Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Theoretical Biology 224 (2003) 399–410

Journal of Theoretical Biology

www.elsevier.com/locate/jtbi

Evolution of cooperation: cooperation defeats defection in the cornfield model

Johan H. Koeslag*, Elmarie Terblanche

Department of Medical Physiology, University of Stellenbosch, P.O. Box 19063 Tygerberg 7505 South Africa

Received 15 January 2003; received in revised form 16 April 2003; accepted 1 May 2003

Abstract

"Cooperation" defines any behavior that enhances the fitness of a group (e.g. a community or species), but which, by its nature, can be exploited by selfish individuals, meaning, firstly, that selfish individuals derive an advantage from exploitation which is greater than the average advantage that accrues to unselfish individuals. Secondly, exploitation has no intrinsic fitness value except in the presence of the "cooperative behavior". The mathematics is described by the simple Prisoner's Dilemma Game (PDG). It has previously been shown that koinophilia (the avoidance of sexual mates displaying unusual or atypical phenotypic features, such as mutations) stabilizes any inherited strategy in the simple or iterated PDG, meaning that it cannot be displaced by *rare* forms of alternative behavior which arise through mutation or occasional migration. In the present model *equal* numbers of cooperators and defectors (in the *simple* PDG) were randomly spread in a two-dimensional "cornfield" with uniformly distributed resources. Every individual was koinophilic, and interacted (sexually and in the PDG tournaments) only with individuals from within its immediate neighborhood. This model therefore tested whether cooperation can outcompete defection or selfishness in a straight, initially equally matched, evolutionary battle. The results show that in the absence of koinophilia cooperation was rapidly driven to extinction. With koinophilia there was a very rapid loss of cooperators in the first few generations, but thereafter cooperation slowly spread, ultimately eliminating defection completely. This result was critically dependent on sampling effects of neighborhoods. Small samples (resulting from low population densities or small neighborhood sizes) increase the probability that a chance neighborhood comes to consist predominantly of cooperators. A sexual preference for the most common phenotype in the neighborhood then makes that phenotype more common still. Once this occurs cooperation's spread becomes almost inevitable.

© 2003 Elsevier Ltd. All rights reserved.

Keywords: Natural selection; Prisoner's dilemma game; Spatial PDG; Units of selection; Mate choice; Koinophilia

1. Introduction

A major unsolved biological problem is the evolution of cooperative behavior (Maynard Smith, 1964; Williams, 1966). Cooperation is defined here as any inherited behavior that enhances the fitness of a group (e.g. a community or species), but which, *by its nature*, can be exploited by selfish individuals, meaning, firstly, that selfish individuals derive an advantage from exploitation which is greater than the average advantage that accrues to unselfish individuals. Secondly, exploitation has no intrinsic fitness value except in the presence of the "cooperative behavior" practiced by the other members of the group.

It is easy to see that cooperation cannot therefore be evolutionarily stable. Since cooperation can, by definition, be exploited by selfish individuals, such selfishness will, if it is inherited, quickly invade and replace the group advantageous cooperation.

The mathematics is described by the simple Prisoner's Dilemma Game (PDG). The game is played by two contestants, who can either play ***C*** (cooperate) or ***D*** (defect). There are therefore four outcomes: ***C****C*, ***C****D*, ***D****C*, and ***D****D*, where the first letter (in bold) denotes the protagonist's move, and the second his opponent's. If the payoffs to the protagonist are such that ***D****C* > ***C****C* > ***D****D* > ***C****D* (Maynard Smith, 1989; Axelrod and Dion, 1988; Milinski, 1993; Sigmund, 1993; Nowak et al., 1994a, b), then cooperation is a group advantage, as

---

*Corresponding author. Tel.: +27-21-938-9390; fax: +27-21-938-9476.

E-mail address: jkoeslag@sun.ac.za (J.H. Koeslag).

0022-5193/03/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.

doi:10.1016/S0022-5193(03)00188-7

defined above. Typical fitness payoffs to the protagonist are, therefore, five points for ***DC***, three for ***CC***, one for ***DD***, and zero for ***CD***. (It is usual that the reward for ***CC*** is greater than [***DC*** + ***CD***]/2.) For the group (the pair of contestants) ***CC*** is clearly the best mutual strategy, earning them a total of six points. This situation can be exploited by playing ***D*** (only while the other continues to play *C*), which earns the protagonist five points (against three if he had cooperated). However, when both individuals play ***D*** then neither earns more than one point. ***D***, however, never loses.

It is now widely recognized that if the game is played repeatedly (the so-called iterated Prisoner's Dilemma Game, IPD), and players remember each other's last moves, then tit-for-tat (TFT) is a robust strategy which promotes reciprocal cooperation (Axelrod and Hamilton, 1981; May, 1981, 1987; Maynard Smith, 1984, 1989; Axelrod and Dion, 1988; Godfray, 1992; Sigmund, 1993). It consists of playing ***C*** in the first round, and then doing whatever the other player did in the previous round. Mistakes, however, severely undermine TFT's effectiveness, especially if both partners follow this strategy, giving rise to long runs of ***DD***, which can only be broken by a subsequent pair of mistakes (Boyd, 1989). When individuals remember both players' moves in the previous round, then PAVLOV (formerly known as SIMPLETON) sometimes does very well (Kraines and Kraines, 1989; Milinski, 1993; Nowak and Sigmund, 1993a, b). It repeats its previous move after ***CC*** and ***DC***, but switches behavior after ***CD*** and ***DD***. It therefore cooperates only after ***CC*** and ***DD***. It ousts TFT, and all-***C***, but cannot win against all-***D***. Since small clusters of TFT can invade populations of all-***D***, and PAVLOV beats TFT, which in turn can be invaded by all-***D***, an endless chaotic cycle of strategies occurs (Boyd and Lorberbaum, 1987; Nowak and Sigmund, 1989, 1990, 1992, 1993a, b; Koeslag, 1997). However, when PAVLOV is the majority strategy, it beats all-***D*** mutants if, after ***DD***, it defects with a small probability, w (Nowak and Sigmund, 1993a).

When individuals are placed in a two-dimensional spatial array, or lattice, and every individual plays the simple PDG with self and its immediate neighbors, cooperation and defection are found, under certain circumstances, to reach stable, quasi-stable, dynamic fractal or chaotically varying polymorphisms which can persist indefinitely (Nowak and May, 1992, 1993; Nowak et al., 1994a, b). In these models there are only two kinds of players; those who always cooperate and those who always defect. No account is taken of past or likely future encounters, so no memory is required and no strategic considerations arise. In each round every individual plays the PDG with a defined set of immediate neighbors, including self. The payoffs in these models differ from the standard payoffs in that ***DD*** is awarded zero (the same as ***CD***) instead of the standard 1. The payoff for ***DC*** is also often set at a value different from the standard 5 (Nowak and May, 1993). After everyone has played the PDG with all of its neighbors, each individual takes on the strategy of the neighbor with the highest score. If self had scored the highest payoff in the neighborhood then it does not change its strategy. The process is repeated for the next generation, and the next, for as long as the distribution of cooperators and defectors keeps changing. This spatial version of the PDG generates striking spatial patterns (described as "evolutionary kaleidoscopes") in which both cooperators and defectors can persist indefinitely. Furthermore, for certain parameter values, it is observed that regardless of the initial conditions, the frequency of cooperators always reaches the same proportion, raising the interesting possibility of the existence of a universal constant governing PDG interactions on a lattice. Many variations on this fundamental theme have now been described (Huberman and Glance, 1993; Nowak and May, 1993; Nowak et al., 1994a, b; Killingback and Doebeli, 1998; Szabo and Tõke, 1998; Hauert, 2001).

Although high payoffs in the Nowak models lead to greater "reproductive success" in neighborhood context, it does not do so in population context. A cooperator surrounded entirely by cooperators earns, in a 3 × 3 neighborhood, a final payoff of $[3 \times 9 = \mathbf{27}]$. A defector surround by eight defectors obtains a score of $[0 \times 9 = \mathbf{0}]$. However, since both are surrounded by individuals who obtained the same payoffs as themselves, the composition of neither neighborhood, nor that of the population as a whole, is affected by this difference in "fitness". The defector with zero fitness leaves as many "offspring" as does the cooperator with a fitness of 27. Furthermore, it is difficult to discern "parents" and offspring, or other aspects of reproductive biology in these models.

A further important feature is that games against self in small neighborhoods (where self constitutes a large fraction of the neighborhood) strongly favor cooperators, while being disadvantageous to defectors. The payoff matrix used in these models is similarly far more favorable to cooperators than is the standard PDG payoff matrix of 5 points for ***DC***, 3 for ***CC***, 1 for ***DD***, and zero for ***CD***. Thus, on a straight-line boundary between a group of cooperators and a group of defectors, in a 3 × 3 neighborhood, a cooperator earns $[(3 \times 0) + (6 \times 3) = \mathbf{18}]$ points, while a defector earns only $[(3 \times 5) + (6 \times 0) = \mathbf{15}]$ points. This is in marked contrast to what happens when the standard payoffs apply, and there are no self-interactions. Under these circumstances a cooperator earns $[(3 \times 0) + (5 \times 3) = \mathbf{15}]$, while the defector earns a much higher $[(3 \times 5) + (5 \times 1) = \mathbf{20}]$ points. In the Nowak models, therefore, the odds at the boundaries between cooperators and defectors are the reverse of those obtained with standard payoffs, and no self-interactions. Thus, contrary to expectation, an *increase* in the payoff to defectors for ***DC*** from 5 to 6, in the Nowak models, equalizes the fitnesses of cooperators and defectors facing each other across straight-line boundaries. It is particularly this property, and the absence of differences in fitness between clustered cooperators and clustered defectors, that gives rise to the fascinating patterns that characterize these models.

Taking all these factors into consideration, we have adapted the spatial PDG to include sex, mate selection, the production of offspring in proportion to fitness, and the Mendelian inheritance of parental traits. We use only the standard PDG payoff matrix, and proscribe self-interactions. The resulting model, which we term the "cornfield model", therefore differs drastically from the Nowak models. It was designed to simulate as realistically as possible (within the constraints of a computer simulation) a population of sexual creatures with limited mobility interacting socially and then reproducing in a uniformly fertile, square "cornfield". The only similarity between the two models is that individuals are always either committed cooperators or committed defectors, and that the PDG is played only with neighbors. In the Nowak models an individual plays the game with all of their neighbors, and with itself, and then adopts the strategy of whoever obtains the highest score in that neighborhood. The outcome is rigidly deterministic. In the cornfield model each individual plays, on average, 40 games (see Methods for details) with individuals chosen at random from the neighborhood. Neighborhood sizes are generally larger than those of the Nowak models, with the result that an individual does not necessarily play the PDG with everyone in its neighborhood, nor are two or more games with the same opponent excluded. The individual's fitness is then the sum of the payoffs of all of the games played. The scores obtained by the protagonist's opponents accumulate towards their fitness scores. When every individual in our model has played its 40 games, sexual encounters occur between individuals. These happen as follows: individuals are selected randomly one at a time. The probability of being selected is directly proportional to absolute fitness (i.e. the sum of the scores of all the PDGs played by the individual). A potential mate is then selected from the individual's neighborhood. Once again the selection is stochastic, with a probability which is directly proportional to absolute fitness. There is then mate choice, which will be explained later, followed by the production of one offspring (per successful encounter of this nature). Fitness is therefore not parochial as it is in the Nowak models; it is absolute. An individual with a fitness of 120 produces, on average, 3 times more offspring than one with a fitness of 40 (in the absence of mate choice). An individual with zero fitness cannot reproduce, regardless of the distribution of fitnesses within its neighborhood, or of its sexual attractiveness.

Since fit individuals produce multiple offspring, these offspring cannot all be accommodated in the parental niche in the cornfield. (Each niche can accommodate only a single individual.) The rule that was applied was that offspring inherited their "mother's" niche if that niche had not already been allocated to an earlier sib or half-sib. (The term "mother" is explained in the Methods.) If the maternal niche had already been allocated, then the offspring spiraled round the maternal niche in step-wise widening circles until it encountered an unallocated niche, which would then become its niche in the next generation. That niche could be located outside the maternal neighborhood, although most offspring found vacant niches within, or very close to, the maternal neighborhood. Our populations are therefore less viscous than those of the Nowak models, where offspring never settle beyond the "parent's" neighborhood. A population is termed "viscous" if individuals do not disperse far from their places of birth (Hamilton, 1964). The picture is, however, complicated in the Nowak models by the fact that although offspring are seemingly produced, it is difficult to ascertain who the parent or parents might be. Individuals simply change their strategy to the most successful strategy in their neighborhood.

When the cornfield model was run without mate choice (i.e. as described above) defection always eliminated cooperation within 6 generations. (Generation 0 always consisted of equal numbers of cooperators and defectors, distributed randomly across the cornfield.) No trace could be discerned of the stable, quasi-stable, dynamic fractal or chaotically varying admixtures of cooperators and defectors, which characterize the Nowak models. Cooperation was simply, systematically and very rapidly driven to extinction. This result was neither surprising, nor unexpected.

The quest for a fit sexual mate had a very dramatic and surprising influence on the evolution of cooperation in the cornfield.

Fitness is defined as the heritable propensity to produce more grandoffspring and greatgrandoffspring than other individuals. In real life, fitness is therefore defined by future events. It can consequently only ever be measured in retrospect. Current fitness is impossible to measure. A sexual creature searching for a fit mate is therefore obliged to rely on past performance when evaluating competing phenotypic features, and to take calculated risks with regard to their present fitness values (which will only become apparent in two or three generation's time).

Natural selection ensures that, as a general rule, maladaptive traits tend to be concentrated at the lower end of the allelic frequency spectrum, while highly adaptive traits are concentrated at its upper end. This is, in fact, the basis for natural selection as an evolutionary force. Frequency analysis is therefore the only best-bet method of evaluating the phenotypic features of prospective sexual mates. It is also the only way in which mutants, as a general category, can be recognized and avoided. The result is koinophilia: the avoidance of unusual, strange, or unexpected phenotypic features in prospective sexual mates (Koeslag, 1990, 1994, 1995, 1997; Koeslag and Koeslag, 1993, 1994a, b).

In keeping with these theoretical considerations, humans clearly find young, modal or average faces the most attractive (Langlois and Roggman, 1990; Etcoff, 1994; Enquist and Ghirlanda, 1998; Perrett et al., 1998). However, Perrett et al. (1994) found that both men and women found a slightly off-modal female face the most attractive from a wide range of women's faces with neutral expressions and identical hairstyles. When the non-modal features were slightly exaggerated the face was judged to be more beautiful still. Close examination of the photos in Perrett, May and Yoshikawa's paper shows, in fact, that the exaggerated face looks younger than the modal female face (a composite of women's faces aged 22–46 years). The differences in appearance are, however, very small, and, to many people, not immediately obvious. Since the same result was obtained with Japanese subjects, these findings are probably culture independent, and would indicate that people, generally, find *young,* modal female faces sexually the most attractive (Langlois and Roggman, 1990), as expected.

It is easy to see how the avoidance of atypical, unexpected or unusual phenotypic features in a sexual mate encourages stabilizing selection. Consider the simplest form of the PDG. Without mate selection the payoffs for ***CC***, ***CD***, ***DC***, and ***DD*** are, typically, 3, 0, 5 and 1, respectively. Under these circumstances, the only evolutionarily stable strategy (ESS) is consistent defection. However, if, in a population of unconditional cooperators (by whatever means that might have arisen), koinophilia leads to discrimination against a minority of defectors at the time of mating (with, say, a 50% probability of rejection of a potential mate if that mate has been behaving in an unusual manner), then the payoffs for ***CC***, ***CD***, and ***DC*** become 3, 0, and [$5 \times 0.5 = \mathbf{2.5}$], respectively. Note that the payoff for ***DD***, [$1 \times 0.5 = \mathbf{0.5}$], does not occur in this case, or is, by definition, very rare. This renders cooperation an ESS, meaning that a koinophilic population of cooperators cannot be invaded by *small* numbers of defectors (Maynard Smith, 1974, 1982, 1984; Parker, 1984; Vickers and Cannings, 1987). In fact, it has been demonstrated that koinophilia is capable of evolutionarily stabilizing any of the 16 possible IPD strategies against invasion by mutants practicing any one of the 15 competing strategies (Koeslag, 1997). Without koinophilia the model cycles chaotically through the different strategies in the manner described by Boyd and Lorberbaum (1987) and Nowak and Sigmund (1989, 1990, 1992, 1993a, b).

In order to investigated whether koinophilia can do more than just stabilize cooperation against invasion by *small* numbers of defectors, the starting population (Generation 0) in the cornfield model always consisted of *equal* numbers of cooperators and defectors. This model therefore tested whether cooperation can outcompete selfishness in a straight, initially equally matched, evolutionary battle. (We already know that without koinophilia cooperation is rapidly eliminated from the field by the defectors.)

Since koinophilia tends to make a local majority phenotype locally more common still (Koeslag, 1995), the model was designed to test whether this form of mate choice would convert a random distribution of cooperators and defectors (in Generation 0) into a series of clusters, each consisting predominantly of either cooperators or of defectors. If it does so before cooperation goes to extinction, then the individuals making up a cluster of cooperators would each have an average fitness of 120, whereas the individuals making up clusters of defectors would each have an average fitness of only 40. However, where clusters abut one on another, the defectors at the boundary would inevitably outcompete the cooperators on the opposite side of the boundary (see earlier discussion). The overall outcome therefore presumably depends on the model parameters. The model was therefore tested under different strengths of koinophilic discrimination, different neighborhood sizes, and different population densities. We also tested the effect of increasing the cornfield's size, and of varying the accuracy with which individuals remembered which PDG strategy had been the more common in their neighborhood. The PDG payoff matrix was, however, never changed, and self-interactions were never allowed.

2. The model

A two-dimensional spatial array, or lattice, comprising 50×50 identical "niches" (each "niche" could be occupied by no more than a single individual) was created using the Turbo Pascal computer programming language. Since the lattice can be visualized as a uniformly fertile, square cultivated cornfield, it will, for ease of description, henceforth be referred to as "the cornfield", from which the model derives its name.

The cornfield was inhabited, in the default situation, by 1500 haploid hermaphrodites, each occupying one niche. There was no niche sharing. In the default situation only 60% of the available niches were therefore occupied. Each individual could interact with the individuals in its immediate "neighborhood" (a square consisting of 5^2, 7^2, or 9^2 niches). Each individual occupied the niche at the exact geographic center of its neighborhood, unless it lived near or on the edge of the cornfield. In that case the neighborhood size remained the same as elsewhere in the field, but the individual no longer necessarily occupied the central niche. Except at the boundaries, therefore, every individual's neighborhood was unique to that individual, although there was always considerable partial overlap with those of its neighbors, and, to a lesser extent, with those of individuals just outside the protagonist's neighborhood. At the cornfield boundaries 3 individuals shared the same neighborhood when neighborhood size was 5^2, and 5 shared a common neighborhood when neighborhood size was 9^2. Neighborhood size, for any given run of the model was the same for all individuals in every generation. The default neighborhood size was a 9×9 square of (occupied and vacant) niches.

All individuals were genetically identical except for one gene, which occurred in the form of two alleles. The one allele was termed "selfish", and the other "cooperative". Since everyone was haploid, the population therefore consisted of only two phenotypes, also known as selfish and cooperative. Individuals sought mates only from within their neighborhoods. Genetic inheritance was strictly Mendelian. Each parent therefore had a 50% probability of passing on its allele (selfish or cooperative) to their common offspring. In a heterozygous mating there was therefore a 50% probability that the offspring was cooperative, and the same probability that it was selfish. There were no mutations. Generations were discrete, and, for any given run of the model, always of the same size.

In Generation 0 individuals were randomly allocated niches in the cornfield until the desired population size was reached. The individuals' selfish or cooperative genotypes were also randomly assigned, in a 50:50 ratio (Fig. 4a). The stochastic effects were produced by the Real Version 1 Minimal Standard random number generator described by Park and Miller (1988).

At the beginning of each generation each individual initiated 20 PDGs, each time with a different, randomly chosen individual from within its neighborhood, but *never including self*. (Each individual therefore played 20 compulsory games as the initiator or protagonist of those games, plus a further approximately 20 games as an opponent in games initiated by its neighbors. Everyone therefore played, on average, 40 games.) The individuals' strategies were rigidly determined by their individual genotypes. A genetically selfish individual always defected when playing the PDG, whereas a genetically cooperative individual always cooperated. Although the game was played many times over, sometimes, by chance, with the same partners, no one had the capacity to learn from their experience, or to develop strategies that might prove useful in the future. Genotypic cooperators always cooperated, and genetically selfish individuals always defected. Individuals did however remember, with 95% accuracy, how many times their opponents had defected and how many times, in their experience, opponents had cooperated, and which of the two strategies was, in their slightly inaccurate estimation, the more common (in their neighborhood). The "95% accurate memory" was implemented as follows: after each PDG, of which each individual played on average 40, the individual remembered his opponent's strategy, *C* or *D*, correctly with a probability of 0.95. On average, therefore, two opponents' strategies out of 40 would be remembered incorrectly. The individual's estimate of what had been the more common strategy was used to choose a sexual mate.

At the end of each game the pair of players were individually awarded a fitness score which depended on the outcome of the game. The standard fitness payoffs applied: three for ***CC***, one for ***DD***, five for ***DC***, and zero for ***CD*** (May, 1981; Maynard Smith, 1989).

The individual's final fitness was the sum of the scores of all the games played by that individual. A cooperative individual who happened to be surrounded entirely by cooperative phenotypes would have, on average, a final fitness of 120; whereas a defector could accumulate a final fitness score of approximately 200 if similarly surrounded by cooperators, but only 40 if surrounded entirely by defectors. A cooperator entirely surrounded by defectors would always have a final fitness of zero, and would therefore not reproduce.

On completion of all the tournaments, individuals were chosen randomly from the population to search for a mate. The probability, P_1, of being chosen for this activity was $P_1 = [Fitness_i/Fitness_P]$, where $Fitness_i$ is the individual's fitness (i.e. the individual's final score for all the games it had played) and $Fitness_P$ is the sum of the fitnesses of all the individuals in the entire population. The selected individual (whom we term, for ease of description, the potential mother) then searched for a mate from within her neighborhood. The probability, P_2, of a given neighbor meeting the potential mother was $P_2 = [Fitness_m/Fitness_N]$, where $Fitness_m$ is the neighbor's fitness, and $Fitness_N$ is the sum of the fitnesses of all the individuals in that neighborhood, *ex*cluding the potential mother.

Pairs of individuals "met" in the manner described above, but such encounters did not automatically lead to mating and reproduction. Every member of the population had a mate preference for the most common phenotype in its experience. Thus, if, according to the individual's memory, defection had been the most common strategy in its neighborhood, then, regardless of that individual's own phenotype, it preferred mates that had consistently defected during the tournaments. Similarly if cooperation had been the more common strategy, then the individual preferred a mate who had been a consistent cooperator. If cooperation and defection had been equally common according to the individual's memory, then that individual showed no mate preference. Note that this is not assortative mating, but koinophilic mating. An individual, regardless of its own phenotype (cooperator or defector), preferred mates who, according to its slightly faulty memory, had displayed the most common behavior in its experience (excluding its own behavior). This experience was limited to encounters with individuals from its neighborhood.

The sexual attractiveness, A, of perceived majority behavior during the tournaments was always unity. The sexual attractiveness of minority behavior was 0.3 in the default situation. Since both individuals in any given sexual encounter practiced mate preference, the probability of mating after such an encounter was the product of the two attractivenesses. Thus, if both individuals had indulged in majority behavior (as individually perceived by both partners) then the probability of mating (after the encounter) was 1. If both had indulged in minority behavior the probability of mating was $0.3^2 = 0.09$. If only one had indulged in minority behavior (as judged by the other) the probability of mating was 0.3. Mating always resulted in a single offspring, who settled (on reaching maturity—i.e. in the next generation) in the same niche in the cornfield as its mother. If that site was already occupied by an older sib or half-sib (from an earlier mating of the mother), then that offspring migrated (on reaching maturity, when the current generation was replaced by its successor) to the geographically nearest vacant niche, which could, potentially, be far removed from the parental neighborhood. In most cases however, offspring found vacant niches within or very close to the maternal neighborhood.

Future niches were allocated, in the computer program, to offspring at the time of "birth". Since these occurred sequentially, in the program, the offspring of the very first mating in any given generation was always assured of inheriting its mother's niche in the next generation, because all future niches in the entire cornfield were still unallocated. If, by chance, that same mother produced the second baby of that generation, it would be allocated an adjoining niche, since the ancestral niche had already been allocated to the older sib, or half-sib. This meant that should the individual who currently owned that niche produce a child of her own, it would have to migrate to the geographically nearest unallocated niche on reaching maturity. Thus, the last offspring to be born to the outgoing generation was very likely (but not necessarily) to have to migrate some distance from its parental niche to find an unallocated site where it could settle to form part of the next generation.

The random selection of pairs in proportion to their final fitness scores, mate choice, reproduction and the allocation of niches (in the cornfield of the new generation) continued until a new population of 1500 individuals was created.

The size of the cornfield remained unchanged for all runs of the model. Population size varied, however, between 500, 1000, and **1500**. The value in bold denotes the default value. Neighborhood sizes varied between 5^2, 7^2, and $\mathbf{9^2}$ niches. The strength of discrimination, d, varied between 0.0, **0.7**, 0.8 and 0.9. The strength of discrimination, d, denotes the probability that an individual will find a potential sexual partner, who had displayed minority behavior, unattractive enough *not* to mate with her (i.e. $[d = 1 - A]$, where A is sexual attractiveness, described above). A strength of discrimination of zero indicates that individuals are 100% willing to mate with phenotypically unusual partners, and that there is therefore no mate choice. The accuracy of memory varied between 0.90, **0.95**, and 1.00.

The model was always run (either under entirely default conditions, or with a single parameter changed from its default value) until one or other genotype had been eliminated from the population. Each experiment (under a given set of parameter values) was repeated 50 times. The results of the 50 experiments are reported, in successive generations, as the mean (±SEM) proportion of cooperators in the population (Figs. 1–3). The starting condition was always 50% cooperators.

3. Results

In the absence of koinophilia (i.e. the strength of discrimination = 0.0) defection eliminated cooperation within 6 generations from the start of every single run of the program (Fig. 3). There was no tendency to form the stable, quasi-stable, dynamic fractal or chaotically varying admixtures of cooperators and defectors that persisted indefinitely, as described by Nowak and May (1992, 1993) and by Nowak et al. (1994a, b). Cooperation was very simply rapidly driven to extinction.

The results of the operation of koinophilia are depicted in Figs. 1–3. In all cases the concentration of cooperators rapidly decreased from the starting value of 50% to a mean of between 5% and 27% during the first 2–4 generations, but thereafter increased. If the initial decrease was to not less than *about* 21 ± SD 3.36% then cooperation appeared always eventually to go to fixation. This meant that in an individual run cooperation would almost certainly ultimately win if only $[\text{mean} - 2\ \text{SD}] = [21 - (2 \times 3.36) = \mathbf{14.3\%}]$ of the population consisted of cooperators in Generation 3. (It should however be noted that this is

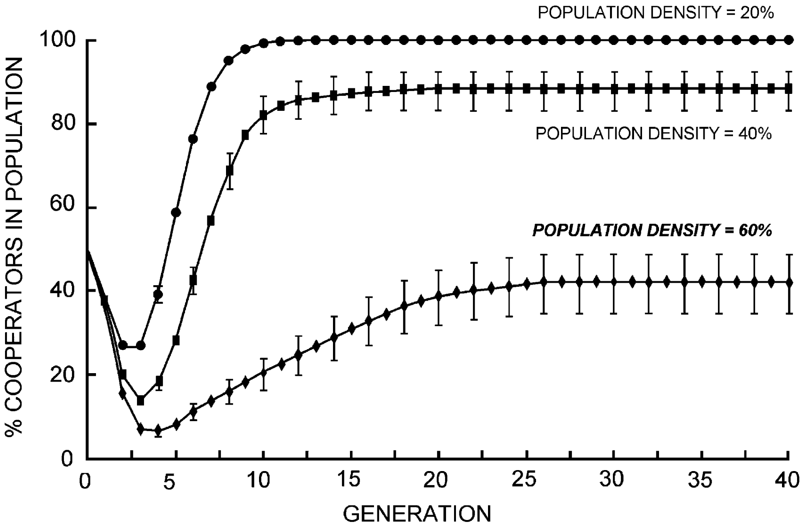


Fig. 1. The evolution of cooperation in the cornfield model (see text) under default conditions (**population density = 60%**, neighborhood size = 9^2 niches, strength of discrimination = 0.7, accuracy of memory = 0.95), and when population density was 40% and 20%, with the other parameters kept at their default values. Each run of each experiment started in Generation 0 with a situation similar to the one depicted in Fig. 4a, and ended when either cooperation or defection had reach 100% concentration. The results show the mean ±SEM concentration of cooperators in successive generations of 50 runs of the same experiment under the identical parameter conditions.

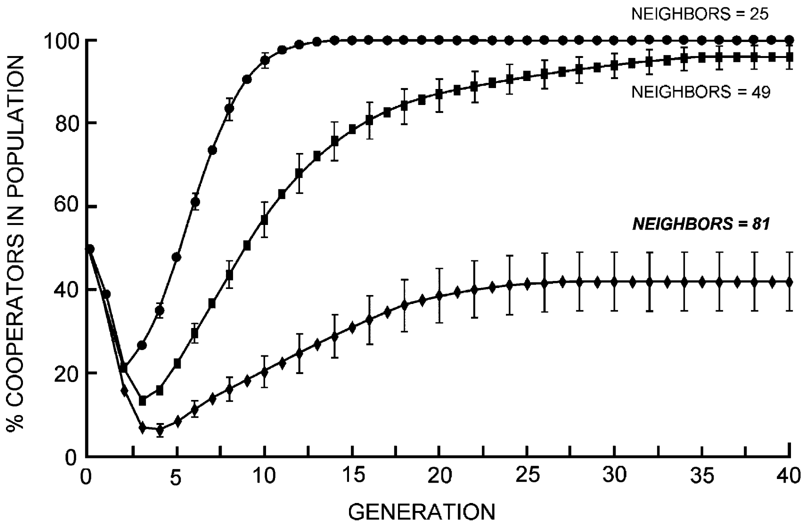


Fig. 2. The evolution of cooperation in the cornfield model (see text) under default conditions (population density = 60%, **neighborhood size = 9^2 niches**, strength of discrimination = 0.7, accuracy of memory = 0.95), and when neighborhood size was 7^2 and 5^2 niches, with the other parameters kept at their default values. Each run of each experiment started in Generation 0 with a situation similar to the one depicted in Fig. 4a, and ended when either cooperation or defection had reach 100% concentration. The results show the mean ±SEM concentration of cooperators in successive generations of 50 runs of the same experiment under the identical parameter conditions.

an approximation derived from the results of the parameter settings that produced the lowest dip in the concentration of cooperators, but then nevertheless still produced a final 100% cooperator fixation rate. It was not possible, in this stochastic model, to deterministically manipulate the nadir of the cooperator's fortunes.) Presumably if all 215 (= 14.3% of 1500 individuals) cooperators formed a single cluster then they were very likely to eventually eliminate defection; but if scattered across many smaller clusters they might not be able to do so. The results, however, seem to indicate that, when there were 215 or so surviving cooperators in Generation 2–4, then they were nearly always grouped in such a way that they could exploit the group benefits of cooperation, and thus outcompete the rest of the field which was by then earning only

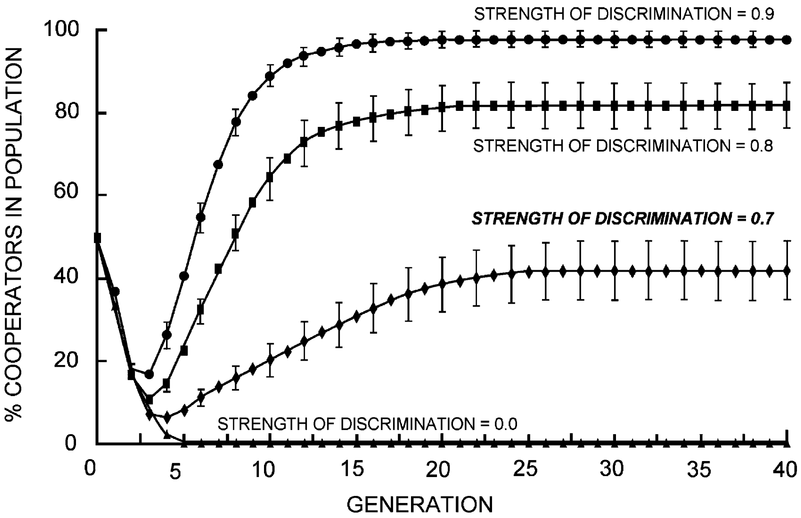


Fig. 3. The evolution of cooperation in the cornfield model (see text) under default conditions (population density = 60%, neighborhood size = 9^2 niches, **strength of discrimination = 0.7**, accuracy of memory = 0.95), and when the strength of discrimination was 0.0, 0.8 and 0.9, with the other parameters kept at their default values. Each run of each experiment started in Generation 0 with a situation similar to the one depicted in Fig. 4a, and ended when either cooperation or defection had reach 100% concentration. The results show the mean ± SEM concentration of cooperators in successive generations of 50 runs of the same experiment under the identical parameter conditions.

one third of the fitness points that the clustered cooperators were earning.

Under default conditions (including the stated starting configuration: a 50×50 niche cornfield size, and a 50:50 random distribution of cooperators and defectors) cooperation became fixed in 42% of populations. A decrease in population density (Fig. 1), a reduction in neighborhood size (Fig. 2), and an increase in the strength of discrimination (Fig. 3) could each, independently, cause cooperation to attain 100% (or very close to 100%) fixation rate. The effect of changes in these parameter values was dose dependent.

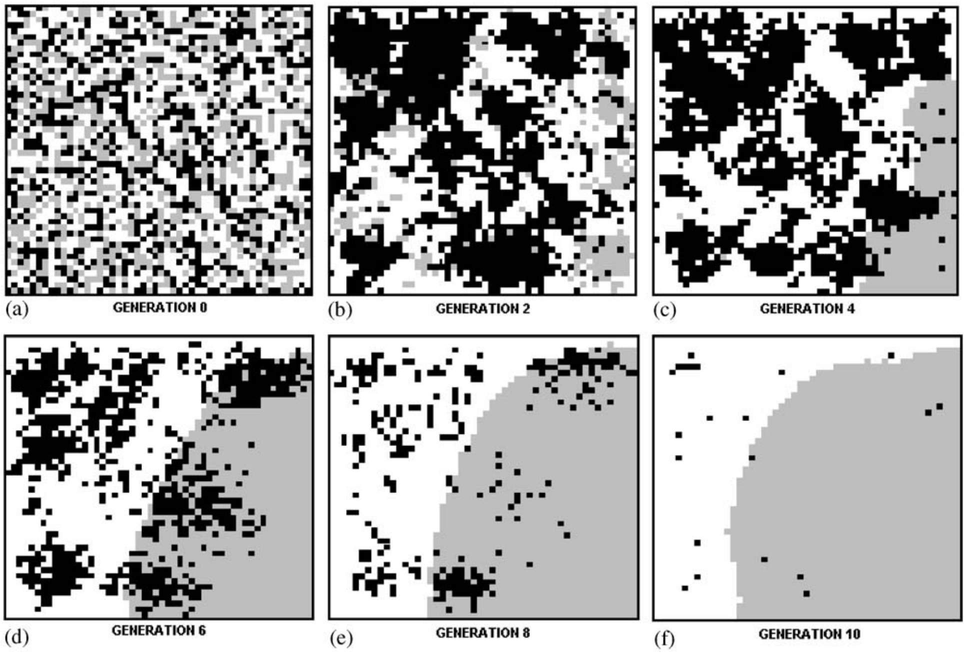
This suggests that the evolution of cooperation in the cornfield model relied heavily on the statistics of sampling. Small samples (small neighborhoods or low population density) result in large sampling "errors" (i.e. large variations in the $C : D$ ratios of those samples), causing some neighborhoods to consist, by chance, predominantly of C and others mainly of D, though most have, of course, intermediate compositions. The greater these random differences in $C : D$ ratio the easier they could presumably be amplified by koinophilia into the clear almost homogeneous clusters seen in, for instance, Fig. 4c. This seems to be a prerequisite for the koinophilia aided evolution of cooperativity in this model: the greater the chance of a randomly high neighborhood $C : D$ ratio somewhere on the cornfield, the greater the number of times that cooperation ultimately won the evolutionary race (Figs. 1–3). Random mating, within neighborhood context, never accentuated any random clustering, and led to an evolutionary zero success rate for cooperation.

An illustration of koinophilia aided clustering in operation can be seen in Figs. 4a–f which depicts the distribution of cooperators and defectors in Generations 0, 2, 4, 6, 8, and 10 of a randomly chosen run of the model (neighborhood size = 7^2; population density = 60%; strength of discrimination = 0.7; memory accuracy = 0.95). Three small clusters of cooperation had developed on the right-hand side of the field in Generation 2. These coalesced by Generation 4, at which time the rest of the cornfield was almost devoid of cooperators. From there on cooperation expanded slowly but inexorably to eliminate defection in 9 further generations.

It is interesting to note that even in Generation 10 there were still the occasional defectors deeply embedded in cooperative territory. Despite their very high fitness (200 compared to the 120 of the average cooperator in a cluster) they were eliminated in the next 2 generations because of their perceived unconventional behavior within the cooperative clusters.

Perfect memory, and therefore more accurate judgement as to which strategy had been the more common in an individual's experience, increased koinophilia's ability to eliminate defectors. Thus, with perfect memory under default conditions cooperativity went to fixation in 60% of cases instead of in only 42% when memory was 95% accurate. A 90% accurate memory led to cooperativity winning in only 34% of the 50 runs.

Enlargement of the size of the cornfield from a 50×50 lattice to a 100×100 lattice increased koinophilia's ability to eliminate defection from 42% to 90%, under



(a) GENERATION 0 (b) GENERATION 2 (c) GENERATION 4

(d) GENERATION 6 (e) GENERATION 8 (f) GENERATION 10

Fig. 4. (a–f) The distribution of cooperators (▨) and defectors (■) in a cornfield of 50 × 50 niches, in Generations 0, 2, 4, 6, 8, and 10 of a randomly chosen run of the model (neighborhood size = 7^2; population density = 60%; strength of discrimination = 0.7; memory accuracy = 0.95). In this particular run cooperation eliminated defection in 13 generations. The distribution of cooperators and defectors (in a 50:50 ratio) is random in Generation 0. By Generation 2 three small clusters of predominant cooperation had developed on the right hand side of the field. These coalesced in Generation 4, by which time the rest of the cornfield was almost devoid of cooperators. From there on cooperation expanded slowly but inexorably to eliminate defection in 9 further generations. Note that even in Generation 10 there were still the occasional defectors deeply embedded in cooperative territory. Since they are the fittest individuals in the cornfield they would, in the absence of koinophilia, rapidly proliferate, and quickly eliminate cooperativity from the field.

otherwise default conditions. The larger lattice presumably increased the probability that a sufficiently large cluster of cooperators would develop somewhere on the field.

4. Discussion

In the spatial PD games described by Nowak and May (1992, 1993) and Nowak et al. (1994a, b) stable, quasi-stable, dynamic fractal and chaotically varying admixtures of cooperators and defectors persisted indefinitely, indicating that equilibrium situations between cooperation and defection are possible when the simple PDG is played with neighbors on lattices similar to our cornfield. The persistence of cooperation in these models is, however, entirely dependent on the populations' enormous viscosity (Nowak and May, 1993; Killingback and Doebeli, 1998; Hauert, 2001), and on games played against self in small neighborhoods, in combination with a modified PDG payoff matrix which strongly favors cooperators, while strongly disadvantaging defectors. This is especially critical at the boundaries between groups of cooperators and defectors.

Our model relied on sexual reproduction, and the production by fit individuals of multiple offspring who, while preferring to remain close to their "mother's" niche in the cornfield, sometimes had to migrate considerable distances to find an unoccupied niche for when it became their turn, in the next generation, to engage in social interactions (PD games), leading to the search for mates and the production of offspring of their own. Our populations were therefore considerably less viscous than those of Nowak and May (1992, 1993) and Nowak et al. (1994a, b). There were, furthermore, no self-interactions, and only the standard PDG payoff matrix applied. Under these circumstances, in the absence of koinophilia, defection always rapidly eliminated cooperation within a few generations. There are no theoretical considerations to suppose that the model should have behaved any differently.

The results show that, under these deliberately unfavorable circumstances, koinophilia was not only able to stabilize cooperation, but to turn it into an aggressive evolutionary force, which was capable of eliminating defection altogether. Only under very special parameter circumstances do the other spatial models accomplish this.

Like the other spatial models (Nowak and May, 1992, 1993; Nowak et al., 1994a, b; Killingback and Doebeli, 1998; Hauert, 2001), koinophilia produces these results by causing phenotypic clustering. The cause of the clustering is, however, more pedestrian and intuitively obvious than in the other spatial games. A preference for mates displaying the most common phenotype will inevitably cause that phenotype to become more common still (Koeslag, 1995). In the presence of a moderate degree of population viscosity, random variations in the local frequencies of defectors and cooperators are therefore rapidly turned into near-homogeneous clusters of one phenotype or the other. Since the clusters of cooperators are fitter than their selfish counterparts, the cooperative clusters will produce more offspring than the defectors, and therefore grow at the expense of the selfish clusters. In our models cooperativity's evolutionary success was, therefore, dependent largely on whether it could survive the first few generations before koinophilia was able to amplify a random partial cluster of cooperators into a dense non-random knot of cooperators (see Figs. 4a–c). Once such a cluster of cooperators had formed it generally grew, and aggressively invaded the rest of the field like a rapidly growing cancer (Figs. 4d–f).

Ironically, the factor in the traditional spatial models that causes the clustering and is partially responsible for cooperativity's survival, namely the high population viscosity, is also the factor that prevents cooperativity from growing at defection's expense. In the Nowak models, the fitness of a cooperator inside a cooperative cluster is exactly the same as that of a defector inside a selfish cluster. Each round of PD games dictates that they both stay the way they were. It is only at the boundaries between clusters of cooperators and defectors that evolution occurs. Not so in our model. Cooperative individuals in the middle of cooperative clusters produce 3 times as many offspring as do defectors in the middle of selfish clusters. Generally, this more than compensates for the disadvantages incurred by cooperators at boundaries with clusters of defectors (see, for example, Fig. 4c). Here cooperators are less fit than their cheating neighbors (see the Introduction), and are hardly protected by koinophilia, especially if they are on straight or convex boundaries. The evolutionary dynamics at the boundaries between cooperation and defection in the Nowak models differ substantially from those in the cornfield model. Here cooperators generally have the advantage over defectors, or are equally matched (see the Introduction). The result is the "evolutionary kaleidoscopes" for which these models have become famous.

Figs. 4e and f indicate that koinophilia continues to play a crucial role in the evolution of cooperation after the initial formation of clusters. Isolated defectors in the middle of cooperative clusters (see Figs. 4e and f) are the fittest individuals in the field. Their fitnesses are at least 67% higher than those of the fittest cooperators. (These defectors have an average fitness of 200 versus the 120 of the fittest cooperators.) Under normal circumstances they would therefore rapidly proliferate, and ultimately replace the cooperative cluster from within, emphasizing the evolutionary conundrum posed by cooperativity. This does not happen if the population is koinophilic. The unusual behavior of the defectors, in the midst of a cluster of cooperators, leads to discrimination against them during mate seeking. This then reduces their chances of reproduction. Indeed, if the strength of discrimination is greater than 0.4, then the advantage they gain from cheating is canceled out (except in relation to the halo of cooperators in their immediate vicinity who had played PD games with them, and who had therefore earned zero points for those particular games—in a 7×7 neighborhood their average fitness is 117.6 instead of 120). With koinophilia these isolated defectors can therefore fare worse in the evolutionary game than the surrounding cooperators. In our models with strengths of discrimination of $\geqslant 0.7$ they never posed an evolutionary threat to the cooperators.

If the conditions of our model are considered not to be overly contrived, then it is likely that cooperation, in its many and varied guises, could be widespread in nature, even in the absence of inclusive fitness considerations (Hamilton, 1964; Maynard Smith, 1964), or the special conditions required by the Nowak spatial models. Both would, of course, contribute to the evolution of altruism, but here we have shown that there is an aggressively powerful additional evolutionary force that can, potentially, maintain cooperativity and altruism single handedly. This, therefore, suggests that it is individual advantages that are evolutionarily unstable in the presence of their rival group advantages, rather than the other way around. The "unit of natural selection" (Maynard Smith, 1964, 1987; Williams, 1966; Lewontin, 1970; Wilson, 1983; Boyd and Richerson, 1990), in such cases, is therefore the community rather than its individual members.

In our model, in fact, the group advantage (all-***C***) eliminated its rival individual advantage (all-***D***) with an ease that is normally associated with competitions between highly beneficial traits and their maladaptive counterparts (Figs. 1–3). This suggests that even if koinophilia is only a part determinant of mate choice it would have major evolutionary consequences, that would require drastic reappraisals of the teleologies that are normally advanced for biological observations. In particular, the requirement that phenotypic features (physical or behavioral) may only be explained in terms of selfish advantages, becomes highly questionable.

The results of this study suggest that outright group advantages are not only evolutionarily permissible, but might even dominate behavioral biology. Flying in V-formation, hunting in packs, alarm signals, the swarming of a school of fish on being threatened by a predator, expressions of the emotions, ritualized behavior, and dawn or evening choruses are all more easily explained in terms of *exploitable* cooperation than in terms of purely selfish imperatives. It has, however, always been assumed that the group advantageous explanations for these phenomena must be resisted at all costs (Krebs and Dawkins, 1984; Dawkins 1989). The present work suggests that the group advantage explanation for these and many other phenomena might well apply.

In other words, to use an extreme example, behavior that promotes population control (Wynne-Edwards, 1962), while apparently flying in the face of Darwinian first principles, *may well*, if these results are correct, be the direct and inescapable result of hard-core Darwinian selection.

Acknowledgements

The authors would like to thank two anonymous reviewers for their helpful suggestions and encouragement.

References

Axelrod, R., Hamilton, W.D., 1981. The evolution of cooperation. Science 211, 1390–1396.

Axelrod, R., Dion, D., 1988. The further evolution of cooperation. Science 242, 1385–1390.

Boyd, R., 1989. Mistakes allow evolutionary stability in the repeated Prisoner's Dilemma game. J. Theor. Biol. 136, 47–56.

Boyd, R., Lorberbaum, J.P., 1987. No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. Nature, London 327, 58–59.

Boyd, R., Richerson, P.J., 1990. Group selection among alternative evolutionarily stable strategies. J. Theor. Biol. 145, 331–342.

Dawkins, R., 1989. The Selfish Gene. Oxford University Press, Oxford.

Enquist, M., Ghirlanda, S., 1998. The secret of faces. Nature, London 394, 826–827.

Etcoff, N., 1994. Beauty and the beholder. Nature, London 368, 186–187.

Godfray, H.C.J., 1992. The evolution of forgiveness. Nature, London 355, 206–207.

Hamilton, W.D., 1964. The genetical evolution of social behaviour. J. Theor. Biol. 1, 1–52.

Hauert, C., 2001. Fundamental clusters in spatial 2 × 2 games. Proc. R. Soc. London B 268, 761–769.

Huberman, B.A., Glance, N.S., 1993. Evolutionary games and computer simulations. Proc. Natl Acad. Sci. USA 90, 7716–7718.

Killingback, T., Doebeli, M., 1998. Self-organized criticality in spatial evolutionary game theory. J. Theor. Biol. 199, 335–340.

Koeslag, J.H., 1990. Koinophilia groups sexual creatures into species, promotes stasis, and stabilizes social behaviour. J. Theor. Biol. 144, 15–35.

Koeslag, J.H., 1994. Koinophilia replaces random mating in populations subject to mutations with randomly varying fitnesses. J. Theor. Biol. 171, 341–345.

Koeslag, J.H., 1995. On the engine of speciation. J. Theor. Biol. 177, 401–409.

Koeslag, J.H., 1997. Sex, the Prisoner's Dilemma Game, and the evolutionary inevitability of cooperation. J. Theor. Biol. 189, 53–61.

Koeslag, J.H., Koeslag, P.D., 1993. Evolutionarily stable meiotic sex. J. Hered. 84, 396–399.

Koeslag, J.H., Koeslag, P.D., 1994a. Koinophilia. J. Theor. Biol. 167, 55–65.

Koeslag, P.D., Koeslag, J.H., 1994b. Koinophilia stabilizes bi-gender sexual reproduction against asex in an unchanging environment. J. Theor. Biol. 166, 251–260.

Kraines, D., Kraines, V., 1989. Pavlov and the Prisoner's Dilemma. Theory Decision 26, 47–79.

Krebs, J.R., Dawkins, R., 1984. Animal signals; mind-reading and manipulation. In: Krebs, J.R., Davies, N.B. (Eds.), Behavioural Ecology. An Evolutionary Approach. Blackwell Scientific, Oxford, pp. 380–402.

Langlois, J.H., Roggman, La., 1990. Attractive faces are only average. Psychol. Sci. 1, 115–121.

Lewontin, R., 1970. The units of selection. Ann. Rev. Ecol. Syst. 1, 1–18.

May, R.M., 1981. The evolution of cooperation. Nature, London 292, 291–292.

May, R.M., 1987. More evolution of cooperation. Nature, London 327, 15–17.

Maynard Smith, J., 1964. Group selection and kin selection. Nature, London 201, 1145–1147.

Maynard Smith, J., 1974. The theory of games and the evolution of animal conflicts. J. Theor. Biol. 47, 209–221.

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

Maynard Smith, J., 1984. Game theory and the evolution of behaviour. Behav. Brain Sci. 7, 95–101.

Maynard Smith, J., 1987. How to model evolution. In: Dupré, J. (Ed.), The Latest on the Best. Essays on Evolution and Optimality. MIT Press, Cambridge, MA, pp. 119–131, 147–149.

Maynard Smith, J., 1989. Evolutionary Genetics. Oxford University Press, Oxford, pp. 168–169, 181–183.

Milinski, M., 1993. Cooperation wins and stays. Nature, London 364, 12–13.

Nowak, M.A., May, R.M., 1992. Evolutionary games and spatial chaos. Nature, London 359, 826–829.

Nowak, M.A., May, R.M., 1993. The spatial dilemmas of evolution. Int. J. Bifurcation Chaos 3, 35–78.

Nowak, M., Sigmund, K., 1989. Oscillations in the evolution of reciprocity. J. Theor. Biol. 137, 21–26.

Nowak, M., Sigmund, K., 1990. The evolution of stochastic strategies in the Prisoner's Dilemma. Acta Appl. Math. 20, 247–265.

Nowak, M., Sigmund, K., 1992. Tit for tat in heterogeneous populations. Nature, London 355, 250–253.

Nowak, M., Sigmund, K., 1993a. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature, London 364, 56–58.

Nowak, M., Sigmund, K., 1993b. Chaos and the evolution of cooperation. Proc. Natl Acad. Sci. USA 90, 5091–5094.

Nowak, M.A., Bonhoeffer, S., May, R.M., 1994a. Spatial games and the maintenance of cooperation. Proc. Natl Acad. Sci. USA 90, 7716–7718.

Nowak, M.A., Bonhoeffer, S., May, R.M., 1994b. More spatial games. Int. J. Bifurcation Chaos 4, 33–56.

Park, S.K., Miller, K.W., 1988. Random number generators: good ones are hard to find. Commun. Assoc. Comput. Mach. 31, 1192–1201.

Parker, G.A., 1984. Evolutionarily stable strategies. In: Krebs, J.R., Davies, N.B. (Eds.), Behavioural Ecology. An Evolutionary Approach. Blackwell Scientific, Oxford, pp. 30–61.

Perrett, D.I., May, K.A., Yoshikawa, S., 1994. Facial shape and judgements of female attractiveness. Nature, London 368, 239–242.

Perrett, D.I., Lee, K.J., Penton-Voak, I., Rowland, D., Yoshikawa, S., Burt, D.M., Henzi, S.P., Castles, D.L., Akamatsu, S., 1998. Effects of sexual dimorphism on facial attractiveness. Nature, London 394, 884–887.

Sigmund, K., 1993. Games of Life. Oxford University Press, Oxford, pp. 180–206.

Szabo, G., Tõke, C., 1998. Evolutionary prisoner's dilemma game on a square lattice. Phys. Rev. E 58, 69–73.

Vickers, G.T., Cannings, C., 1987. On the definition of an evolutionarily stable strategy. J. Theor. Biol. 129, 349–353.

Williams, C.G., 1966. Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thoughts. Princeton University Press, Princeton, NJ.

Wilson, D.C., 1983. The group selection controversy: history and present status. Ann. Rev. Ecol. Syst. 14, 159–187.

Wynne-Edwards, V.C., 1962. Animal Dispersion in Relation to Social Behaviour. Oliver & Boyd, Edinburgh.