University of Tartu

Institute of Philosophy and Semiotics

Relevance of Evolutionary Biology and Game Theory in Ethics

Master's Thesis in Philosophy

Meenachi Naidu

Supervisor: Kadri Simm

Tartu

2021

Table of Contents

1. Introduction 3
2. Evolutionary origins of morality 7
   - 2.1 The process of natural selection 7
   - 2.2 Selfish gene view to explain altruism 7
   - 2.3 Dilemma of altruism and morality 8
   - 2.4 Kin selection and reciprocal altruism 9
3. Game theory explanation for the evolution of morality 10
   - 3.1 Evolutionarily stable strategy (ESS) 10
   - 3.2 The caveat: hybrid model of morality in humans 14
4. Enhance morality beyond survival to promote ethics 17
   - 4.1 Game theory based research on reciprocal altruism 17
5. Integration of evolution theory into ethics 22
   - 5.1 Brief history on evolutionary ethics 22
   - 5.2 Evolution is not progressive 24
   - 5.3 Naturalistic fallacy and scope for evolutionary ethics 25
6. Cultural evolution and teaching ethics to promote moral behavior 29
   - 6.1 Ethics-teaching as a meme 32
7. Conclusion 35
8. Abstract 37
9. Bibliography 38

1. Introduction

This thesis aims to explore the evolutionary origins of morality, its components, function, and the valuable insights our biological past might offer for further advancement in the field of ethics. I take the route of biology and evolution based on the well-established fact that although altruism is evidently exhibited by other organisms, morality as a full range of attitude and behavior is only clearly expressed in humans. This is an outcome of our evolutionary history through the process of natural selection and mutual cooperation.

Morality in this thesis is meant to describe the system through which humans discern right and wrong conduct. This system constitutes of conscious as well as unconscious moral intuitions, emotions, and psychological traits that ultimately aid individuals to make moral choices and take actions for the welfare of themselves and others. It also includes the tendency to form and follow moral rules and to engage in society. Ethics is meant to imply the study of morality and inquiry into human nature to consistently understand and improve human behavior to achieve moral progress (Pojman 1990).

I explore the drift occurring in the discipline of moral philosophy, from its reliance on traditional philosophical methods to it taking on a more interdisciplinary approach welcoming insights from prominent fields such as moral psychology, neuroscience, evolutionary ethics, anthropology, and various other similar domains. The apparent divide between science and ethics can partly be attributed to the misconstruction of the perceived difference in the disciplines, restricting the scope of influence, which was popularized by Stephen Jay Gould with his notion of NOMA, i.e. Non-Overlapping Magisteria whereby he claims that the realms (magisteria) in which science and religion operate are clearly demarcated, each having the authority to teach in their respective domains without any possibility for overlap. He maintained that although science could inform moral philosophers of certain facts about the natural world, it could not claim higher insight into moral truth from any superior knowledge of the world's empirical constitution (Gould 1999).

Another reformation that has been taking place in the background is the reshaping of our communities from smaller villages to multinational societies

with globalization and the outreach of the internet and technology, calling for an upgrade in the way we cooperate. Joshua Greene has remarked on this change in his book, 'the Moral Tribes', and has argued that our ability for cooperation fails reliably at scale, as we have evolved merely to cooperate with close kin and in-group members. Humans struggle to cooperate with individuals outside their own groups. He terms this as the tragedy of the commonsense morality (Greene 2013, 5).

I draw inspiration for the explanatory parts of my thesis from Richard Dawkins' work on ethology and evolutionary biology which have particularly aided me in comprehending the key concepts and putting forward my claims. I am especially motivated by his emphasis on the secondary kind of evolution taking place in the form of cultural evolution whose units of inheritance are memes, which could supply us with the ability to combat the constraints that genetic evolution imposes upon our ethical aims. Finally, I agree with the opinion put forth by Michael Ruse and other experts in the field that evolutionary science has significant contributions to yield for the progress of ethics and needs to be given its due share of attention.

Organization of the thesis

The structure of the thesis is as follows, in this chapter I introduce the aim of the thesis and establish its scope. In chapter 2, I provide an overview of the evolutionary origins of morality expanding on how we evolved the capacity for morality as an adaption due to its value in promoting cooperation and thus increasing fitness and survival. In chapter 3, I introduce the game theory aspect explaining the various strategies that enable mutual cooperation tying it in with evolutionary stable strategy thus introducing the hybrid model of morality that that exists in humans. In chapter 4, I use Robert Axelrod and WD Hamilton's research work and tournaments as a model to show how we can harness the power of morality beyond survival, to promote wellbeing and reduce suffering. I put forward the argument that expanding on our moral attitude and behavior should be our aim as a society. In chapter 5, I address the naturalistic fallacy to make the claim that although we cannot derive an ethical framework based directly on the theory of evolution, there are valuable insights that we can take away from studying it, enhancing our current understanding of human behavior,

social cooperation, and morality. In chapter 6, I propose a possible solution to promote the aims discussed in chapter 4 through cultural evolution and teaching ethics.

Overview of the thesis

Charles Darwin's evolutionary theory along with Maynard Smith's game theory provides an apt model of how morality may have evolved in humans to aid social cooperation which ultimately leads to an individual's survival and reproduction. Even though the origins of morality are based on gene selfishness: for its struggle to survive and become numerous, morality has emerged to be a powerful adaptation that has benefited us in the long run. Humans have an innate capacity for being genuinely altruistic and care deeply for the wellbeing of others (even one's rivals[1], evolutionarily speaking) but the constraint attached to this altruistic behavior is that the net benefit to self must be higher than others in evolutionary time. The caveat is that we have been hardwired to help those that we stand to gain from, in the long run. This means that biologically, we have evolved to care for some humans more than others, specifically our own family and social groups.

In this thesis, I argue for the following three claims:

1) Morality doesn't operate in a vacuum. Given that it evolved as a tool naturally selected due to its overall benefits, it operates within a set framework where it is bound by its net payoff in terms of survival and reproduction. The evolution of morality is a dynamic process that can be explained by 'game theory' which is a branch of mathematics. I argue in chapter 4 that as a society we must aim for moral progress, promote wellbeing and reduce suffering.

2) Even though morality is an epiphenomenon of evolution through natural selection, I argue in chapter 5 that evolutionary theory provides us no direct basis for concluding moral standards by how we have evolved. Natural selection is an unconscious process that is not progressive and branches out

[1] Member of one's specie is a biological rival, competing for the same resources which tend to be finite

in several directions. From studying extinct species, it can be deduced that survival of the fittest doesn't mean survival of the best and does not give always give us the best form from the basic. Hence any form of ethics (like Social Darwinism) which tries to establish its metaethical framework on progress owing to evolution, 'that which is moral to do is what evolution does' is inappropriate (Ruse 1986, 95-97). However, I also argue that applying discoveries from evolutionary science to rephrase our understanding of the field of ethics does not lead us into committing logical fallacies. There is a general reproach for the field of evolutionary ethics due to its history for having asserted questionable morals, but this understanding is based on an incorrect and outdated interpretation of the theory of evolution and needs revival based on the improvements that have occurred in the field.

3) Morality emerged as a biological adaptation- a tool for social cooperation which aids us in mutual cooperation but also limits us from cooperating with non-kin and those outside our groups. Despite this limitation, we can utilize our innate biological capacity for morality through conscious foresight. I argue in chapter 6 that through culture and ethics teaching we can raise awareness about our innate proclivities to cut out the struggle for existence and gain from the benefits of mutual cooperation.

2. Evolutionary origins of morality

This chapter explores the evolutionary origins of morality and focuses on the relationship between morality and survival which is key to understanding why humans have developed the capacity for morality. The constraints evolution presents are this: only the traits that support its goal, i.e., survival of the genes (organism), get selected and passed down, while those traits that do not promote or adversely affect survival, are not retained. This feature when applied to the evolution of morality in humans, gives rise to the paradox of altruism i.e., promoting another's welfare at a cost to self. In this chapter, I explore the various concepts, laying the foundation for the discussion of how morality evolved in humans. The primary advantage of selecting for morality is its function to enhance social cooperation in turn promoting fitness and survival in individuals who could strike the balance between altruistic and selfish behavior.

2.1 The process of natural selection

Charles Darwin first proposed the idea of evolution through 'natural selection' where he hypothesizes that species adapt and evolve over time to be better suited to survive in an ever-changing environment. All of life on earth are interlinked and share a common ancestor. Variation in traits occurs in organisms through genetic mutation or recombination through reproduction which gives rise to individuals with diverse adaptations, strengths, and abilities. As the number of resources available in nature is limited, each individual is competing to maximize its chances of survival and genetic future with its enhanced features and advantageous traits. Therefore, those organisms that are better suited and adapted to compete with their biological rivals in a given environment and adapt faster to keep up with the changing environment survive and pass their genes down the generations (Darwin 1859).

2.2 Selfish gene view to explain altruism

There are several versions of the selection theory such as kin selection, group selection, etc., that account for how natural selection operates to give rise to altruism. The view that gathers most consensus amongst biologists is that selection occurs at the gene level, formulated by Richard Dawkins in his book 'The Selfish Gene' (Dawkins 1989). He postulates that it's not the whole species or the

organism itself that is trying to survive, rather the individual gene in the cells that is trying to be in as many bodies as possible. Each gene competes with other genes in the gene pool for its slot on the chromosome, hence natural selection occurs at the gene level. Since genes in a cell are responsible for the body they create through embryonic development, the genes are directly responsible for their future survival. Therefore, it becomes necessary for each gene to collaborate with other genes that enable to build survival machines (bodies) apt to function together and increase the body's chances of survival and reproduction which in turn enables the body to leave as many copies of its genes as possible. It is due to this phenomenon of genes working together, it appears as though the whole organism is trying to maximize its fitness and chances of survival and spreading its genetic inheritance (Dawkins 1989, 46).

Genes tend to work in unison to ensure the future survival of the bodies they create, and through which they disperse and proliferate. Genes are not selected as good in isolation but as good at working together forming gene systems, and in turn, adding to the survival of the survival machine (body) that they build together. A single gene can give rise to several effects in a body and at the same a single effect can be brought about by several gene systems working together. Therefore, a gene will be selected if it has the ability to collaborate and complement other genes in the survival machine it shares enabling survival and propagation (Dawkins 1989, 47).

2.3 Dilemma of altruism and morality

In this section, I present the dilemma of altruism which is crucial in establishing the role of game theory in understanding social cooperation and the model of morality that has evolved in humans.

Altruism is defined as any action performed by an individual at the cost of self-fitness to promote another's fitness. Several examples of altruism include young soldiers laying down their lives during war for their fellow countrymen, birds signaling others in the flock of an approaching predator at the risk of calling attention to themselves, strangers donating to charity for the welfare of individuals across continents that they may never meet or benefit from. This altruistic behavior gives rise to the dilemma, which is 1) If a gene's sole aim is its own survival and genetic future, why do we have altruistic individuals? Why did

natural selection favor the altruistic gene? How did altruistic individuals not get extinct by their own tendency of self-sacrifice? 2) If altruism comes at the cost of reducing self's fitness, why is it that we are not overpopulated by non-altruistic individuals who are better at survival than altruists?

It is possible to attempt to resolve this dilemma through the selfish gene view with two phenomena: 1) kin selection and 2) reciprocal altruism.

2.4. Kin selection and reciprocal altruism

Evolutionarily, kin selection is the tendency of a gene to help the replicas of itself (gene) in other bodies without exceeding the benefit to itself in the present body. An example is an organism helping its siblings' child but not if it comes at the cost of one's child's fitness, as one shares more genetic material with their child. Since genes compete with their rival counterparts in the gene pool, natural selection favored those genes that could help copies of themselves sitting in other bodies, this is why familial altruism evolved. The chance of relatedness and sharing genes between kin is as follows: ½ (parents, children, siblings), ¼ (grandparents, aunt, uncle, grandchildren), etcetera. There had to be a mechanism for the genes to recognize copies of themselves and the most efficient way to model this would be to promote familial altruism as one shares most copies of their genes amongst close kin. When humans lived in hunter-gatherer bands, most members were kin (Boehm, 2012) hence genes for altruism evolved. Even though some favor was rendered to non-kin, the net benefit of altruism overall outweighed the losses hence genes for altruistic behavior became numerous in the gene pool. Morality and associated emotions emerged as a result of cooperation and kin altruism (Dawkins 1989, 122).

Understood this way, it can be theorized that 'moral altruism'-selfless behavior benefits us in the long run and therefore fosters 'biological altruism'- the tendency to help others for a benefit to self in the long run (Boehm 2012).

In summary, altruism was naturally selected for its benefits to survival, which led to an increase in cooperation. Capacity for morality was a side effect of altruism and survival through cooperation.

3. Game theory explanation for the evolution of morality

This chapter aims to explain the dynamics through which social cooperation takes place among individuals as discussed in chapter 2 and to present an explanatory model on how we have evolved to possess the apparatus for social cooperation replete with our conscious as well as unconscious abilities to calculate overall costs and benefits, propped by our emotions and intuitions such as guilt, jealousy, happiness, anger, hatred, and so on, which are an accidental byproduct of the cooperation mechanism owing to its survival value. In chapter 4, I build on these concepts from game theory to argue that 'cooperation' is a better strategy for the society than 'aggressive and 'non-cooperative' strategies.

I explain the key concepts and present the argument originally formulated by biologist and game theorist Maynard Smith, in section 3.1. The salient point of this chapter, as discussed in section 3.2, is that human altruism is a derivative of the notion that our genes are selfishly[2] trying to get numerous in the gene pool. That we have the capacity for morality and altruism, which is a happenstance of our evolutionary journey and the survival advantage it procures. However, the flip side being that we have evolved to inherit a hybrid model of morality which I explain in section 3.2, wherein although it's in our biology to be altruistic towards kin and those who we stand to gain from, we have also evolved the mechanism to override it for obvious gains in the shorter run when we lose sight of the long-term benefits. This hybrid model also disables individuals from extending altruism and moral attitude towards non-kin and members outside one's shared geographical, cultural, or communal terrain if not motivated by reasons other than merely relying on one's intuitions and natural predispositions. I will elaborate more on the latter point in chapter 6.

3.1 Evolutionarily stable strategy (ESS)

As explained in the previous chapter, through the process of natural selection, genes not only aim to get excessive in the gene pool but also compete with their alleles[3] to get a slot on the chromosome and thus build robust survival machines, the organism itself, which is used as a vessel for the genes' survival and

[2] Metaphorically, as genes are not conscious

[3] Complementary (competing) gene- for example a gene for green eyes is an allele to a gene for blue eyes

propagation. This is achieved by forming efficient gene complexes by individual genes pairing with other genes. When coupled together in this manner, genes are efficient in building a resilient body that is highly likely to survive and reproduce. It is also important to note that natural selection favors genes that can control the survival machines they embody utilizing the body's ability to effectively alter and control the environment to achieve the gene's aims (Dawkins 1989, 86). For an organism carrying genes that it must help propagate, any other organism, apart from its own child or close kin, is a part of the environment that it must try to manipulate for its own gains. The challenge however is that it meets with other survival machines (organisms) who are also attempting to accomplish similar aims and hence there is a complex system of influences that organisms impose upon each other. Although there is an influence from organisms outside one's species, for example, predators on prey, the impact of the competition from members of one's own species are more prominently felt, first, due to the race for securing potential mates to produce more survival machines and second, the way in which genes are stored (DNA) is similar and therefore there is direct competition for resources necessary for life. In this sense, another organism of one's species is a direct rival for one's fitness in biological terms (Dawkins 1989).

Drawing on the idea discussed, it might appear that the most beneficial pursuit for an organism would be to eliminate its rivals, to maximize its fitness, however, this is hardly the case as inferred from studies in the wild or human interactions. Animals and humans alike (setting aside outlier cases) are known to indulge in formal tournaments and contest in a refined, systematic manner compelled to adhere to the rules, whereby individuals evoking deceit and malice are collectively punished and rooted out.

The reason organisms do not go all out and kill rival members of their own species is because there are costs and benefits attributed to the aggressive behavior. As pointed above, an intra-species rivalry forms a complicated web with organisms posing undue influence on each others' fitness. Dawkins provides the following example to illustrate the idea. If individuals A, B, and C belong to the same species and share a biological rivalry in the manner we have been discussing so far, A would benefit from eliminating B, however, A would be charged for it in terms of resources and its own fitness and may even be badly injured in the

encounter. However, C, will now not only benefit from the removal of B without paying any costs but will also be in a position to threaten A. Other things being equal, there is no obvious benefit in extricating one's rivals. On the other hand, it might be useful for A to challenge B if the outcome will endow A with a valuable resource from B, so that the overall benefits from the fight from A's perspective outweigh the costs and the threat posed by C (Dawkins 1989). Even though a simplistic example, it helps understand the large and complex system of rivalries found in nature and the underpinning of 'cost-benefit' calculations upon which the social cooperation institution rests.

Maynard Smith, first introduced the idea of 'evolutionarily stable strategy' in collaboration with G. R. Prince and G. A. Parker, using Game theory, a branch of mathematics, to formalize the method for deriving strategies that individuals stake to navigate the landscape of biological rivalry. A 'strategy' is defined as a pre-programmed behavioral policy, which may be conscious or unconscious in an individual. An example of a strategy is 'Attack first; if the opponent attacks, retreat, if the opponent flees pursue them'. Evolutionary stable strategy (ESS) is a concept that captures this idea as follows. In any given population, only one strategy, namely the ESS, when adopted by most members of that population, offers the maximum fitness value for an individual in that population. Any alternative strategy not only does worse than the ESS but also cannot affect or alter the ESS for that population once the ESS strategy is reached. For instance, a population could have two or more strategies as potential candidates for forming an ESS, however, based on the strategy the majority of the individuals in the population gravitate toward ESS is reached, and once the ESS is achieved, it is not prone to a threat from an alternate strategy, as the individuals choosing alternate strategies are treated as deviants since their cumulative fitness will be far worse than those adopting ESS and are not favored by natural selection. They simply do not make it into the next generation due to the blow to their fitness affecting their survival and reproduction. There is however some leeway to this point, which I discuss in chapter 4. To sum up, given the complex web of influences emanating from biological rivalry between individuals, the best strategy for an individual depends on what the majority of the population opts to do, given the fact that each organism is invariably striving to increase one's fitness (Smith 1982).

Following is a simple hypothetical case offered by Smith to apply the concept of ESS to aggression and cooperation. Let us assume that there are only two fighting strategies possible for a population, called 'hawk' and 'dove'. The names of the strategies are only meant to indicate the nature of the fighting style attributed to humans and not the species itself after which they are named. Hawks always fight as hard and unrestrained as they can and retreat only when seriously injured. Doves on the other hand, only threaten and never attack. When a hawk meets a dove, the hawk always wins as the dove flees; when a hawk meets a hawk, the fight progresses until one is seriously injured; when a dove meets a dove, they go on threatening each other until one of them concedes, although the victor gains the resource of the contest both are penalized with a small cost for the time wasted in threatening. If we assign arbitrary values for calculating the net fitness for ease of comprehension as follows:

Win	Loss	Seriously injured	Wasting time
+50	0	-100	-10

The points can directly be ascribed to survival value as the individual with the highest net payoff will be able to leave more genes in the gene pool and thus more individuals with that specific strategy type in the subsequent generations. Thus, Smith worked out that if a population consisted of only one type of strategy say that of doves, the average net pay-off for each individual would be +15 in a 'dove-only' population, whereas it would be -25 for a 'hawk-only' population. It would appear to be the case that we would do far better if we had to mutually conspire to adopt a dove-only strategy and cooperate maximally. The reason we fail to achieve this high level of cooperation is due to the following evolutionary phenomenon.

In a 100% dove population, if a mutant hawk arises, the hawk will do much better than any dove as the hawk would win every fight gaining +50 points invariably without ever losing. Under this condition, due to the way natural selection works, hawk genes will see a rapid rise, and the population would oscillate toward a 100% hawk strategy. Once again if a mutant dove arises in this population, the dove would do better than any hawk fleeing from every fight with 0 points for losing each time but still doing better than any hawk whose average net payoff is -25. Thus, dove genes will continue to rise, and the population will

once again swing toward the dove strategy. This indicates that neither 'hawk-only' nor 'dove-only' is a stable strategy. Smith worked out that the ESS for this population would be 5/12 doves to 7/12 hawks. When this stable ratio is achieved, the average payoff for hawks is exactly equal to that of doves. This implies that both hawk and dove genes will be selected, and neither are exterminated from making it into the next generation. Therefore, the population continues at a stable strategy. It should be noted that ESS will be achieved in a manner that an individual will have the ability for both dove and hawk strategies and the net payoff is calculated presuming that they will opt to play dove in half of their fights and hawk in the other half. The average net pay-off at ESS is +6.25. One can extrapolate from this that every individual in the population would do better in the long run if they willingly resorted to a 'dove-only' strategy (net pay off +15) at all encounters, but due to the constraints ushered in by natural selection and individual's shortsightedness of the long-term benefits, the population rather functions at an ESS with evidently lower net payoff (+6.25). I discuss more on ways in which we could overcome this in chapters 4, and 6.

3.2 The caveat: hybrid model of morality in humans

Emerging research from promising fields such as cognitive sciences, moral psychology, and behavioral economics has invariably emphasized that humans demonstrably show a contradiction when making decisions, especially moral ones. That individuals do not always act according to the values they hold, that reason for holding a value doesn't entail that the necessary action will follow. This issue is often raised by moral psychologists from studies of human psychology and subsequent moral choices and behaviour (Haidt 2001). In this section, I argue that this view is compatible with both the evolution theory and the game theory model of social cooperation.

Michael Ruse has popularized the notion of the hybrid model of morality in humans resulting from our evolutionary history with selection pressure working towards a parsimonious yet efficient process for making strategic decisions on whether one must cooperate with another individual or not, upon each interaction. Even though humans are seeded with morality for their overall survival advantage, they are also faced with eternal, ongoing contemplations of how they must interact with other biological rivals, as argued in the previous

section. Most of the time, the conflict settlement between social instincts and self-preservation instincts is an unconscious, automated process taking place in the background with only part of it emerging in the inner dimension of experience in the form of moral intuitions and emotions that are built into our system. Even if these cost-benefit analyses and calculations are not always conscious, they are nevertheless crucial to an individual's fitness and survival, hence the mechanism continuously develops and evolves. In other words, we are hardwired to be moral for its function in survival, but we also can override this tendency (Ruse, 1986).

Taking account of evolution in other animals, we could have either developed a rigid model of morality as seen in social insects such as ants, known to be the highest form of biological altruism, through genetic hardwiring. They are machinelike, programmed to work on their innate dispositions triggered by chemicals with rigid automaticity that they inherit through their genetics. The biological advantages to this model are the efficiency, elimination of individual ants from the need for learning which reduces mistakes. However, the downside is that there is no flexibility whatsoever, if circumstances change, individual ants are unequipped to react. This may be a smaller risk to presume in the case of ants as they are cheaper to produce biologically but would fail for biologically expensive species such as humans. At the other extreme, we could have evolved to have 'superbrains' granting us the ability to work out the most accurate, best strategy possible injecting every detail into the 'cost-benefit calculations. In addition to consuming inordinate amounts of energy, such a system would also be mightily slow and would not have served us well in the background against which we have evolved. Therefore, we have been handed a middle-of-the-road model of biological altruism. On the one hand, we have developed to have a conscience supplied with intuitions and emotions[4], but on the other hand, we possess the flexibility to choose to override our innate programming supplanting it with alternate actions of our choosing. Ruse summarizes this attribute as *"we are inclined to behave morally but not predestined to obey the dictates of our conscience"* (Ruse 1986).

[4] unique to our species due to the scaling of intellect and the ability to process excessive amounts of information

The evolutionary account as put forward by Ruse is that ethics is a collective illusion of the human race, fashioned and maintained by natural selection in order to promote cooperation which in turn promotes survival. (Ruse 1986, 102)

To summarize, we indulge in cooperative and altruistic behavior as it tends to favor us in the longer run, and although we have moral capacity hardwired in us through genetic inheritance, the model of morality we have evolved to possess also makes it flexible for humans to override the innate mechanism for morality whenever it suits our fitness.

4. Enhance morality beyond survival to promote welfare

As established in earlier chapters, humans, along the evolutionary journey, have developed the capacity for morality merely because it helps us to cooperate and in turn increase our chances of survival. Morality evolved just like our eyes or feet for the selective advantage it ensues. It is an adapted feature that could get us to the point we are at, on account of it being economical by a) avoiding the enormous repercussions of blind action and b) cutting back on the expenses of operating a superbrain. Individuals that could cooperate well did better than those that couldn't, and hence altruistic individuals survived, reproduced, and passed on their genes. Altruism evolved facilitated by morality. (Ruse 1986, 99)

Models of altruistic evolution show that those individuals at either end of selfishness or selflessness will not have been naturally selected. The balance lies in being altruistic without exceeding the cost to self in the long run. 'Moral altruism'- literal selfless behavior can facilitate 'biological altruism'- help given to others for the benefit of self. (Dawkins 1989)

There was selection pressure towards i) kin selection- altruism to close kin and ii) reciprocal altruism- to those who can directly/indirectly benefit self. As a result of the benefits of cooperation, morality evolved in humans in a hybrid model: a combination of a) automatic- the capacity to internalize epigenetic rules and act based on genetic hardwiring like highly functional ant colonies and b) deliberate- the capacity to use conscious reasoning to override our innate settings. (Ruse 1986, 98)

To summarize, although we are hardwired biologically to cooperate with others, we also have the tendency (biologically) to act selfishly if we can get away with it (Boehm 2012). We also have the flexibility to choose whether to act upon this hardwiring or not, through reason and training. This theory of moral intuitions correlates with the dual-process thinking theorized by Daniel Kahneman. (Kahneman 2011)

4.1 Game theory based research on reciprocal altruism

Even when bound by the evolutionary constraint of survival and reproduction, with which natural selection limits our capacity for altruism, it is still possible to promote well-being and reduce suffering within reasonable amounts if we

cooperate and further enhance our capacity for morality. I lead to this idea using Robert Axelrod and W.D. Hamiton's research on a gambling game called 'Prisoner's Dilemma' in its iterated form (Axelrod, 1984). It demonstrates that a society of 'Nice' individuals that are cooperative, forgiving, non-envious, remorseful, and able to retaliate when mistreated has higher-order stability than a society filled with 'aggressive' individuals. This could be an extension as well as a justification to the notion of why we as a society should aim at promoting moral behavior and altruism. The means that propagating morality in this sense may strike one as unconventional, being at odds with what is traditionally held as moral attitude and behavior. However, it makes a compelling case for welcoming insights from evolutionary science into ethics which I discuss in chapter 5 but makes for an important note at this juncture.

Evolutionarily stable strategy (ESS) as already stated is a strategy, when adopted by a population in a changing environment cannot be overthrown by an alternate strategy. That is, no deviant individuals have the ability to benefit from such a strategy. For example, in a population of highly aggressive and highly passive strategies, neither are stable as the aggressive strategy will benefit and grow when the rest are passive, easily preying on them. As a result, aggressive strategies will continue to grow as they do well and reproduce. When the population thus becomes clouded by aggressive strategies, nice strategies do better, and the environment once again becomes populated by nice strategies and vice versa. In nature, this stability is achieved through ESS, for some time there are fluctuations, but once equilibrium is reached no new or alternative strategy can do better than the existing strategy. The existing strategy cannot be displaced (Smith 1982)

Axelrod and Hamilton identified that the 'Prisoner's Dilemma' is repeatedly played by many animals and plants in evolutionary time. In the original version, the game is played as follows. Two players play against each other and the banker being the neutral body settles and pays out the winnings to the players. The players have only two cards, 'cooperate' and 'defect'. The players ought to simultaneously choose a card to play without divulging their decision to the opponent. The banker reveals the cards and distributes the winnings based on the combination reached. The winnings are calculated as follows (arbitrary values):

Player A / Player B	Cooperate	Defect
Cooperate	Both A & B get +300	A gets +500, B gets -100
Defect	A gets -100, B gets +500	Both A & B get -100

Overall, it appears that the best bet for both players when unaware of the opponent's motivations is mutual cooperation as the winnings are respectable, eliminating the threat of the fine. However, opting for cooperating raises the stakes for both the players as it comes with the uncertainty of the opponent's intentions, which if against the player, could mean a much higher fine than had they to opt for defect. Thus, although 'cooperate' yields better results for both players, 'defect' appears as the most rational choice when the opponent's intentions are not apriori known.

Although it could appear that the most rational bet for a player is to defect in the single instance of the game, stemming from a lack of information of the opponent's intentions, the terrain quickly transforms when the games are iterated. It is as though the faculty for trust is integrated into the game with the repetition of the games between the same players indefinitely. They quickly learn that in the longer run, both players benefit more from cooperating (Greene 2013, 29-65). Despite this, to maximize one's total winnings from a game session, there are numerous strategies that the players can adapt to benefit from mutual cooperation and have a higher net payoff than their opponent.

Axelrod set himself the task to hypothesize which strategy would be the best. He advertised a tournament, inviting game theory experts to submit strategies to pick the winning strategy. He received fourteen strategies and added a random strategy which formed the baseline, against which the rest of the strategies were evaluated. All the strategies were paired against the rest in a round-robin fashion and the winning strategy was called 'Tit for Tat', it begins by cooperating on the first move and thereafter mirrors the opponent. This strategy was 'nice' and also could correct mistreatments by the opponents. Axelrod suggested another version of the 'Tit for Tat' called 'Tit for Two Tats' where the strategy begins by being nice and forgives defect by the opponent one time and thereafter repeats

their play. This strategy would have done even better than 'Tit for Tat', although it was not submitted in the first tournament (Axelrod 1989, 27-40).

Axelrod ran a second tournament, and the programmers were informed of the previous tournaments' conclusions. 62 strategies were received. Depending on the results of the first tournament, it seemed that there were equal amounts of both 'nice' strategies based on its success, and 'aggressive' strategies perhaps anticipating the environment to be filled with nice strategies. The competition was run and once again 'Tit for Tat' was the winning strategy. Even though 'Tit for Two Tats' was submitted this time, it couldn't do well in an environment filled with aggressive strategies, however since 'Tit for Tat' had the ability to retaliate without forgiving it did well and won once again. So, it seems that the effectiveness of a strategy depends on what strategies it has to encounter. That is, the environment and the strategies of the rest of the players, underline how well a strategy does and ultimately survives and gets passed on (Axelrod 1986, 48-54).

Axelrod tried to adapt this into an ESS by running the received strategies into a computer simulation with payouts in terms of reproduction, just as natural selection would. The strategies that did well survived and continued down generations (further rounds of the simulation) and after about 200 generations, equilibrium was reached and at this point, all the surviving strategies kept playing cooperate and even though it wasn't possible to distinguish them since both the player and the opponent kept cooperating, cooperation kept passing on into the generations that followed (Dawkins 1989, chapter 12).

Axelrod's tournaments give a key representation of how morality could have evolved due to its benefit for cooperation and survival. Forms of moral strategies that do not account for survival will not be selected and hence not passed on. Even though challenging, it should be possible to account for morality that promotes well-being and a good life whilst ensuring survival. Any overt form of morality that doesn't take evolution into account, even if promoting well-being and noble views of altruism, is not evolutionarily stable. Even if the environment itself is shaped to account for and promote high morals among individuals it would be vulnerable to clustering and overthrow by deviant individuals (Dawkins 1989, chapter 12).

Axelrod and Hamilton have conducted further research in this field and have remarked that even in an environment that reaches an ESS with any form of 'aggressive' strategy where players are always defecting, i.e., a highly aggressive environment, is susceptible to being taken over by clusters of individuals that play 'Tit for Tat'. If a few 'nice' individuals playing 'Tit for Tat' strategies continue to grow in this environment, they will continue flourishing and once a critical point is reached, the population starts shifting towards an ESS with 'nice' strategies as aggressive strategies no longer do well. The reverse is however not possible; therefore 'Tit for Tat' is a better strategy in terms of stability and doesn't face the threat of being taken over by aggressive strategies (Axelrod 1984, 27-205).

The essence of this chapter is to establish the relevance of game theory in understanding how morality evolved and why we form the norms and traditions as we typically do. This is to shape our societies so that majority of the individuals are aiming towards cooperation. Cooperation also has a second-order stability that aggressive societies lack and the net gain from cooperation is also higher than a society operating at an ESS or an aggressive strategy. Therefore, we should aim at promoting welfare and reduce suffering by working within the constraints of natural selection.

5. Integration of evolution theory into ethics

In this chapter, I discuss the interdisciplinary study of evolution and ethics. Several attempts have been made previously to derive an ethical framework directly from the evolution theory with drastic outcomes due to which the field of evolutionary ethics has generally received a negative response in philosophy and other fields (McCampbell 1983). I give a brief history of the field of evolutionary ethics, presenting the various ideologies, and proceed to make the claim that the naturalistic fallacy which prohibits the derivation of a value from empirical facts, also called as the fact-value gap, discussed in section 5.3, needs to be further challenged to not be a constraint on evolutionary ethics[5], but of providing a different outlook for it in the field of ethics. Therefore, it is essential to study the field for the value it can provide in understanding human motivations and behavior based on the way we have evolved and is crucial for the advancement of ethics.

5.1 Brief history on evolutionary ethics

There was notable enthusiasm for the field of evolutionary ethics in the nineteenth century which rose with the leap in enthusiasm for a scientific explanation for how conscience is a unique feature to humans and not explicitly demonstrated in other animals. However, towards the end of the century, the enthusiasm declined. At the time, the theory was not in discussion in its current form. That is, instead of the Darwinian form, two other forms were more popular and actively discussed.

First, was the Lamarckian evolution theory which was popular among evolutionists that pursued an ethical understanding. Lamarck had argued that evolution happens through characteristics that are acquired from the preceding generations. Thus, a parent could acquire certain traits intentionally such as physical strength, mental abilities, morality and so on which they then pass on to their offspring. It seemed to focus attention on effort and acquisition of moral traits and therefore appealed to the religious and ethically-minded individuals.

[5] the discipline that attempts to formulate ethical principles based on research from evolutionary sciences

This theory eventually was rejected by evolutionists due to it being based on concepts such as will and intent and not being based on reliable, empirical facts (McCampbell, 1983).

Second, was the cosmic evolution theory that was supported by those that believed that the process of evolution extends to the entire universe. Herbert Spencer argued that evolution was a cosmic law and belief in it promoted belief in progress. Theologians did not readily accept this theory as it took charge away from God and gave natural selection the sole authority (McCampbell 1983, 162).

The rise in evolutionary ethics branched into two schools of thought largely based on the cosmic evolution theory. The first kind were the advocates of the theory that sought to explain how humans alone came to develop a conscience in the evolutionary process and thus evolved as moral beings. The second attempted to derive a moral code from the process of evolution itself, thus giving rise to the notion that nature has a procedure and tendency to promote progress, and therefore humans should stand aside and let it work, i.e., no ethical intervention outside of nature was required to promote moral progress. As nature promotes the survival of the fittest, the fit is better and superior. This gave rise to the movement of social Darwinism which was spearheaded by Herbert Spencer, although not supported by Darwin himself (McCampbell 1983, 168). Under this movement, the practice of eugenics was encouraged, and hedonistic tendencies were promoted. Survival and happiness were likened to superiority as enforced by natural selection and therefore ethical. The central idea of the philosophy was that since natural selection is working for progress, humans should stand aside and let nature do its work. The inferior and unfit would be destroyed through the struggle and their incapacity for existence. That our obligation is to prize the strong and successful and to let the weakest go to the wall (Ruse 1986, 96).

This doctrine, however, was strongly refuted by T.H. Huxley who believed that our interactions with nature may have given rise to our conscience and capacity for morality but it surely wouldn't make us better humans (McCampbell 1983, 170). He used the difference between a cultivated garden and the wild brush in the forest as an example of the difference between ethical activity and nature. He compared the effort needed in cultivating a garden to the role of ethics in human progress which can be destroyed if abandoned just as weed tends to destroy a

garden when left unattended. Darwin himself believed that conscience was a product of the conflict between social instincts and self-preservation instincts. An individual is constantly faced with the challenge of taking care of one's needs versus making a sacrifice for the wellbeing of a fellow companion. And that any animal that could evolve to the point of possessing high intellect as humans do would be able to develop a conscience and therefore that this isn't a valid argument for allowing nature or evolution to drive our moral code (McCampbell 1983, 164-167).

5.2 Evolution is not progressive

I agree with Huxley in saying that evolution is not progressive[6] and therefore cannot be the basis for any normative or metaethical framework. Evolution branches out in several directions and does not follow a linear path of progression. It is impossible to offer true assessments of top and bottom, higher and lower, better and worse. For example, depending on the change in the environment which sometimes the organisms themselves impose upon their environment, they either progress into a better form from the previous one or they tend to evolve parallelly into different species which may or may not be better than their previous form. Similarly, sometimes they may have to shed some of their previous superior traits or adaptations if it no longer suits the environment and is not tenable for their survival. Among today's organisms, venereal disease thrives, whereas the great apes stand near extinction (Ruse 1986, 96). If one were to go by evolution-based metaethics, then gonorrhea would seem to be superior to the chimpanzee. When extrapolated to humans, such a metaethical premise would not be humane therefore not moral as we understand morality currently, because it would ask one to walk over the weak and the sick, the very young and the very old. It wouldn't improve our moral attitude but would create a society founded on aggression and competition, which is the exact opposite of the morality that has served humanity thus far in cooperation and coexisting reasonably, even if not throughout our history.

[6] progressive is used to denote advancement and improvement

5.3 Naturalistic fallacy and scope for evolutionary ethics

A common criticism to evolutionary ethics has been the naturalistic fallacy as put forward by G. M. Moore based on Hume's conclusion that one cannot go straight from talk of facts (evolution) to talk of obligations, i.e., one cannot go from using 'is' language to 'ought' language.

> In every system of morality, which I have hitherto met with, I have always remark'd, that the author proceeds for some time in the ordinary way of reasoning, and establishes the being of a God, or makes observations concerning human affairs, when of a sudden I am surpriz'd to find, that instead of the usual copulations of propositions, is, and is not, I meet with no proposition that is not connected with an ought, or an ought not. This change is imperceptible; but is, however, of the last consequence. For as this ought, or ought not, expresses some new relation or affirmation, 'tis necessary that it shou'd be observ'd and explain'd; and at the same time that a reason should be given, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it (Hume 1978, 469).

G. M. Moore backed this point in his Principia Ethica, arguing that all who would derive morality from the physical world stand convicted of the 'naturalistic fallacy'. As Teehan and DiCarlo have put forth in their paper, (Teehan, DiCarlo 2004, 38)

> The concern for those who oppose evolutionary ethics seems to be that if we allow evolutionary thinking into our ethics, we are going to end up with a reactionary moral system which supports an oppressive patriarchal value system in which women are consigned to the kitchen, the poor and disadvantaged to the fringes of society, the diseased and disabled to die, all in the name of the natural moral order. The naturalistic fallacy cuts off any such strategy by pointing out that simply because something has played a certain role in the evolution of the species it does not follow that it ought to continue to play that role, or that it can play no other role. How we ought to behave is a moral question which cannot simply be read out of the world of facts.

To proceed further, I analyze Moore's idea. He states that a simple value like good cannot be defined, as in defining it, one can ask the question whether the definition is good enough. He treats good as a supersensible property that is not definable. Thus, in defining the undefinable, one commits the naturalistic fallacy. It is important to note that the Naturalistic fallacy is committed not only when one takes factual claims and turns them into moral obligations on the basis of them being good, but also when one takes value claims, be it religious or philosophical, and attempts to turn them into moral obligations, based on them being good. In that sense, both the claims of the following nature would commit the naturalistic fallacy.

X is the natural purpose of Y; therefore, Y ought to do X

X is the representation of Ultimate Reality: therefore, X is morally correct

What Moore seems to imply is that there could be no objects, be it physical or metaphysical, out there, that could give us a justification and substantial evidence of values (good) as values (good) simply cannot be defined. Therefore, it is not possible to derive a moral code directly based on what is found in the physical world or through metaphysical thinking. However, in turn, this implication paves way for the idea that there are no concrete value objects out in the physical world that need to be found, but rather facts that serve as a guide for us to base our values on, which in turn stems out of our concern for the values we deem as the most crucial. Ruse also backs this idea of ethics and infers that there is no reasoned justification for morality and ethics based on a foundation from which it arises, to which one can appeal to while making moral judgments (Ruse 1986, 102)

This does not mean that there are no moral truths, it just means that we as humans get to collectively choose which values we ought to focus on based on the facts that the natural world presents. Values arise from the complex interactions of individuals with their environment and the phenomenon of valuing itself which comes from the evolutionary process that led us to act on the inputs and stimulus that nature provides us with and aids us with choosing mechanisms for survival and cooperation with others that may not value the same things as we do. *Morality is twofold, it affects as well as is the effect of a complex web of social interactions*. I claim that if we are needed to constantly act on the cues provided by nature it would seem reasonable to develop a code of conduct taking into account the way nature influences us. That is, if we both act and are being acted upon by nature, then deriving facts that may aid us in upholding values that we see as significant becomes a response to the way nature acts upon us, the latter also being a blind and non-rational process. For example, once we jointly agree that integrity is a value, we can take inputs from the real world in understanding the social or natural triggers that could deviate us from achieving this goal. It then would help us adopt a better framework to achieve our aim, i.e., the values we cherish. As suggested first by Aristotle through the doctrine of virtue ethics and more recently advocated by Ruse and other ethicists, morality needs to be seen as applied science (Ruse 1986), and from this, it naturally follows that any science that lets us learn more about our value judgment mechanism, our cognitive processes, emotions, and innate inclinations, and explains the history of our

moral intuitions has a lot to contribute to moral philosophy. Evolution is such a science and needs to be included to contribute to ethics rather than being barred.

As an illustration, in his article, "The Evolution of Cooperative Behavior and Its Implications for Ethics" (Morris 2009), Stephen G Morris makes the argument that individuals that can manage to maintain a positive affective state such as happiness have the potential to be more altruistic and therefore tend to be more ethical. He argues that a person's propensity to behave cooperatively is positively correlated with the degree to which they maintain a positive affective state (Morris 2009, 923). From an evolutionary standpoint, positive mental states such as happiness and contentment are good indicators of having one's basic needs met so individuals can behave altruistically without paying the cost that would decrease one's own fitness in behaving so, as altruism is not advantageous for survival especially when individual's basic needs are not met. Studies have shown a direct link between positive states and cooperation. Therefore, evolutionary studies can help ethicists concerned with improving social cooperation by giving insights into the science behind the factors that influence positive affective states in human beings. Evolution could help us understand the strategy behind the selection of negative affective states such as sadness, jealousy, hatred, fear and help us explore practical means to shape our social infrastructure to make it plausible to promote cooperation (Lewis 2011).

Morris provides an example for such a research space that could help public policymakers to create an environment in which there are fewer triggers for a negative response, the study of our tendency to derive our self-worth based on social comparison and others' evaluation of our social position. It served us in the past when living in hunter-gatherer bands to foster group identity (Boehm 2012), but in recent times has been reported to affect our psychological wellbeing. Although these tendencies are genetically rooted and are difficult to overcome, recognizing that these are also shaped by the environment can help us design environments that can curb these negative triggers and can help us get closer to meeting our ethical goals. In this endeavor, evolutionary science has major contributions to make.

In summary, the field of evolutionary ethics traditionally has received negative criticism for being based on a misconstrued account of evolutionary theory giving

rise to several serious social issues such as eugenics, sterilization of unfit individuals, hedonism, and so on. These were based on the archaic accounts of evolution and do not depict the advancement in science hence the field needs to be explored and given appropriate attention. I argue that although evolution doesn't give us any basis to directly translate it into an ethical framework, it can still inform us on the methods we could employ to resolve practical ethical issues that we aim to resolve. Understanding the evolution of our innate traits may help us better understand our moral dispositions and judgments and to help us make internal shifts as well as help us shape the environment (society). In applying ethics in this manner, one doesn't commit the naturalistic fallacy and is not violating any logical code. Evolutionary ethics, as a field, has a lot of significant contributions to make for the advancement of moral philosophy and ethics.

6. Cultural evolution and teaching ethics to promote moral behavior

In this chapter, I draw on Dawkins' hypothesis on 'cultural evolution' unique to humans which is a type of evolution similar to genetic evolution where instead of genes, certain entities, such as language, ideas are transmitted and evolve. I explain the concept and argue that cultural evolution can help us in achieving our ethical aims of promoting wellbeing and reducing suffering, as cultural evolution is much faster than genetic evolution and can be passed across individuals and not just between generations or related individuals helping us overcome the limitation with altruism being limited only to kin and close group members. There are several examples of cultural transmission such as language, art, science, technology, food, etc. and I propose that we consider teaching ethics as a type of cultural transmission and use it to promote ethical values and goals. However, since there's always the dependency of limitations posed by genetic evolution, we need to understand the co-dependencies of evolutionary and cultural evolution.

Dawkins argues that culture is unique to our species and that cultural transmission being similar to genetic transmission gives rise to another kind of evolution. The unit of cultural transmission is 'meme'. Examples of memes are ideas, tunes, catchphrases, clothes, fashion, ways of making pots, or of building arches. Memes also propagate themselves in the meme pool by leaping from brain to brain through a process called imitation. For example, a scientist who hears about a good idea passes it on to his colleagues and students, mentions it in his work, and talks. If the idea is good, it spreads from brain to brain. Brains can be parasitized by planting memes in them, and brains act as the vehicle for meme's propagation (Dawkins 1989, chapter 11).

The meme is realized not just metaphorically, but physically millions of times over as a structure in the nervous systems of individuals in the world. He gives the example of 'the God' meme, even though it's not known how the idea originated, it could have originated several times via several individuals, much like independent mutation and this meme replicates itself with the help of other memes such as profound language (speech, written word), great music and great art. Not only has the meme survived for a very long time and has remained stable

but has also deeply penetrated human culture and plays a significant role in an individual's lives and identity. The reason for this is the meme's great psychological appeal. It gives us persuasive answers to existential questions that can otherwise be unsettling and disturbing, suggesting that injustices in this world may be rectified in the next. Just as the doctor's placebo is sometimes effective despite being imaginary, so is the effect of the 'God meme' in providing comfort and easing out emotional pain. These he says are some of the reasons why the idea of God is replicated so easily by successive generations of individual brains. "God exists, if only in the form of a meme with high survival value, or infective power, in the environment provided by human culture" (Dawkins 1989, 250). Dawkins expands further on the idea stating psychological appeal means having appeal to brains and brains are determined by the natural selection of genes in the gene pool. Thus, perhaps there are genetic advantages to having a brain like this as it improves gene survival.

There are similarities between a meme and a gene which can also be the constraints that affect a meme's transmission, speed of replication, longevity, and stability. Similar to the competition in the gene pool to occupy a body and get passed on to the next generation, memes also compete in the meme pool to occupy space and get spread across brains. Memes that are idea memes, that is the memes that are capable of being transmitted from brain to brain, compete with the rival memes to occupy the attention and focus of the brain they are planted in. They also could compete for space on the internet, media, news, books, museums, servers, databases, or any other such information holding platforms.

Just as genes that cooperate well with each other form gene units that produce well-functioning bodies that can survive longer, memes also form meme units by forming a cluster with memes that foster each other. For example, the 'God' meme, goes well with the 'hell' meme as they both have a deep psychological impact and reinforce each other, that is, they are mutually exclusive giving rise to stronger memes, such as the 'blind faith' meme or the 'religion' meme. Some memes can transmit rapidly but will not survive longer, while others may have existed for long without significant acknowledgment but see a spike in

transmission suddenly, due to another meme or a cultural event triggering their relevance and making them significant.

Just like the natural selection of genes and forming of gene complex is a blind, non-rational process passed on merely by the ones that manage to survive and replicate, so are the memes that can promote survival of the bodies they inhabit and can stay relevant in the brains that they have been planted. Other memes can collaborate with idea memes to form meme complexes that foster the idea and collectively survive and stabilize and are replicated through the generations. For example, religious institutions make use of art, language, music, costume, rules, rituals, architecture, written and verbal memes to coexist as a collectively stable set of mutually assisting memes.

Memes and genes often reinforce each other but may sometimes oppose each other and yet the meme can continue to exist if the value of the meme in fostering other memes in its meme complex and therefore overall survival of associated genes in the gene pool is high. This is also possible if the time spent actively transmitting and fostering a meme in the brain is very high such that no other memes are allowed to share that space. “Selection favors memes that exploit their cultural environment to their advantage.” For example, a gene for celibacy will never succeed in the gene pool however celibacy can be fostered through the idea meme of ‘celibacy’. As priests spend a lot of time transmitting the idea of celibacy to young boys using other memes in the meme complex that can be quite powerful, and that do not usually allow opposing memes to take space, the idea of celibacy gets transmitted. As this meme is an important part of the ‘religion’ meme complex, at least for some faiths, it tends to consistently get transmitted despite it not creating bodies that will pass on their genes. This is an example of how the use of memes can supersede the constraints of survival and reproduction that genetic evolution entails. However, it must be noted that memes have their own constraints that need to be adhered to. This could also be a result of biological evolution which could indeed be accidental and possible only in humans due to us having evolved to have high intellectual capacity when compared to other species. The exchange is so long as genes can create brains that can be developed and have the capacity to be cultivated with memes, memes carry out the work further of keeping the body alive, ensuring the survival and further

replication of the individual genes in the gene pool. Overall, there can be several combinations for the complex cross evolution of genes and memes, which needs further research and can have a significant impact in helping humans overcome the genetic constraints to promote our ethical goals.

An advantage of memetic evolution (evolution through memes) is that it renders us with conscious foresight, the ability to look ahead and correct our present actions. This tendency keeps improving in humans as can be inferred through the course of our human evolution. As discussed in the third chapter, that all individuals would do far better in terms of fitness through a 'dove-only' strategy rather than the strategy rendered through ESS, humans' tendency to not able to forgo short term goals for long term fitness curbs us. Through our ability to forecast our future, which keeps improving as we evolve, we could make use of culture to help us enter into pacts such as in the example of 'dove-only' strategy from chapter 3. We can understand our own biological and psychological shortcomings, renegotiate, and adapt to collaborate universally through the use of cultural transmission and evolution.

In the following section, I propose one such suggestion, to invoke ethics teaching as a means to surpass our biological and psychological limitations, to achieve a more morally adapt society. As the constraints of genetic evolution still apply, this effort can be considered as a cross between genetic and memetic evolution.

6.1 Ethics-teaching as a meme

Throughout the thesis, I have addressed the evolutionary origins of morality providing an overview of how our moral emotions and intuitions[7] have evolved as a result of their function in facilitating social cooperation. This trait ultimately was retained through the evolutionary process of natural selection due to its benefits for promoting survival. I made the following two claims which I reemphasize as these are essential for the aim of this chapter: 1) even though we inherit our moral behavior and attitude through our genes, we cannot fully rely upon it as it's also in our nature to promote selfish behavior for personal fitness

[7]Moral emotions and intuitions are defined as the unconscious appearance of a moral judgement including an affective valence such as good, bad, shame, guilt etc in the consciousness of an individual without any conscious reasoning.

when suitable opportunities arise. We are altruistic as long as we gain more from it in the long run, achieved by calculations weighing risks against benefits. This process can be both conscious as well as unconscious and the tendency to carry out these calculations is innate in us. 2) We need to adapt and enhance our moral behavior and utilize our capacity for morality beyond mere survival as we fare better by living altruistically as discussed in chapter 4, cooperating through a code of moral conduct by defining our aims and moral values. We could aim at a society with increased cooperation and less suffering if we expand on our moral nature as discussed in chapter 4.

In this section, I make the claim that the constraint the first point presents us in achieving the second point, can be tackled through culture by teaching ethics to some degree. I however do not claim that ethics teaching is the definite or guaranteed solution.

The argument I put forward, is as follows:

1. Morality evolved as it is an important trait that facilitates social cooperation and therefore survival.
2. Even though we are moral by nature, and have innate moral intuitions and emotions, we also have the innate tendency to override our moral intuitions to promote selfish gains.

   Therefore, our innate moral intuitions are not sufficient to promote ethical behavior.
3. Both living altruistically following moral values; and living aggressively can form an ESS (evolutionarily stable strategy) in a given population.
4. The former has higher-order stability in terms of ESS (evolutionarily stable strategy) as it cannot be overthrown by an organized change by the latter, however, vice versa is possible.

   Therefore, we would do much better in the long run by cooperating and living altruistically than if we choose the ESS that is formed through aggressive behavior.
5. There are constraints to practicing moral behavior as described in 1 and 2.
6. These constraints can be overcome through culture, teaching ethics and making individuals aware of the limitations, and helping them acquire the skills to practice moral behavior.

Therefore, teaching ethics is beneficial and could be effective in promoting moral attitude and behavior among individuals.

I have mentioned in this chapter that cultural evolution happens much like genetic evolution and we could bank on this phenomenon to propagate the teaching of ethics in the form of memes. For this endeavor, we would need to study the design of memetic evolution and its dependency on genetic evolution in order to devise memes that help us achieve our ethical aims.

This study needs to be done in collaboration with evolutionary biologists and game theorists but would be beyond the scope of this thesis.

7. Conclusion

In this thesis, I argued that morality in humans is the result of an evolutionary adaptation of altruistic behavior for the survival value it offers. Biological altruism has equipped us with a range of psychological traits that enable us to cooperate in a social setting. This in turn has made us adept at engaging in moral behavior. Although our moral intuitions and emotions are innate in us, due to the need for balancing off our self-preservation instincts against our social instincts, we have also biologically evolved the flexibility to override these natural predispositions, and are able to manipulate the situation to our advantage whenever we stand to gain from being non-altruistic. This is achieved through a cost-benefit analysis mechanism that occurs in the human mind which could be conscious or unconscious. Due to this, relying merely on our moral intuitions does not suffice to make us moral, it needs to be complemented by culture and teaching of ethics.

Next, I argued that the field of evolutionary ethics has come to acquire the reputation of being notorious and deleterious and is generally not well received. This criticism is based on a primitive, inadequate account of the evolution theory and the field itself. Both the theory and the field have significantly progressed since the nineteenth century and have corrected for the gaps in the scientific method and philosophical research. In its current form, the field has significant insights to offer to the understanding of human attitude and behavior.

I also used a game theory model produced by Robert Axelrod and W.D. Hamilton to justify why it is better for humans to cooperate in the long run than indulge in aggressive behavior. Their model shows how although both a society formed on a combination of 'aggressive' as well as 'nice' strategies has the potential to form an ESS, the latter has the capacity to overthrow the former through a process called clustering while the vice versa is not possible. The society compromising of individuals exercising 'nice' strategies has second-order stability that the 'aggressive' one lacks and the individuals in the former society have a higher net payoff than individuals operating at an ESS. I present this reasoning as my argument for why we ought to collectively strive for a society founded on morality that attempts to promote well-being and reduce harm and suffering since it bodes well for us in the long term.

Finally, I elicited that there's a secondary form of evolution that has emerged in the form of cultural evolution, whose unit of inheritance is 'meme' (art, language, ideas, etc.) and occurs much like genetic evolution by design and process. Memes impose an imperative on the genes to produce individuals that are capable of being vectors for their propagation and hence there is coevolution that occurs between genetic and memetic evolution. I suggested that we could tap into this possibility, to teach ethics utilizing the various memes available in the form of philosophical ideas, language, art, music, to shape genetic evolution and expand our moral landscape.

8. Abstract

Recent research from the fields of evolutionary biology, game theory, cognitive sciences, and anthropology claims that human morality is a result of evolution and the selective advantage of cooperation among individuals in their survival and fitness. Although humans have the innate propensity to cooperate and thus be moral, they are limited by the tendency of promoting self-fitness whenever possible. This hybrid model of morality restricts us from cooperating in the long-term when the benefits of not cooperating in the short term appear more appealing. Also, our innate morality only equips us to cooperate with kin and ingroup members and fails when we are required to extend it to members outside one's group. In this thesis I argue that 1) we should aim for increased cooperation as a society since it benefits us in the long run, 2) evolutionary ethics has significant insights to offer for the expansion of our moral behavior, 3) cultural evolution can be a suitable means for us to achieve our ethical endeavors overcoming the constraints posed by genetic evolution.

9. Bibliography

Darwin, Charles. (1859). On the origin of species, John Murray, London.

Dawkins, Richard. (1989). The Selfish Gene, Oxford University Press, New York.

Boehm, Christopher. (2012). Moral origins: The evolution of virtue, altruism, and shame, Basic Books, New York.

Greene, Joshua David. (2013). Moral Tribes: Emotion, Reason, and the Gap between Us and Them, The Penguin Press, New York.

Kahneman, Daniel, (2011). Thinking, Fast and Slow, Farrar, Straus and Giroux, New York.

Ruse, M. (1986). Evolutionary ethics: a phoenix arisen. Zygon, 21(1), 95–112.

Smith, Maynard (1982). Evolution and the Theory of Games, Cambridge University Press, New York.

Pojman, Louis P. Ethics. (1990). Discovering Right and Wrong, Wadsworth Pub. Co., Belmont, Calif.

McCampbell, Duane. (1983). The Development and Failure of the 19th Century Evolutionary Theory of Ethics. Restoration Quarterly, vol. 26, no. 3, pp. 161–171.

John Teehan, Christopher DiCarlo. (2004). On the Naturalistic Fallacy: A Conceptual Basis for Evolutionary Ethics. human-nature.com/ep – 2004. 2: 32-46

William S. Lewis. (2011). Evolutionary Psychology in the Service of Moral Philosophy: A Possible Future for Ethics?. The Journal of Speculative Philosophy, vol. 25, no. 1, 2011, pp. 48–63.

Stephen G. Morris. (2009). The Evolution of Cooperative Behavior and Its Implications for Ethics. Philosophy of Science 76, no. 5: 923.

Gould, S. J. (1999). Rocks of Ages: Science and Religion in the Fullness of Life, The Library of Contemporary Thought, The Ballantine Publishing Group, New York.

Axelrod, R. (1984). The evolution of cooperation, Basic Books, New York.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. Psychological Review, 108(4), 814–834.

Moore, G. E. (1903). Principia Ethica, Cambridge University Press, Cambridge.

Hume, D. (1740). A Treatise of Human Nature, P.H. Nidditch, ed. 2nd edition, Oxford University Press, Oxford.

Non-exclusive licence to reproduce thesis and make thesis public

I, Meenachi Naidu,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to

reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

Relevance of Evolutionary Biology and Game Theory in Ethics,
supervised by Kadri Simm

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons licence CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Meenachi Naidu
15/05/2021