Flemming, T., Heimbauer, L., Talbot, C. F., ... &

- Wilson, B. J. (2011). Responses to the Assurance game in monkeys, apes, and humans using equivalent procedures. *Proceedings of the National Academy of Sciences*, *108*(8), 3442-3447.
- Bueti, D., & Walsh, V. (2009). The parietal cortex and the representation of time, space, number and other magnitudes. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *364*(1525), 1831-1840.
- Caldwell, R. L. (1986). The deceptive use of reputation by stomatopods. In Mitchell, R. W. & Thompson, N. S., (Eds.), *Deception: Perspectives on human and nonhuman deceit* (pp. 129-145). SUNY Press.
- Campanario, J. M. (2002). The parallelism between scientists' and students' resistance to new scientific ideas. *International Journal of Science Education*, *24*(10), 1095-1110.
- Cappelletti, M., Freeman, E. D., & Butterworth, B. L. (2011). Time processing in dyscalculia. *Frontiers in psychology*, *2*, 364.
- Cerutti, D. T., & Staddon, J. E. R. (2004). Immediacy versus anticipated delay in the time-left experiment: a test of the cognitive hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*(1), 45.
- Church, R. M., Meck, W. H., & Gibbon, J. (1994). Application of scalar timing theory to individual trials. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*(2), 135.
- Clements, K. C., & Stephens, D. W. (1995). Testing models of non-kin cooperation: mutualism and the Prisoner's Dilemma. *Animal Behaviour*, *50*(2), 527-535.

- Clutton-Brock, T. (2002). Breeding together: kin selection and mutualism in cooperative vertebrates. *Science*, 296(5565), 69-72.
- Clutton-Brock, T. (2009). Cooperation between non-kin in animal societies. *Nature*, 462(7269), 51-57.
- Cordes, S., King, A. P., & Gallistel, C. R. (2007). Time left in the mouse. *Behavioural processes*, 74(2), 142-151.
- Crespi, B. J. (2001). The evolution of social behavior in microorganisms. *Trends in ecology & evolution*, 16(4), 178-183.
- Crosby, A. W. (1967). Conquistador y pestilencia: the first new World pandemic and the fall of the great Indian empires. *Hispanic American Historical Review*, 321-337.
- Crystal, J. D. (2001). Nonlinear time perception. *Behavioural processes*, 55(1), 35-49.
- Crystal, J. D. (2006). Long-interval timing is based on a self-sustaining endogenous oscillator. *Behavioural processes*, 72(2), 149-160.
- Dawkins, M. S. (2010). Do asymmetries destabilize the Prisoner's Dilemma and make reciprocal altruism unlikely? *Animal Behaviour*, 80(2), 339-341.
- De Castro, A. V., & Machado, A. (2010). Prospective timing in pigeons: Isolating temporal perception in the time-left procedure. *Behavioural processes*, 84(1), 490-499.
- De Dreu, C. K., Greer, L. L., Handgraaf, M. J., Shalvi, S., Van Kleef, G. A., Baas, M., ... & Feith, S. W. (2010). The neuropeptide oxytocin regulates parochial altruism in intergroup conflict among humans. *Science*, 328(5984), 1408-1411.
- De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and

rational decision-making in the human brain. *Science*, 313(5787), 684-687.

Decibel. (2015). In *Encyclopædia Britannica*. Retrieved from

<http://global.britannica.com/EBchecked/topic/155074/decibel>

Dehaene, S. (2003). The neural basis of the Weber–Fechner law: a logarithmic mental number line. *Trends in cognitive sciences*, 7(4), 145-147.

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, 5(4), 390-407.

Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science*, 320(5880), 1217-1220.

Dicke, M., & Sabelis, M. W. (1992). Costs and benefits of chemical information conveyance: proximate and ultimate factors. In Roitberg, B. D., & Isman, M. B. (Eds.), *Insect chemical ecology: an evolutionary approach* (pp. 122-155). Springer Science & Business Media.

Doebeli, M., & Hauert, C. (2005). Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. *Ecology Letters*, 8(7), 748-766.

Dufour, V., Pelé, M., Sterck, E. H. M., & Thierry, B. (2007). Chimpanzee (*Pan troglodytes*) anticipation of food return: coping with waiting time in an exchange task. *Journal of Comparative Psychology*, 121(2), 145-55.

Dugatkin, L. A., & Alfieri, M. (1991). Tit-for-tat in guppies (*Poecilia reticulata*): the relative nature of cooperation and defection during predator inspection. *Evolutionary Ecology*, 5(3), 300-309.

- Dzieweczynski, T. L., Earley, R. L., Green, T. M., & Rowland, W. J. (2005). Audience effect is context dependent in Siamese fighting fish, *Betta splendens*. *Behavioral Ecology*, *16*(6), 1025-1030.
- Eckardt, W., & Zuberbühler, K. (2004). Cooperation and competition in two forest monkeys. *Behavioral Ecology*, *15*(3), 400-411.
- Fiete, I. R., Burak, Y., & Brookings, T. (2008). What grid cells convey about rat location. *The Journal of Neuroscience*, *28*(27), 6858-6871.
- Frank, R. H. (1988). *Passions within reason: the strategic role of the emotions*. New York: WW Norton & Co.
- Gailliot, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review*, *11*(4), 303-327.
- Gallistel, C. R. (2011). Mental magnitudes. In Dehaene, S. & Brannon, E. (Eds.), *Space, Time and Number in the Brain*, (pp. 3-12).
- Gallistel, C. R., & Gelman, R. (2000). Non-verbal numerical cognition: From reals to integers. *Trends in cognitive sciences*, *4*(2), 59-65.
- Gibbon, J. 1977. Scalar expectancy theory and Weber's law in animal timing. *Psychological Review*, *84*, 278-325.
- Gibbon, J., & Church, R. M. (1981). Time left: linear versus logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes*, *7*(2), 87.
- Gibbon, J., Malapani, C., Dale, C. L., & Gallistel, C. R. (1997). Toward a neurobiology of

temporal cognition: advances and challenges. *Current opinion in neurobiology*, 7(2), 170-184.

Gingins, S., Werminghausen, J., Johnstone, R. A., Grutter, A. S., & Bshary, R. (2013). Power and temptation cause shifts between exploitation and cooperation in a cleaner wrasse mutualism. *Proceedings of the Royal Society B: Biological Sciences*, 280(1761), 553.

Gish, S. (2004). *Desmond Tutu: a biography*. Westport, CT: Greenwood Publishing Group.

Gough, K. (1961) Nayar - Central Kerala. In Schneider, D. M., & Gough, K. (Eds.). *Matrilineal kinship* (298-384). University of California Press.

Green, L., Myerson, J., & McFadden, E. (1997). Rate of temporal discounting decreases with amount of reward. *Memory & cognition*, 25(5), 715-723.

Hamilton, W.D. (1964) The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, 7(1), 1-16.

Hamilton, W. D. (1964). The genetical evolution of social behaviour. II. *Journal of Theoretical Biology*, 7(1), 17-52.

Holden, C. J., & Mace, R. (2003). Spread of cattle led to the loss of matrilineal descent in Africa: a coevolutionary analysis. *Proceedings of the Royal Society of London B: Biological Sciences*, 270(1532), 2425-2433.

Holden, C. J., Sear, R., & Mace, R. (2003). Matrilineal descent as daughter-biased investment. *Evolution and Human Behavior*, 24(2), 99-112.

Johnson, K. O., Hsiao, S. S., & Yoshioka, T. (2002). Book Review: Neural coding and the basic law of psychophysics. *The Neuroscientist*, 8(2), 111-121.

- Kheifets, A. (2014). Time and probability in the mouse (Doctoral dissertation). Retrieved from <http://dx.doi.org/doi:10.7282/T3154F4T>
- Kokko, H., Johnstone, R. A., & Clutton-Brock, T. H. (2001). The evolution of cooperative breeding through group augmentation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 268(1463), 187-196.
- Lahti, D. C., & Weinstein, B. S. (2005). The better angels of our nature: group stability and the evolution of moral tension. *Evolution and Human Behavior*, 26(1), 47-63.
- Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation, and numerical magnitude comparison. *Child Development*, 78(6), 1723-1743.
- Li, J., & Kendall, G. (2014). The Effect of Memory Size on the Evolutionary Stability of Strategies in Iterated Prisoner's Dilemma. *IEEE Transactions on Evolutionary Computation*, 18(6), 819-826.
- MacKay, D. M. (1963). Psychophysics of perceived intensity: A theoretical basis for Fechner's and Stevens' laws. *Science*, 139(3560), 1213-1216.
- Mackevicius, E. L., Best, M. D., Saal, H. P., & Bensmaia, S. J. (2012). Millisecond precision spike timing shapes tactile perception. *The Journal of Neuroscience*, 32(44), 15309-15317.
- Macmillan, N. A. (2002). Signal detection theory. In Pashler, H., Wixted, J. (Eds.), *Stevens' handbook of experimental psychology* (pp. 43-90). John Wiley & Sons, Inc.

Margulis, L. (1981). *Symbiosis in cell evolution: Life and its environment on the early earth.*

San Francisco: W. H. Freeman & Co.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(3), 320.

Miller, G. (2010). The prickly side of oxytocin. *Science*, 328(5984), 1343-1343.

Mischel, W., Shoda, Y., & Rodriguez, M. I. (1989). Delay of gratification in children.

Science, 244(4907), 933-938.

Møller, A. P. (1987). Social control of deception among status signalling house sparrows

Passer domesticus. *Behavioral Ecology and Sociobiology*, 20(5), 307-311.

Moser, E. I., Roudi, Y., Witter, M. P., Kentros, C., Bonhoeffer, T., & Moser, M. B. (2014).

Grid cells and cortical representation. *Nature Reviews Neuroscience*. 15(7), 466-481.

Mountcastle, V. B., Poggio, G. F., & Werner, G. (1963). The relation of thalamic cell

response to peripheral stimuli varied over an intensive continuum. *Journal of neurophysiology*, 26(5), 807-834.

National Aeronautical and Space Administration, (2015, April 21). *NASA's NExSS Coalition*

to Lead Search for Life on Distant Worlds. Retrieved from <http://www.nasa.gov/feature/nasa-s-nexss-coalition-to-lead-search-for-life-on-distant-worlds>

Nieder, A., & Miller, E. K. (2003). Coding of cognitive magnitude: compressed scaling of

numerical information in the primate prefrontal cortex. *Neuron*, 37(1), 149-157.

- Nowak, M. A., & Sigmund, K. (1992). Tit for tat in heterogeneous populations. *Nature*, 355(6357), 250-253.
- Nowak, M., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, 364(6432), 56-58.
- Nowak, M. A., & Sigmund, K. (1994). The alternating prisoner's dilemma. *Journal of theoretical Biology*, 168(2), 219-226.
- Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, 393(6685), 573-577.
- Piff, P. K., Stancato, D. M., Côté, S., Mendoza-Denton, R., & Keltner, D. (2012). Higher social class predicts increased unethical behavior. *Proceedings of the National Academy of Sciences*, 109(11), 4086-4091.
- Pinto, A., Oates, J., Grutter, A., & Bshary, R. (2011). Cleaner wrasses *Labroides dimidiatus* are more cooperative in the presence of an audience. *Current Biology*, 21(13), 1140-1144.
- Poortinga, Y. H., Schoots, N. H., & Koppel, J. M. (1993). The understanding of Chinese and Kurdish emblematic gestures by Dutch subjects. *International Journal of Psychology*, 28(1), 31-44.
- Portugal, R. D., & Svaiter, B. F. (2011). Weber-fechner law and the optimality of the logarithmic scale. *Minds and Machines*, 21(1), 73-81.
- Press, W. H., & Dyson, F. J. (2012). Iterated Prisoner's Dilemma contains strategies that

- dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26), 10409-10413.
- Rachlin, H., Raineri, A., & Cross, D. (1991). Subjective probability and delay. *Journal of the experimental analysis of behavior*, 55(2), 233-244.
- Radke, S., & de Bruijn, E. R. (2011). The other side of the coin: oxytocin decreases the adherence to fairness norms. *Frontiers in human neuroscience*, 6, 193-193.
- Ramseyer, A., Pelé, M., Dufour, V., Chauvin, C., & Thierry, B. (2005). The temporal limits of reciprocity in brown capuchin monkeys. *Proceedings of the Royal Society of London, B*, 273, 179-184.
- Roberts, W. A. (2006). Evidence that pigeons represent both time and number on a logarithmic scale. *Behavioural processes*, 72(3), 207-214.
- Rodríguez-Gironés, M. A., & Kacelnik, A. (2001). Relative importance of perceptual and mnemonic variance in human temporal bisection. *The Quarterly Journal of Experimental Psychology: Section A*, 54(2), 527-546.
- Rohwer, S. (1977). Status signaling in Harris sparrows: some experiments in deception. *Behaviour*, 61(1), 107-129.
- Shirado, H., & Maeno, T. (2005). Modeling of human texture perception for tactile displays and sensors. In *World Haptics Conference* (pp. 629-630). IEEE Computer Society.
- Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young

children. *Child development*, 75(2), 428-444.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of*

Economics, 99-118.

Singer, T., Seymour, B., O'Doherty, J. P., Stephan, K. E., Dolan, R. J., & Frith, C. D. (2006).

Empathic neural responses are modulated by the perceived fairness of others. *Nature*, 439(7075), 466-469.

Skagerlund, K., & Träff, U. (2014). Development of magnitude processing in children with

developmental dyscalculia: space, time, and number. *Frontiers in psychology*, 5.

Staddon, J. E. R., & Higa, J. J. (1999). Time and memory: towards a pacemaker-free theory

of interval timing. *Journal of the experimental analysis of behavior*, 71(2), 215-251.

Staddon, J. E. R., & Higa, J. J. (2006). Interval timing. *Nature Reviews Neuroscience*, 7(8).

Stephens, D. W. (2002). Discrimination, discounting and impulsivity: a role for an

informational constraint. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 357(1427), 1527-1537.

Stephens, D. W., McLinn, C. M., & Stevens, J. R. (2002). Discounting and reciprocity in an

iterated prisoner's dilemma. *Science*, 298(5601), 2216-2218.

Stevens, S. S. (1960). The psychophysics of sensory function. *American Scientist*, 226-253.

Stevens, S. S. (1961). To honor Fechner and repeal his law. *Science*, 133, 80-86.

Stevens, J. R., Hallinan, E. V., & Hauser, M. D. (2005). The ecology and evolution of

patience in two New World monkeys. *Biology Letters*, 1(2), 223-226.

Stewart, A. J., & Plotkin, J. B. (2013). From extortion to generosity, evolution in the iterated

prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 110(38), 15348-15353.

Strassmann, J. E., Zhu, Y., & Queller, D. C. (2000). Altruism and social cheating in the social amoeba *Dictyostelium discoideum*. *Nature*, 408(6815), 965-967.

Sun, J. Z., Wang, G. I., Goyal, V. K., & Varshney, L. R. (2012). A framework for Bayesian optimality of psychophysical laws. *Journal of Mathematical Psychology*, 56(6), 495-501.

Tibbetts, E. A., & Dale, J. (2007). Individual recognition: it is good to be different. *Trends in Ecology & Evolution*, 22(10), 529-537.

Trompenaars, F. (2003). *Did the pedestrian die? Insights from the Greatest Culture Guru*.

London: Capstone.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases.

Science, 185(4157), 1124-1131.

Van den Bos, R., Hartevelt, M., & Stoop, H. (2009). Stress and decision-making in humans: performance is related to cortisol reactivity, albeit differently in men and women.

Psychoneuroendocrinology, 34(10), 1449-1458.

Van Rijn, H., & Taatgen, N. A. (2008). Timing of multiple overlapping intervals: How many clocks do we have? *Acta Psychologica*, 129(3), 365-375.

Varshney, L. R., & Sun, J. Z. (2013). Why do we perceive logarithmically? *Significance*,

10(1), 28-31.

Wade, N. (2006). *Before the dawn: Recovering the lost history of our ancestors*. New York:

Penguin.

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in cognitive sciences*, 7(11), 483-488.

Wearden, J. H., & Jones, L. A. (2007). Is the growth of subjective time in humans a linear or nonlinear function of real time? *The Quarterly Journal of Experimental Psychology*, 60(9), 1289-1302.

Webb, R., & Dorris, M. C. (2013). A Neural Model of Stochastic Choice in a Mixed Strategy Game. *Social Science Research Network*. Retrieved from <http://dx.doi.org/10.2139/ssrn.1949732>

Wedekind, C., & Milinski, M. (1996). Human cooperation in the simultaneous and the alternating Prisoner's Dilemma: Pavlov versus Generous Tit-for-Tat. *Proceedings of the National Academy of Sciences*, 93(7), 2686-2689.

Wedekind, C., & Milinski, M. (2000). Cooperation through image scoring in humans. *Science*, 288(5467), 850-852.

West, S. A., Griffin, A. S., & Gardner, A. (2007). Social semantics: altruism, cooperation, mutualism, strong reciprocity and group selection. *Journal of Evolutionary Biology*, 20(2), 415-432.

White, D. (2002). *Himalayan tragedy: the story of Tibet's Panchen Lamas*. Tibet Society of the United Kingdom.

Wittmann, M., & Paulus, M. P. (2008). Decision making, impulsivity and time perception. *Trends in Cognitive Sciences*, 12(1), 7-12.

Wrangham, R. W., & Peterson, D. (1997). *Demonic males: Apes and the origins of human violence*. London: Bloomsbury.

Wu, J. J., Zhang, B. Y., Zhou, Z. X., He, Q. Q., Zheng, X. D., Cressman, R., & Tao, Y. (2009). Costly punishment does not always increase cooperation. *Proceedings of the National Academy of Sciences*, *106*(41), 17448-17451.

Yamada, N., & Sakama, C. (2013). Evolution of Self-interested Agents: An Experimental Study. In *Multi-disciplinary Trends in Artificial Intelligence* (pp. 329-340). Springer Berlin Heidelberg.

Yi, L. (2009). Do rats represent time logarithmically or linearly? *Behavioural processes*, *81*(2), 274-279.

Zhang, B., Li, C., De Silva, H., Bednarik, P., & Sigmund, K. (2014). The evolution of sanctioning institutions: an experimental approach to the social contract. *Experimental Economics*, *17*(2), 285-303.