(An Excerpt from **Evolution and Theology of Cooperation**, a research proposal put forth by Professor Sarah Coakley and Professor Martin Nowak)

EVOLUTION AND THEOLOGY OF COOPERATION

Sarah Coakley[1] and Martin A. Nowak[2]

Executive Summary

We propose to study the emergence of altruistic behavior, forgiveness and unselfish love in the context of biological, ethical and theological considerations. We will outline how basic principles of natural selection can lead from competition to cooperation, from conflict to coexistence. The mathematics of evolutionary game theory represents an appropriate tool to study the competition and cooperation of individuals adopting various strategies and phenotypes.
We will discuss the implications of winning strategies from a teleological perspective. This research represents a newly-conceived attempt to understand the evolutionary biology of a world created by God. Studying the principles of evolution from this perspective will generate insights into universal laws guiding the progress of life through space and time. The proposed research project represents a contribution to the fascinating, ongoing dialogue between science and religion that has been largely initiated by the Templeton Foundation.

As for specific goals and objectives:

1. We will establish a new, meaningful dialogue between evolutionary biology, ethics and theology, based on a detailed analysis of the significance of game theory for our respective disciplines.
2. We shall initially explore the evolutionary and theological aspects of moral behavior and ethical choices in the context of reciprocity and fairness, 'cooperation' and 'selfishness'. From here we shall unfold a range of more nuanced theological questions in connection with game theory (see intra), culminating in a set of hypotheses about the relation of God to evolutionary processes.
3. In the first year we shall hire two Post-doctoral Scholars to work with Martin Nowak on scientific aspects of the project, and to initiate interdisciplinary discussion with Sarah Coakley and her team in ethics and theology. The whole Harvard group will

[1] Edward Mallinckrodt 'Jr.' Professor of Divinity, Harvard Divinity School, 14 Divinity Avenue, Cambridge MA 02138.

[2] Harvard Professor, Department of Organismic and Evolutionary Biology and Department of Mathematics; Faculty Director, Program for Evolutionary Dynamics, One Brattle Square Suite 6, Cambridge MA 02138.

meet once per week in term-time to discuss progress and gather shared materials. As of year two we will support four Post-doctoral Scholars, two from ethics and theology, and continue regular meetings with the full research team.

4. In the spring semester of the first year we shall invite one ethicist (Timothy Jackson) as a Visiting Scholar in the program, and in the second full year we expect an established expert in science and religion (Philip Clayton) to be in residence as a Visiting Scholar at Harvard as a member of the research team.
5. We will aim to establish a new, enduring network of people at Harvard who are interested in this particular research and, more generally, in the interaction between science and religion.
6. We will engage graduate and undergraduate students to work on specific aspects of our research projects.
7. We will convene conferences and a seminar series to facilitate collaboration among a wider network of scientists, ethicists and theologians, including visiting speakers who will be invited to give presentations and to take an on-going part in our work.
8. We shall publish at least one book and several research papers in leading refereed journals. We shall also disseminate our findings in the press and in more popular journals (see intra). We shall expect our post-doctoral assistants to work towards producing a textbook under our guidance, and we shall seek to organize publicity sessions on our research through the learned societies to which we belong.
9. The proceedings and published work we produce will help to establish an international network of researchers working on the interface between evolutionary biology and theology.

Proposal Narrative

Evolution of Cooperation

Natural selection is the consequence of competition between reproducing individuals or groups. Central to all of biology is the question of how natural selection can lead to cooperation. Genes cooperate in cells, cells cooperate in organisms, and individuals cooperate in societies. Altruistic interactions among animals and humans are frequent and require specific explanation in terms of Darwinian evolution. Darwin himself realized that altruistic phenomena represent a major challenge to evolutionary theory; but debate continues amongst contemporary evolutionary theorists as to whether evolution will invariably select against genuinely altruistic behaviors (Clayton and Schloss eds. 2004). The research project that we propose will subject this question to a new and rigorous investigation, bringing together the interdisciplinary insights of game theory, ethical theory and theology.

There are three classes of mechanisms that lead to the evolution of cooperation:

1. Kin selection can explain cooperation among individuals that are genetically related. Suppose a particular gene induces altruistic behavior towards other individuals. The donor of the altruistic act pays a cost, c, while the recipient obtains a benefit, b. The currency of this interaction is fitness (reproductive success). Such a gene is favored by natural selection if the cost to benefit ratio, c/b, is less than the coefficient of relatedness, r, among individuals (Hamilton 1998, Hamilton 2002, Frank 1998, Keller 1999).

2. Group selection can lead to cooperation among individuals if not only individuals but also groups reproduce and are under selection. In every group, defectors have a greater fitness than cooperators, but groups that consist mostly of cooperators have a greater fitness than groups that contain mostly defectors. This effect can lead to the emergence and persistence of cooperation provided that individuals cannot move too rapidly between groups and that cooperators do not 'mutate' too often into defectors (Sober and Wilson 1998, Wilson 2002, Wilson 1999, Wilson and Sober 1996, Wilson and Dugatkin 1997, Wilson and Sober 1994a, Wilson and Sober 1994b, Wilson 1990, Wilson 1983, Wilson 1975, Williams 1992).

3. Reciprocity can lead to cooperation among unrelated individuals in the absence of group selection. Reciprocity comes in three flavors. Direct reciprocity means that you help somebody who might help you. Indirect reciprocity means that your cooperation will be returned not by the recipient but by another individual. Spatial reciprocity means that neighbors help each other (Axelrod and Hamilton 1981, Axelrod 1984, Alexander 1987, Sugden 1986, Nowak and Sigmund 1992, Nowak and Sigmund 1993, Nowak and Sigmund 1998, Nowak and May 1992). Biologically, one has to specify the mechanisms or functions of these behaviors in such a way as to explain why detrimental effects to the individual do not select against the behavior over time. One can also consider the possibility that some altruism is pleiotropic and selected for other reasons. We shall give attention to these continuing areas of debate.

The Prisoner's Dilemma (PD) is a game theoretic approach describing the interaction among cooperators and defectors. It is formulated as a game between two people. If both cooperate, then both receive as payoff the reward, R. If both defect, then both receive the punishment, P. If one person cooperates while the other defects, then the defector receives the highest payoff in the game, the temptation, T, while the cooperator receives the lowest score, the so called sucker's payoff, S. The PD is defined by the ranking T>R>P>S.

In the non-repeated PD, defection dominates cooperation. Since T>R, it is best to defect against a cooperator. Since P>S, it is also best to defect against a defector. Defection is a strict Nash equilibrium or an evolutionarily stable strategy (ESS). If everybody defects, than natural selection cannot favor the emergence of cooperation. Moreover, defection is also an 'unbeatable strategy': for any ratio of cooperators and defectors in a population, defectors have a higher fitness.

We will derive a classification of mechanisms leading to cooperation in terms of their moral and ethical implications. We will discuss spiritual and religious aspects of

unselfish behavior and see to what extent they can be grounded in evolutionary biology. We will explore the fundamental structures of evolutionary systems that are required for the emergence of higher level morality. We will then generate a range of possible ethical, meta-ethical and theological explanatory systems which can be brought into creative interaction with current evolutionary theory, and attempt a nuanced adjudication between such systems.

Direct reciprocity

We propose to begin by studying direct reciprocity. In the repeated PD, two players interact more than once. The cooperation of one player can be reciprocated by the cooperation of the other player in subsequent moves. A strategy in the repeated PD is a device for deciding to cooperate or to defect given any history of the game, consisting of a sequence of cooperative and defective moves. For an unending PD, there are infinitely many strategies; most of them are of enormous complexity. We will explore simple strategies that lead to relevant insights.

The two simplest strategies are 'always cooperate' (AllC) and 'always defect' (AllD). Clearly, AllC is dominated by AllD. For any mixture between AllC and AllD individuals, the latter receive a higher payoff and have higher fitness.

Tit-for-tat (TFT) is a strategy that cooperates on the first move and then does whatever the opponent did in the previous round. Two TFT players will cooperate in the absence of mistakes, while TFT facing AllD will only be exploited in the first round and will defect from then on. If the number of rounds is sufficiently large, then TFT playing TFT receives a higher payoff than AllD playing TFT. Vice versa, AllD playing AllD receives a higher payoff that TFT playing AllD. Therefore, in infinitely large populations, neither strategy can invade the other. This makes cooperation evolutionarily stable, but demands the question: how can cooperation emerge in the first place? It turns out that in finite populations, natural selection can favor TFT invading and replacing AllD under certain conditions.

In the presence of mistakes, TFT is weak. An accidental defection by one TFT player will be reciprocated by the other TFT player, which in turn will trigger a defection from the first player. TFT is unforgiving. In a natural setting where errors are possible, TFT is dominated by another strategy called generous tit-for-tat (GTFT). This strategy always starts with cooperation, always cooperates when the opponent has cooperated in the previous round, but sometimes also cooperates if the opponent has defected. Unlike TFT, GTFT can forgive defection and therefore correct occasional mistakes that might arise in form of unintended defection or misinterpreted cooperation.

Both TFT and GTFT, however, can be replaced by the strategy 'always cooperate', AllC. If everybody in a population plays GTFT and some people play AllC, then nobody has a disadvantage. These strategies are neutral variants of each other. Random drift can increase the abundance of AllC players and even lead to the extinction of GTFT.

If most people use AllC, then AllD can invade again. Thus, we have run a full circle from AllD to TFT to GTFT to AllC and back to AllD. These oscillations between

cooperation and defection (between light and darkness, we might say – although this is to beg normative questions that we shall consider critically in our research program) are an essential part of the struggle. Cooperation is never stable. There is always rise and fall, hope and despair.

Another simple strategy, which plays an interesting role in the repeated PD, is win-stay, lose-shift (WSLS). This strategy always starts with cooperation. Thereafter, it repeats its last move when it was satisfied with its payoff (when it made either R or T), whereas it makes the opposite move when it was not satisfied (receiving either P or S). Two WSLS players cooperate almost all the time, and correct occasional mistakes. In contrast to TFT or GTFT, however, WSLS is not neutral with AllC. It will detect that AllC does not reciprocate an accidental defection and from then on defect all the time. WSLS exploits AllC, but thereby stabilizes cooperation in the population.

If we were to rank all strategies mentioned so far according to increasing levels of morality, we may be tempted to propose: AllD< WSLS<TFT<GTFT<AllC. However, for the moment we here only hypothesize some ethical conclusions which our program will discuss from a number of different ethical and meta-ethical perspectives.

Clearly AllD has lowest morality, and the forgiveness of GTFT is nobler than the harsh retaliation of TFT, but all other rankings are problematic. WSLS seems to be worse than TFT because of its property of exploiting AllC. On the other hand, two WSLS players can achieve a higher level of cooperation than two TFT players, because WSLS has an inbuilt mechanism of forgiving defection. Likewise a population of WSLS players can ensure stable cooperation for a longer time period than a population of GTFT players, because the latter allow AllC to spread, which ultimately leads to a total breakdown of cooperation once AllD comes back. Unconditional cooperation (AllC) is in some experiments with humans considered as cheating, because it allows exploitation. But the morality of this matter is one worthy of further debate.

We would also like to study the repeated Prisoner's Dilemma in the context of group selection. How do levels of forgiveness and retaliation evolve under individual selection versus group selection? How does status within a group and moral reputation affect the strategy of direct reciprocity? This leads us directly to our next section.

Indirect reciprocity

Indirect reciprocity is the idea that altruistic acts are not returned by the recipient but by some other individual from the population. Interactions between individuals may not be repeated often enough for direct reciprocity to emerge. Furthermore, individuals may find themselves in very asymmetric roles when it comes to reciprocate help. The injured traveler who barely survived the robbers' attack might never have had the opportunity to reciprocate the Good Samaritan's help. Indirect reciprocity can lead to the evolution of cooperation if reputation is important. Helping is costly but purchases a good reputation and increases the probability of receiving help.

There are various alterations of this theme. A potential donor might base a decision to help only on the reputation of the recipient. In evaluating the reputation of the recipient the donor scores how often the recipient helped somebody else. Other strategies take into account their own reputation: if their own reputation is very low, it might be

best to help unconditionally; if their own reputation is very high, it might be unnecessary to help this time (Nowak and Sigmund 1998). Further complications arise when players take into account whether a potential donor was supposed to help a recipient. For example, if a donor refuses to help a recipient, because the recipient's reputation is low, then this refusal might not impact the reputation of the donor.

These considerations lead to a variety of 'assessment modules'. Suppose for the moment that reputation (or an incremental change in reputation) is described by a single binary parameter, 'good' or 'bad'. A donor may or may not help a potential recipient. The observer will update the reputation of the donor based on this event. 'Scoring', 'standing' and 'judging' are three different assessment modules:

	Scoring	Standing	Judging
Good helps Good	good	good	good
Good helps Bad	good	good	bad
Gad helps Good	good	good	good
Bad helps Bad	good	good	bad
Good does not help good	bad	bad	bad
Good does not help bad	bad	good	good
Bad does not help good	bad	bad	bad
Bad does not help bad	bad	bad	bad

Scoring gives a good reputation to the donor if and only if the donor has helped (has performed an altruistic act). Standing allows the donor to remain in good reputation even if the donor has refused help to someone in bad reputation, who did as a consequence not deserve help. Thus, standing attempts to distinguish between unjustified defections and punishment. Judging differs from standing in so far as it does not allow the recipient to help someone who does not deserve it. Thus helping someone who is in bad reputation is seen by judging as cheating on the system. The game becomes especially interesting with imperfect information about the reputations of others and attempts to deceive about one's own reputation and the reputation of others.

We will study game theoretic aspects of such strategies and evaluate the type and level of morality, and the motivation of strategies, in the various cases of indirect reciprocity.

Cooperation in groups

Most work on cooperation and defection is formulated in terms of interactions between two individuals. The Public Goods Game is an extension of the PD to more than two people. Each person has the option to make an investment into a common pool which then gets multiplied by a factor greater than one and distributed to all individuals of the group irrespective of whether or not they have made a contribution. The temptation is to free ride on the contribution of others.

Let us define cooperators as individuals that provide a benefit to the members of the group at some cost to themselves. In contrast, defectors attempt to exploit the common enterprise, avoid the costs of helping and provide no benefit. Suppose that groups of cooperators and defectors are formed by binomial sampling from a large population. In the limit of an infinitely large population, we can use the deterministic replicator dynamic (Taylor and Jonker 1978, Hofbauer and Sigmund 1998, Hofbauer and Sigmund 2003). We will consider synergy or discounting of the contributions of cooperators. For example, if only one cooperator is present in the group the benefit to each member of the group member is b, but if two cooperators are present, then the benefit is $b(1+w)$. If $w>1$ the contributions of the two cooperators are synergistic. If $w<1$ the contributions are discounted. With these basic assumptions, we can show that the interaction between cooperators and defectors can lead to situations where cooperators coexist with defectors or even dominate them. This is particularly interesting for the present proposal where we wish to find games that can be 'won' by unconditional cooperators.

We will explore the moral and theological dimensions of this setting.

Finite populations

Most work on evolutionary game dynamics is formulated for deterministic dynamics of infinitely large populations (Maynard Smith 1982, Hofbauer and Sigmund 1998, Fudenberg and Tirole 1991, Weibull 1995). Here we propose to develop an evolutionary game theory for finite populations. As mentioned before, in finite populations, AllD can be replaced by TFT under positive selection. More precisely a single individual using TFT can invade and take over a population of AllD players with a probability greater than the reciprocal of the population size, which is the corresponding fixation probability of a neutral mutant (Nowak et al 2004).

Evolutionary game dynamics in finite populations can be modeled by a classical Moran process (Moran 1962) with the extension of frequency dependent fitness. There is a population of N players consisting of i players that use strategy A and N-i players that use strategy B. At any one time step of the discrete stochastic process, an individual is chosen for reproduction, at random but proportional to fitness. An offspring is being produced, which replaces a randomly chosen individual from the population. Thus the total population size remains strictly constant. The stochastic process is a birth-death process with two absorbing states: all-A and all-B. We want to calculate the probability that a single A individual will take over the whole population. If this probability is greater that 1/N then selection favors the replacement of B by A (Nowak et al 2004).

We want to explore many properties of evolutionary game dynamics in finite populations. At present we are trying to understand the evolutionary oscillations of AllD, AllC and TFT. In a finite population, in the presence of mutation, it is possible that the time average of these oscillations is almost totally concentrated on TFT although AllD is the only strict Nash solution (the only ESS) in this setting. Finite population size can favor cooperation in unexpected ways. We would like to develop a comprehensive understanding of such observations.

Learning cooperation

Learning cooperative behavior in particular and moral systems in general is not a trivial task. There is much work on cognitive and computational aspects of learning human language. We propose to develop a similar approach for learning altruistic behavior. We will explore the case for the logical necessity of an innate moral grammar. We will combine learning approaches in computer science (Gold 1967, Vapnik and Chervonenkis 1971, Valiant 1984, Nowak et al 2002) with learning approaches in game theory (Fudenberg and Levine 1998).

Human language and cooperation

There are various evolutionary approaches to understanding properties of human language (Pinker and Bloom 1990, Bickerton 1990, Jackendoff 2002, Nowak and Krakauer 1999, Nowak et al 2001, Nowak et al 2002, Nowak et al 2000). Humans and chimpanzees separated some 5-7 million years ago. Chimpanzees have a complex system of a conceptual understanding, and they have rich social interactions, but they are limited in the sorts of information they can exchange. They do not have a communication system comparable to human language. The central question of the origin of human language is which genetic modifications led to changes in brain structures that were decisive for the ability to generate human language. Given the enormous complexity and specialization of human language, we should expect several modifications and incremental steps guided by natural selection. The structure of existing human language also enables us to speculate about the progression of stages from simple symbols to hierarchical structures and complex grammar. At the same time, we should keep in mind that evolution will have re-used cognitive features that evolved long ago and for other purposes.

The general goal of studying language evolution is not only to reconstruct the sequence of historical events that led to human language, but to formulate a theoretical framework explaining how Darwinian dynamics yield fundamental properties of human language such as arbitrary signs, lexicons and syntactic communication. The aim is to design a program for how formal language theory and learning theory can be brought together with the theory of evolutionary dynamics.

Formal language theory begins with the definition of an alphabet as a set containing a finite number of symbols. Possible alphabets for natural languages would be the set of all phonemes or the set of all words of a language. For these two choices one obtains formal languages on different levels, but the mathematical principles are the same. A sentence is defined as a finite string of symbols. A language is a set of sentences. Among all possible sentences some are part of the language and some are not. A finite language contains a finite number of sentences. An infinite language contains an infinite number of sentences. There are infinitely many finite languages, as many as integers. There are infinitely many infinite languages, as many as real numbers; they are not countable. Hence, the set of all languages is not countable.

A grammar is a finite list of rules specifying a language. A grammar is expressed in terms of ‘rewrite rules’, which are of the form: a certain string can be rewritten as another

string. Strings contain elements of the alphabet together with so-called ‘non-terminals’, which are place holders. After iterated application of the rewrite rules the final string will only contain symbols of the alphabet. There are infinitely many grammars, but only as many as integers: any finite list of rewrite rules can be encoded by an integer. Since there are uncountably many languages, only a small subset of them can be described by a grammar. These languages are called ‘computable’.

There is a correspondence between languages, grammars and machines. The set of all computable languages is described by ‘phrase structure’ grammars which are equivalent to Turing Machines.

Children learn their native language by hearing grammatical sentences from their parents or other members of their speech community. From this ‘environmental input’, children construct an internal representation of the underlying grammar. Children are not told the

grammatical rules. Neither children nor adults are ever aware of the grammatical rules that specify their own language.

Chomsky and others pointed out that the environmental input available to the child does not uniquely specify the grammatical rules. This phenomenon is known as ‘poverty of stimulus’. ‘The paradox of language acquisition’ is that children of the same speech community reliably grow up to speak the same language. The proposed solution of the paradox is that children learn the correct grammar by choosing from a restricted set of candidate grammars. The ‘theory’ of this restricted set is ‘universal grammar’ (UG).

The concept of an innate, genetically determined UG was controversial when introduced some 40 years ago and has remained so. The mathematical approach of learning theory, however, can explain in what sense universal grammar is a logical necessity.

Learning a language is inductive inference. The learner is presented with data and has to infer the rules that generate these data. The difference between ‘learning’ and ‘memorization’ is the ability to generalize beyond one's own experience to novel circumstances. In the context of language, the child learner will generalize to novel sentences never heard before. Any child can produce and understand sentences that are not part of his previous linguistic experience.

Learning theory describes the mathematics of learning with the aim to outline conditions for successful generalization. It can be shown that successful generalization requires an a-priori limited search space.

There are various mathematical approaches toward unifying formal language theory, learning theory and evolutionary game dynamics. In all these approaches, it has been assumed that the individuals cooperate with each other. Here we do not wish to take cooperation for given. Instead we would like to study the coevolution of communication and cooperation. Cooperation is required for the evolution of communication, and vice-versa communication can promote the evolution of cooperation. For indirect reciprocity, language seems to be essential. Moreover, we would like to investigate the concept of a universal grammar for learning cooperative behavior and for evolving cooperative behavior. The former is a feature of individuals. The latter a feature of nature.

Attempting to formalize some aspects of ethics in a basic model

We will formalize basic questions of ethics and morality by considering the following model (which is similar to a half round of an alternating PD). There are two players, you and somebody else. You can choose between action A and action B. The other person is the passive recipient of whatever you do. The other person has no choice and does not act (in this game). The payoff matrix is the following

	you get	the other person gets
you do A	0	0
you do B	1	-1

Thus you can choose between action A, which yields zero payoff to both players or action B which provides a reward for you at a cost for the other person. Without any additional assumptions, game theory says that your 'rational' choice is action B. Evolutionary game theory says that natural selection favors those individuals who choose action B in your role. But we can ask which *theory of ethics* is compatible with choosing action B, and so greatly complexify the discussion. Some additional considerations can immediately be added. For example, (i) the other is a declared enemy who would certainly choose action B in your position; your friends expect you to chose B; not choosing B would lower your reputation with your friends (ii) the other person is a stranger; you have never met him, you will never meet him again; nobody will witness your choice; the other person does not know you and will not recognize you again (iii) the other person is a friend; (iv) action B is unlawful or despised in your society; (v) action B is a completely accepted norm.

We can also consider the following payoff matrix in the same five scenarios:

	you get	the other person gets
you do A	0	0
you do B	10	-1

There is a much greater reward for you than there is a cost for the other person. How does this affect your choice? Notice that you can argue if two games were played in reversed roles than by choosing B, both players end up with a total payoff of +9 instead of 0. Hence, it is 'fair' to choose B in your position. (This option may remind us of Rawls's original position [Rawls 1972] , in which one purportedly chooses a moral or political system before knowing whether one will be powerful or powerless in a society.)

We will develop a quantitative evaluation of various ethical theories and study the role of evolution in forming ethical norms. We will study conditions that lead to the evolution of players who choose action A. At the same time we shall consider the possibility that ethics itself is part of *cultural* evolution, arguably obeying different laws than biological evolution: meme theory is clearly relevant, and needs our critical

attention. One needs to consider whether meme theory could be quantified in terms of expected benefits and rewards, or of the likely development of an ethical meme over time.

Certainly, real-world situations are usually much more complicated than what is initially described here. This is not necessarily a problem precluding our formal analysis; but we shall take into account the critiques that object to formal analysis precisely because they abstract from real-world complexities and non-rational considerations.

It has often been argued that evolutionary explanations of altruism are equivalent to 'taking the altruism out of altruism'. Hence, an apparently altruistic behavior is eventually of benefit to the individual (because others will reciprocate, for example). Any evolutionary explanation requires a cost to benefit analysis and the final sum must be positive. We will revisit this issue, because it is not completely correct. We will show that a biological trait (a gene, a behavior) can spread in a population with a certain probability although it *reduces* the biological fitness of the carrier.

Wisdom, trust, forgiveness and 'unlimited love'

In our game theoretic approach to cooperation we have encountered concepts like trust and forgiveness. Trust is usually associated with cooperating in the first round of a repeated Prisoner's Dilemma in the context of direct reciprocity. You have no data yet on the other person, but you trust that your cooperative opening will be rewarded. Another possible way to think about trust is in the context of a continuous PD. In each move, you have the option to cooperate on a scale anywhere between 0 and 100%. High levels of cooperation make you most vulnerable to exploitation and require trust that your partner will cooperate.

Forgiveness emerges in generous tit-for-tat (GTFT). This strategy always cooperates when the opponent has cooperated and sometimes cooperates when the opponent has defected. The probability to cooperate after an opponent's defection defines the level of forgiveness. A high level of forgiveness leads to a high payoff for GTFT interacting with other GTFT strategists. But there is a maximum level of forgiveness that is still compatible with resisting invasion of defectors.

We propose to explore how additional concepts of theology might arise in the game theoretic approach. These concepts include love, wisdom, hope, dignity and sanctity. We cannot assume that cooperation, as defined above, will necessarily involve these other virtues; cooperation may indeed initially be entirely selfishly motivated, even given its necessary 'costliness'.

Loving somebody, however, seemingly implies that in addition to cooperation with that person, you also trust the other person, and that you can forgive him/her. Love means that the gain of the other person is also your gain. This is an interesting point. In the kin selection of selfish genes, the actors are bound together by genetic relatedness, whereas in love they are bound together by empathy.

Wisdom could entail the interesting property of TFT-like strategies never to try 'winning' against any given opponent. In the game with any other player, a TFT player

can never obtain a higher payoff than the other person. TFT never tries to exploit. The success of TFT is to establish cooperation with many other strategies and thereby receive a higher cumulative payoff than others. For example, TFT might cooperate with strategies X and Y, whereas the game between X and Y may not be fully cooperative. This is an aspect of wisdom.

We point out that evolutionary or game theoretic explanations of certain behaviors do not take away dignity or sanctity. They just show that the highest ethical imperative, perhaps even when considered as a divine command, is compatible with logical mathematical rules that are woven into the world. Showing that some theological concepts might emerge from considerations of evolution and mathematics is a powerful and convincing argument that underlines the elegance and beauty of the world and its omnipotent Creator.

In Christian theology, there is a quest for unconditional cooperation and unlimited love. We are not satisfied with strategies that lead to cooperation only because this action increases one's own payoff. The Good Samaritan does not help in order to receive help. If you give to the poor then, according to Jesus's teaching, your left hand should not know what your right hand is doing. We should help out of love and not in order to accumulate treasures in this life. We would like to study the question whether evolutionary biology can explain such 'unlimited love' (in contradistinction from more limited cooperation).

It is inconceivable that Jesus could ever 'defect' (in the exploitative sense defined earlier); but he did punish the merchants in the Temple, and did drive them away. Similarly, a loving parent will limit the freedom of a child at times. Unconditional cooperation is not necessarily called for in any human interaction. We propose to study this question by extending existing work on altruistic punishment. Thus, a strategy could be to cooperate all the time, but punish a defector when necessary. Cooperation and defection, on the one hand, and punishment, on the other, are two different channels of interaction. We can also complexify the approach by allowing more than two channels of interaction.

Experimental theology

We would like to begin to explore collaborations with experimental economists and psychologists to study altruistic behavior of humans in conjunction with religious beliefs and theological motivations. Are believers and practitioners of a religion more altruistic than others? If not, why not? Can we distinguish between those who merely assent to certain religious beliefs and those who lead a saintly life? This enterprise could be the very beginning of what we might call an 'experimental theology', in which evolutionary biology and the social sciences play equally important roles in establishing the conditions of varying degrees and sorts of cooperation.

Is love in the fabric of the universe?

Moving from these initial starting points and items for discussion, we shall go on to study which fundamental principles of evolutionary systems can support the emergence of true unselfish love as promoted by Christianity and other religions (cf. Post et al 2002, Post 2003). Whilst our research proposal is obviously attuned primarily to the Christian tradition, we expect and hope to bring into the discussion experts in, and representatives of, at least two non-Christian traditions to broaden the comparative perspective.

Meta-Ethical and Theological Considerations

1. Our proposal so far deliberately does not make a choice *re* what ethical or meta-ethical theory we propose to favor. This choice is crucial, however, for what sort of *theism* might be implied in any particular application of game theory (and, in principle, as we point out, the 'game' can be set up any way we wish). So ethical theory, meta-ethical theory, and potential choices about different views of 'God' go together. We shall need to extend, and complexify, the range of systematic options that we consider in these three areas as we proceed in our interdisiplinary working-group, and it is a conscious feature of this proposal that we do not wish to dictate in advance where the ethical/theological die will finally be cast in our joint conclusions. However, we may consider at the outset at least the following three (crudely expressed) 'types' of option which present themselves – and are arguably latent in what has already been discussed, above:

i. Virtue ethics (founded in a 'natural law' perspective) – Thomas Aquinas, *par excellence*. God, on this view, as creator and sustainer of all that is, is indirectly traceable in the workings of nature itself, and hence in the evolutionary system; but it is in revelation – supremely in Jesus Christ – that we find the 'perfection' of what we can read off nature by reason, and which provides us with the key to *virtue*; virtue consists in conforming oneself to the standards of Christ's life and death over and above what can be known from natural law, although natural law acts as a rational basis. Note that human freedom on this view involves cooperating with God's grace, *not* exercising purely human 'autonomy'.

ii. Enlightenment *a priori* ethics of moral autonomy and universal goods (sometimes called an 'ethics of duty') – Immanuel Kant, *par excellence*. God, on this view, is what must be presumed (by humans already engaged in autonomous acts of moral choice, guided by reason and conscience) as a human 'moral postulate' to guarantee the *summum bonum*. Freedom involves a timeless act of supervenience over conditioning or context.

iii. Utilitarianism ('the end justifies the means') – Jeremy Bentham, or John Stuart Mill, *par excellence*. The 'ethical' act is simply what allows one to arrive at the desired moral *goal*; how to balance this goal with the demands of *justice* (for all) remains somewhat problematic on this view. It is also less than obvious what 'God' looks like on a utilitarian view – perhaps something like the overseer of the *panoptikon* that Bentham proposed for his newly-conceived prison-system?

Now: it would seem that, *much* of the time, game theorists presume utilitarianism; and certainly the Darwinian interest in 'higher reproductive success' seems to be interested in the end rather than the means. One of the first tasks of our research group, then, will be to probe critically what ethical (and meta-ethical) presumptions are being made, more or less unconsciously, by current game theorists, and to examine the implications. If at times such theorists assume some form of utilitarianism, at other times a simpler form of 'ethical egoism' may be what is taken for granted – again without specific argumentation. But when we get to more complex strategies delineated above that involve deliberate acts of self-denial, one of the other two meta-ethical theories seems implicitly to come into view. We could debate that. (Whoever wrote the introductory material for the Templeton meeting in October 2003, for instance, definitely assumed the Kantian view – that 'God' appears as a necessary 'construct' at some point in human ethical development; and our earlier talk of the 'emergence of cooperation', etc. (above) *might* seem to suggest that line: out of that emergence would eventually come the need for a 'God' concept.) However, when we then consider the possibility of 'unselfish behaviour' being 'grounded in evolutionary biology', a hypothesis we have already hinted at, there is then the potential for a tilt to a Thomist perspective. Indeed, the great value of the Thomist view is that it sees a continuum, not a disjunction, between 'natural' life and human morality – both are 'grounded' in God. Some ethicists would argue, in fact (see Jackson 2003, e.g.), that we do not need to make a disjunctive choice between the three basic types of meta-ethical position outlined, since they describe different *perspectives* on the human and the moral.

It follows, of course, that the quest for the 'highest level of morality' (see above) will have to declare what meta-ethical theory it favors before delivering a judgment. This will form an important part of our opening deliberations. However, further reflection suggests that our initial three-part typology of meta-ethical theories may not be subtle enough:

2. For here comes another complexification: if we are also interested in New Testament ethics, it is not *clear* which of the meta-ethical theories so far mentioned would come out top. Some of the explicit ethical claims of Jesus could be compatible with more than one theory (e.g., the 'golden rule': 'Do unto others as you would have them do to you'). But the radical saying of Jesus found in the source *Q (Quelle)* that Matthew and Luke have in common demands a love of *enemies*, and in other strands of the gospels there is also a *radical* demand to forgive and to go on forgiving. Even John and Paul draw back from these demands at points; and we could debate what 'game' could possibly be 'won' by these strategies, and whether *they* are the 'highest morality' of pure unselfishness. At this point in our discussion we would have to bring into consideration further meta-ethical theories which do not merely bring in divine revelation to 'complete' reason, but in some instances apparently to confound it: 'divine command theory', and the 'ethics of love' are the obvious cases, and ostensibly they may seem to present greater difficulties for potential compatibility with evolutionary theory than the first three meta-ethical theories we considered (though see the work of Post, 2002 and 2004). We shall assess their possibilities, using the specific insights of game theory (though not without questioning the possible blind-spots or denials of game theory itself). We shall be asking: could a love ethic in any way be written into the stuff of evolution

itself, or must it be seen as the culmination of a conscious ethical development by *homo sapiens* in response to divine command and *over against nature*? If I believe that there is a God who rewards/punishes even what I do 'in secret', or that acting unlovingly actually causes to cut 'against the grain of the universe', how would I differently calculate or strategize the kinds of +1/-1 game described above?

3. Another vital point of clarification thus immediately arises here: at what point in the evolutionary process does it make sense to talk of *volitional* 'strategy'? A lot seems to hang on this. When we say that 'we will develop a mathematical framework for classifying "morality" in situations of direct reciprocity', we need to clarify whether we are operating at the level of physicalism and *natural* evolutionary 'bias', or at the level of *homo sapiens* and volitional, ethical procedure in the strict sense. 'Good reputation', e.g., in 'indirect reciprocity', only makes sense on the latter presumption. Game theory as applied to evolution seems to 'anthropomorphize' the 'strategies' of non-conscious beings by a sort of sleight of hand. *Biologically*, motivations are only relevant through their behavioral results; but this is a presumption that will need critical discussion. Similarly, have talked above – daringly – about 'light and darkness' (=? cooperation and defection) when we are still ostensibly at the physicalist evolutionary level (see above). These points need ethical clarification.

4. If, however, we do want to talk about 'light' *in advance* of human volition, whose light is it? God's? Are we then suggesting obliquely – what some claim – that 'science' can only ever be agnostic about – *viz.*, that there is some *cosmic* tendency towards the 'light'? This would bring 'God' in earlier in the discussion than as a merely Kantian guarantor of the *summum bonum*, or as a Benthamite surveyer of human goal-directed strategies. Perhaps this clarifies why the Thomist option seemingly looks so interesting from the perspective of evolutionary theory: it is in 'natural law' that 'games' begin, but in human volition (guided by divine revelation and desire for 'virtue') that 'games' reach a higher level of purposive structure. However, as stressed above, our research group will not *a priori* tilt towards one meta-ethical theory over another. We shall deliberately include in our discussions noted exponents of different kinds of meta-ethical theory, different kinds of theism (a-temporal of emergentist), and indeed different kinds of physicalism. In short, we shall be fundamentally concerned as a research group with *what* metaphysical hypotheses best explain evolutionary phenomena in the light of game theory.

5. The above considerations in turn raise the interesting question of 'delayed gratification' (or *eschatology* in Christian theological terms). The more complex games we have described above require more foresight, wisdom, and *ascetical* virtue. Or so it would seem – and this is worthy of comment and discussion. Is that why, interestingly, many rational behaviors – predictable on game theory – are *not* the ones that 'fallen' humans actually choose? Questions of 'sin' and 'fall' interestingly insert themselves here. And so, likewise, does the ever-contentious problem of 'other minds', since only an animal cognitively conscious-enough to recognize an 'other mind' is going to make certain choices over others. This is the point at which game-theoretical models seem to intersect, fascinatingly, with issues of specific neurological capacity.

6. Also, surely, in the more complex games we cannot *simply* equate defection with 'darkness/despair' and cooperation with 'light/hope' (see again, above). The

example of Bonhoeffer's willing death (often discussed by MN and SC) is of interest here. An initial 'defection' of this sort can constitute a self-destructive act now, maybe seemingly worthless, but ultimately it may 'win' morally.

7. Sacrifice: Does 'true' unselfish love necessitate *conscious* sacrifice? This would not (obviously) be validated by a utilitarian approach, but – for a Christian at least - more likely by a 'Christic' perception of being. This would be worth discussing in the light of some of the more complex and ramified game-theory 'strategies' already outlined in the paper. It could then open up to a comparative ethics discussion, bringing in parallel – or different – ethical theories from non-Christian religious traditions.

8. Finally, a meta-meta-question (otherwise known as a question from the sociology of knowledge): Whose idea is game theory, anyway? Why does game theory seem to have so much explanatory power to *this* generation of post-modern Western scholars? (On this, see Mary Douglas's cultural theory: is a particular sort of game theory happily compatible with Thatcherite/Bushite competitive individualism and global capitalism? Does it insidiously collude in 'baptizing' these facets of contemporary life with the force of evolutionary 'necessity'? (see Douglas 1992, 1996) Is game theory intrinsically patriarachal or sexist? Is its current fascination for economists, biologists and mathematicians an unthinking collusion in a notable set of ethical mistakes? SC persistently rasies these questions; but MN would reply, *no*: i. because in principle we can set up the game in many different ways, and the atavistic will to power doesn't have to be valorized; and ii. because it is striking in any case (so say the economists using game theory) that folk *don't* 'play' rationally (i.e., to their advantage) much of the time. Volitional and affective dimensions often lead to ostensibly non-rational choices in our 'games'. Nonetheless, this fundamental criticism should be aired and assessed.

Evolution and Theology in General

To sum up the ambitious goals of this research collaboration: We will take the opportunity in this project to create a new discussion of the general implications of Darwinian evolution for theology and ethics, and vice versa. When Darwin proposed his ideas, it seemed to him and many of his contemporaries that evolution was at variance with the Christian and Jewish narrative of the creation of the world and humanity's place in this creation. One could either be faithful to Christianity or accept the new teachings of evolution. Now, a century and a half later, it is obvious to many people that evolution and religion are not, as such, intrinsically incompatible. However, the *precise* mode of their relation is unclear and continues to be debated. We want to establish a proper form of communication between evolution and religion, and to do so in the most sophisticated terms available to us.

First, it has to be understood that evolution as such is no 'religion'. Many questions that are important to humans are not addressed by evolution (or science in general), whereas religion does address such questions. Evolution (or science) seeks structural and functional explanations for empirical questions, but it would seem that this is not the focus of 'religion'. Thus, it would be incorrect to interpret Genesis as a 'scientific' theory for the origin of the universe and life within. Conversely it is incorrect to interpret Darwinism as a 'religion' contradicting Genesis. Nonetheless, we cannot

separate science and religion, and thus the altruism debate (as outlined above) certainly presents the *possibility* of a genuine conflict between scientific and religious approaches, as Darwin himself recognized.

In the Jewish, Christian and Muslim traditions, there is confidence that Genesis contains the narrative that God wants us to have about how He conceived of the world. For the religious 'person of the book', Genesis represents the appropriate description of how the Supreme Being initiates creation and the flow of time, and the Bible in general describes God's omnipresence and continuing interaction with His creation. For the scientist who is also religious, one could say that it has pleased God to reveal natural laws that guide this creation and its dynamical processes. However, in the past, different questions were asked by the theologian and the scientist, and correspondingly different answers were received. Our attempt is to bridge this divide.

From the perspective of science, we want to study to what extent evolution shows a gradient of 'progress' or a tendency toward increasing complexity. Yet evolution, as we now understand it, is blind with respect to a long-term perspective. Whatever survives, we might say, 'hangs around'; whatever cannot survive, disappears. Evolution *as such* does not optimize fitness. Mutations can become fixed in a population, although they reduce the average fitness of individuals in that population. Yet there seems to be an (unexplained!) gradient toward increasing complexity. Evolution on earth has progressed from simple cells (prokaryotes), to more complex cells (eukaryotes), to multi-cellular organisms, to a species (humans) which invented a new mode of evolution. This new mode is cultural evolution. Some higher animals have a rudimentary form of cultural evolution, but only human language allows (comparatively) unlimited cultural evolution.

There is no mechanistic theory yet that explains why genetic evolution *did* progress toward more complexity (on this scale) and toward cultural evolution in a being that is conscious about evolution and wonders about its role in the world. To contemporaries of Darwin it seemed that his theory degraded the human person from a central position in the biota to one of many species an offshoot of primates. Reflecting on human invention of cultural evolution, however, the human person is again unique among all animals.

Christianity at times asks the question whether the creation is progressing in a certain direction - and specifically toward God - a matter on which a variety of theological opinions have been classically voiced. But science *per se* does not obviously either seek or find teleology (or if it does, we might say that it is already tending towards a covert theology). For many scientists, the world falls from nothing into nothing, adopting random, transient forms in the meanwhile. It is seemingly religion, then, that provides teleology (or 'eschatology'). Here the world has a meaning. Each life has a purpose. In the beginning we were with God. God is the Alpha and the Omega, the first and the last. In the end, we will fall into the hands of our maker.

Our research program, then, will be concerned at its widest with the question of whether there is a larger, teleological purpose to evolution which can only be speculated about, and thematized, in religious/theological terms. We shall be concerned, as we have indicated, to generate a range of possible meta-scientific teleologies for consideration; but we shall also keep under review the contested area of 'immanent teleology', asking

whether our game-theoretical approach already nudges us towards questions of evolutionary 'purpose'. This matter remains an open question in the current conversations between MN and SC.

We will compare in some detail, then, the light that science and religion can shed on the world of evolution. The communication between these two approaches has potentially much to offer. Keeping in mind, however, that God's ways are mysterious and that His thoughts are above our thoughts, it has to be determined to what *extent* science can represent an approach toward understanding God. This matter must be discussed with careful nuance, and with greater reflection on the apophatic dimensions of God-talk than is characteristic of some work in the science and religion debate. If God reveals himself in whatever way he chooses, attempting to out-calculate God is like building a tower of Babel that tries to reach the heavens. On the other hand, if God's intrinsic character (as love), as well as God's initial authorship, are written into the nature of the universe from the very start, we should not be surprised to find some resistant features in 'evolution' that could at least *beg* teleologicial/theological questions.

We would like to emphasize that religion and evolution are not at variance with each other. We remember that Darwin was theologically motivated. In the 'Origin of Species' he contrasts two different possibilities: (i) God created the world and embedded it with elegant and precise laws that unfold His Creation and (ii) God has to perform a little miracle whenever He wants to create a new species. Darwin collected evidence for (i) and against (ii). It is also our intention to study those natural and divine laws of evolution.

Summary: Research Questions and Hypotheses

Our central research questions and hypotheses may thus be summed up as follows:

1. We shall return to the unanswered questions about cooperation already identified by Darwin, and ask if game theory can provide an adequate solution to them without also (re)introducing the question of God.

2. We shall examine the current state of game theory as applied to the question of cooperation, and critically probe the ethical *presumptions* made by such game theory.

3. We shall delineate different 'meta-ethical' theories that may underlie game theoretical reflection, and show their connections to different possible perceptions of the relation of God to the world/nature. (Here we shall compare natural law, divine command, and utilitarian theories, with variations, and ask if we need to make disjunctive choices between them.)

4. We shall explore different possible meanings of 'cooperation', and carefully distinguish self-interested cooperation, altruism, unselfishness, (unlimited) love, wisdom, hope and forgiveness. We shall ask which of these can be seen as products of evolution on a game theoretic approach, and delineate the precise connection implied to divine creativity.

5. We shall look at aspects of the Golden Rule which have emerged in game theoretic models of fairness (in the context of the Ultimatum game), and consider whether a Golden Rule ethic may be seen as in accordance with nature and/or God's will. Can a purely economic rationality do the work of the 'practical reason' of ethics, or is there a covert reduction here?

6. We will extend our study of direct reciprocity and explore the ethical and theological implications of strategies like tit-for-tat, generous tit-for-tat and win-stay-lose-shift.

7. We will study indirect reciprocity and ask if evolution can explain love of strangers.

8. Jesus asks us to love our enemies. We shall enquire which evolutionary game (if any) can possibly be won by such a strategy.

9. We will aim to establish an experimental game that can identify cooperative or altruistic behavior in people with different cultures and religions. Can 'saintly' behavior be measured cross-culturally?

10. We shall look at how cooperation might lead to an additional reward in the form of inner happiness and peace. We shall explore how this possibility might affect games of cooperation or defection.

11. We will study the impact of group selection on the evolution of cooperation.

12. We will study the mathematics of evolution in finite populations.

13. We will formulate the first model for co-evolution of cooperation and communication (human language).

14. We will aim to combine learning theory with evolutionary game theory.

15. We will attend to the possible gender associations of various forms of cooperation (intra-culturally and cross-culturally), and examine the relevance of feminist ethics to the discourse of evolutionary biology. We shall also ask if game theory itself is 'masculinist' in its presumptions, and if so in what sense.

16. In sum, we will work on the general implications of evolutionary biology for theology and vice versa, in order to initiate a new and sophisticated 'neo-Darwinian' discussion in the university.

Strategic Dissemination

Our work will be disseminated in a number of ways. First, we plan at least one co-authored book which will make available to an interested lay readership the novelty of what we have to say about *Games and God.* We might expect such a book to attract some attention in the current intellectual and cultural climate, especially if carefully 'pitched' and marketed. In addition, we shall publish related, more technical, articles in peer-reviewed journals, and edit the proceedings from our various conferences. Our aim

will be to organize presentations which will disseminate our findings not only at Harvard, but in our relevant learned associations and conferences (e.g., the *American Academy of Religion,* where SC is involved in both the Systematic Theology and the Philosophy of Religion group). We shall investigate the possibility of assembling readings for a textbook on game theory and theology from our weekly seminars; we would envisage asking a couple of our post-doctoral assistants to work on the editing of such a textbook under our supervision. We expect to make much of the material from our conferences available on the web, as discussed elsewhere in this proposal. We also plan a number of articles in more popular – but serious – formats such as *The New York Times*, *The Boston Globe*, *Time*, *The Tablet*, *The Christian Century*, *Commonweal* and *Harvard Magazine.* We would hope also to make some radio appearances to launch our book together. Finally our work together will certainly inform our own teaching, and our Visiting Scholars and Post-docs who teach while they are at Harvard will also be involved in purveying our findings to Harvard students.

References and Additional Bibliography

Alexander, Richard D. ***The Biology of Moral Systems*** Aldine de Gruyter (Berlin, Germany) 1987.

Arnhart, Larry **Darwinian Natural Right**: **The Biological Ethics of Human Nature** SUNY 1998.

Axelrod, Robert M. ***The Evolution of Cooperation*** Basic Books (New York, NY) 1984.

Axelrod, Robert & Hamilton, William D. (1981) The evolution of cooperation. ***Science*** 211: 1390-1396.

Bickerton, Derek ***Language & Species*** The University of Chicago Press 1990.

Binmore, K. ***Game Theory and the Social Contract*** MIT Press 1994.

Brandt, H. & Hauert, Christoph & Sigmund, Karl (2003) Cooperation, punishment and reputation in spatial public goods games. ***P Roy Soc Lond B Bio*** 270: 1099-1104.

Brooke, John Hedley ***Science and Religion: Some Historical Perspectives*** Cambridge University Press 1991.

Clayton, Philip ***God and Contemporary Science*** William B. Eerdmans Publishing Company (Grand Rapids, MI) 1997.

Clayton, Philip & Peacocke, Arthur eds. **In Whom We Move and Have Our Being: Panentheistic Reflections on God's Presence in a Scientific World** Eerdmans 2004.

Clatyon, Philip & Schloss, Jeffrey P. eds. **Evolution and Ethics: Human Morality in Biological and Religious Perspective** Eerdmans 2004.

Conway Morris, Simon ***The Crucible of Creation: The Burgess Shale and the Rise of Animals*** Oxford University Press 1998.

Conway Morris, Simon ***Life's Solution: Inevitable Humans in a Lonely Universe*** Cambridge University Press 2003.

Davies, Paul C.W. ***The Mind of God: The Scientific Basis for a Rational World*** Simon & Schuster (New York, NY) 1992.

Douglas, Mary **Risk and Blame: Essays in Cultural Theory** Routledge 1992.

_____, **Thought Styles: Critical Essays on Good Taste** Sage Publications 1996.

Farber, Paul Lawrence **The Temptations of Evolutionary Ethics** University of California Press 1994.

Fehr, Ernst & Fischbacher, Urs (23 OCT 2003) The nature of human altruism. ***Nature*** 425(6960): 785-791.

Frank, Steven A. ***Foundations of Social Evolution*** Princeton University Press 1998.

Frankena, William K. **Ethics** Prentice-Hall 1973.

Fudenberg, Drew & Levine, David K. ***The Theory of Learning in Games*** MIT Press 1998.

Fudenberg, Drew & Tirole, Jean ***Game Theory*** MIT Press 1991.

Gold, E.M. (1967) Language identification in the limit. ***Inform Control*** 10: 447–474.

Grant, Colin **Altruiusm and Christian Ethics** Cambridge University Press 2001.

Hamilton, W.D. ***Narrow Roads of Gene Land, Volume 1: Evolution of Social Behaviour*** Oxford University Press 1998.

Hamilton, W.D. ***Narrow Roads of Gene Land, Volume 2: Evolution of Sex*** Oxford University Press 2002.

Hauert, Christoph & De Monte, Silvia & Hofbauer, Josef & Sigmund, Karl (10 MAY 2002) Volunteering as red queen mechanism for cooperation in public goods games. ***Science*** 296(5570): 1129-1132.

Hauert, Christoph & Doebeli, Michael (8 APR 2004) Spatial structure often inhibits the
evolution of cooperation in the snowdrift game. ***Nature*** 428: 643-646.

Haught, John F. ***Science and Religion: From Conflict to Conversation*** Paulist Press (Mahwah, NJ) 1995.

Haught, John **God After Darwin: A Theology of Evolution** Westview 2000.

Hefner, Philip ***The Human Factor: Evolution, Culture and Religion*** Fortress Press (Minneapolis, MN) 1993.

Henrich, J. & Boyd R. & Bowles S. & Camerer C. & Fehr E. & Gintis H. & McElreath R. (MAY 2001) In search of homo economicus: behavioral experiments in 15 small-scale societies. ***Am Econ Rev*** 91(2): 73-78.

Hofbauer, Josef & Sigmund, Karl ***Evolutionary Games and Population Dynamics*** Cambridge University Press 1998.

Hofbauer, Josef & Sigmund, Karl (2003) Evolutionary game dynamics. ***B Am Math Soc*** 40(4): 479-519.

Jackendoff, Ray ***Foundations of Language: Brain, Meaning, Grammar, Evolution*** Oxford University Press 2002.

Jackson, Timothy P ***The Priority of Love: Christian Charity and Social Justice*** Princeton University Press 2003.

Keller, Laurent ed. ***Levels of Selection in Evolution*** Princeton University Press 1999.

Killingback, Timothy & Doebeli, Michael (7 APR 1998) Self-organized criticality in spatial evolutionary game theory. ***J theor Biol*** 191(3): 335-340.

Lewin, Roger & Foley, Robert A. ***Principles of Human Evolution*** Blackwell Publishing (Oxford, UK) 2004.

Maynard Smith, John ***Evolution and the Theory of Games*** Cambridge University Press 1982.

Midgley, Mary **The Ethical Primate: Humans, Freedom and Morality** Routledge 1994.

Moran, P.A.P. ***The Statistical Processes of Evolutionary Theory*** Clarendon Press (Oxford, UK) 1962.

Murphy, Nancey **Anglo-American Postmodernity: Philosophical Perspectives on Science, Religion, and Ethics** Westview Press 1997.

Nowak, Martin A. & Komarova, Natalia L. & Niyogi, Partha (5 JAN 2001) Evolution of universal grammar. ***Science*** 291(5501): 114-118.

Nowak, M.A. & Komarova, Natalia L. & Niyogi, Partha (2002) Computational and evolutionary aspects of language. ***Nature*** 417: 611-617.

Nowak, Martin A. & Krakauer, David C. (JUL 1999) The evolution of language. ***P Natl Acad Sci USA*** 96: 8028-8033.

Nowak, M.A. & May, Robert M. (1992) Evolutionary games and spatial chaos. ***Nature*** 359: 826-829.

Nowak, Martin A. & Plotkin, Joshua B. & Jansen, Vincent A.A. (30 MAR 2000) The evolution of syntactic communication. ***Nature*** 404: 495-498.

Nowak, M.A. & Sasaki, Akira & Taylor, Christine & Fudenberg, Drew (2004) Emergence of cooperation and evolutionary stability in finite populations. ***Nature*** 428: 646-650.

Nowak, M.A. & Sigmund, Karl (1992) Tit for tat in heterogeneous populations. ***Nature*** 355: 250-253.

Nowak, M.A. & Sigmund, Karl (1993) A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. ***Nature*** 364: 56-58.

Nowak, M.A. & Sigmund, Karl (1998) Evolution of indirect reciprocity by image scoring. ***Nature*** 393: 573-577.

Numbers, Ronald ***Darwinism Comes to America*** Harvard University Press 1998.

O'Hear, Anthony **Beyond Evolution: Human Nature and the Limits of Evolutionary Explanation** Clarendon Press 1997.

Peters, Ted ***Playing God? Genetic Determinism and Human Freedom*** Routledge (London, UK) 1997.

Pfeiffer, Thomas & Schuster, Stefan & Bonhoeffer, Sebastian (20 APR 2001) Cooperation and competition in the evolution of ATP-producing pathways. ***Science*** 292(5516): 504-507.

Pinker, Steven ***The Language Instinct: How the Mind Creates Language*** William Morrow and Company (New York, NY) 1994.

Pinker, Steven & Bloom, Paul (DEC 1990) Natural language and natural selection. ***Behav Brain Sci*** 13(4): 707-784.

Plotkin, Henry **Evolution in Mind: An Introduction to Evolutionary Psychology** Harvard University Press 1997.

Polkinghorne, John ***The Faith of a Physicist: Reflections of a Bottom-Up Thinker*** Princeton University Press 1993.

Polkinghorne, John C. ***Belief in God in an Age of Science*** Yale University Press 1998.

Post, Stephen G. **Unlimited Love: Altruism, Compassion and Science** Templeton Foundation Press 2003.

Post, Stephen G. & Underwood, Lynn G. & Schloss, Jeffrey P. & Hurlbut, William B. eds. ***Altruism and Altruistic Love: Science, Philosophy, and Religion in Dialogue*** Oxford University Press 2002.

Rawls, John **A Theory of Justice** Clarendon Press 1972.

Ruse, Michael ***Darwin and Design: Does Evolution Have a Purpose?*** Harvard University Press 2003.

Russell, Robert John & Clayton, Philip & Wegter-McNelly, Kirk & Polkinghorne, John eds. ***Quantum Mechanics: Scientific Perspectives on Divine Action - Volume 5*** Vatican Observatory Publications (Vatican City State) 2001.

Schloss, J.P. "Sociobiological Explanations of Altruistic Ethics: Necessary, Sufficient, or Irrelevent Perspective on the Human Moral Quest" In **Investigating the Biological Foundations of Human Morality** James Hurd ed. The Edwin Mellen Press (New York) 1996.

Schloss, J.P. "Of Natural Theology and Natural Theology: Evolutionary Accounts of Altuistic Morality and the Quandary of Goodness by Design" In **Mere Creation** William Dembski ed. Intervarsity Press 1998.

Schloss, J.P. "Wisdom Traditions as Mechanisms of Homeostatic Integration: Evolutionary Perspectives on Organismal 'Laws of Life'" In **The Science of Wisdom and the Laws of Life** Warren Brown ed. Templeton Foundation Press 2000.

Smith, Christian ***Moral, Believing Animals: Human Personhood and Culture*** Oxford University Press 2003.

Sober, Elliott & Wilson, David Sloan ***Unto Others: The Evolution and Psychology of Unselfish Behavior*** Harvard University Press 1998.

Stark, Rodney ***One True God: Historical Consequences of Monotheism*** Princeton University Press 2001.

Stark, Rodney ***For the Glory of God: How Monotheism Led to Reformations, Science, Witch-Hunts, and the End of Slavery*** Princeton University Press 2003.

Sudgen, Robert ***The Economics of Rights, Cooperation, and Welfare*** Blackwell Publishing (Oxford, UK) 1986.

Taylor, P.D. & Jonker, L.B. (1978) Evolutionary stable strategies and game dynamics. ***Math Biosci*** 40(1-2): 145-156.

Templeton, John Marks ed. ***How Large is God? The Voices of Scientists and Theologians*** Templeton Foundation Press (Radnor, PA) 1997.

Valiant, L.G. (1984) A theory of learnable. ***Commun ACM*** 27: 436–445.

Vapnik, V.N. & Chervonenkis, A.Y. (1971) On the uniform convergence of relative frequencies of events to their probabilities. ***Theor Probab Appl*** 17: 264–280.

Weibull, Jörgen W. ***Evolutionary Game Theory*** MIT Press 1995.

Williams, George Christopher ***Natural Selection: Domains, Levels, and Challenges*** Oxford University Press 1992.

Wilson, D.S. (1975) Theory of group selection. ***P Natl Acad Sci USA*** 72(1): 143-146.

Wilson, D.S. (1983) The group selection controversy – history and current status. ***Annu Rev Ecol Syst*** 14: 159-187.

Wilson, D.S. (SEP 1990) Weak altruism, strong group selection. ***Oikos*** 59(1): 135-140.

Wilson, D.S. (JUL 1999) A critique of R.D. Alexander's views on group selection. ***Biol Philos*** 14(3): 431-449.

Wilson, David Sloan ***Darwin's Cathedral: Evolution, Religion, and the Nature of Society*** University of Chicago Press 2002.

Wilson, David Sloan & Dugatkin, Lee A. (FEB 1997) Group selection and assortative interactions. ***Am Nat*** 149(2): 336-351.

Wilson, D.S. & Sober, E. (DEC 1994a) Reintroducing group selection to the human behavioral-sciences. ***Behav Brain Sci*** 17(4): 585-608.

Wilson, D.S. & Sober, E. (DEC 1994b) Group selection – the theory replaces the bogeyman – authors response. ***Behav Brain Sci*** 17(4): 639-647.

Wilson, D.S. & Sober, E. (DEC 1996) More on group selection and human behavior. ***Behav Brain Sci*** 19(4): 782-787.