The Evolution of Friendship

Limited partner choice, selective niceness and the emergence of cooperation

Daniel J. Hruschka*
dhrusch@sph.emory.edu

&

Joseph Henrich
jhenric@emory.edu

Department of Anthropology
Emory University
Atlanta, GA 30322

*Corresponding Author: Department of Anthropology, Emory University, 1557 Dickey Drive, Atlanta, GA 30322. Tel: (404) 636-0144. Fax: (404) 727-2860. Email: dhrusch@sph.emory.edu.

Manuscript Information: Figures 4, Tables 2

Word Count: Abstract 207 words, Body of text 6,786

Abbreviations: B/C, benefit-to-cost of cooperating; IPD, Iterated Prisoner’s Dilemma; TFT, Tit-for-Tat.

Revision: 12/15/04
ABSTRACT

The evolution of cooperation is a central problem in biology and the social sciences. While theoretical work using the iterated prisoner’s dilemma (IPD) has shown that cooperation among non-kin can be sustained among reciprocal strategies (i.e. Tit-for-Tat) these results are sensitive to errors in strategy execution, to cyclical invasions, and to the existing ecology of strategies. Moreover, at least in its application to humans, the IPD unrealistically assumes that an individual’s probability of playing the PD game with other strategies is independent of the decisions made by these strategies in the past. Here, we remove the independence assumption on partner choice by studying a more plausible cooperative dilemma in which players can preferentially interact with a limited set of known partners and deploy longer term accounting strategies that can counteract the effects of random errors. We show that cooperative strategies readily emerge and persist in a range of noisy environments, with successful cooperative strategies (henceforth, cliquers) maintaining medium-term memories for partners and low thresholds for acceptable cooperation (i.e. forgiveness). The strategy’s success relies on its cliquishness—a propensity to defect with new partners if it already has an adequate number of partners. Notably, this combination of medium-term accounting, forgiveness, and cliquishness fit with empirical studies of friendship among humans.

Key words: Iterated Prisoner’s Dilemma, Partner Choice, Friendship, Altruism, Cooperation.
The emergence of cooperation among unrelated individuals poses a continuing challenge to theories in biology and economics (Hammerstein, 2003). Much research has focused on how natural selection might favor strategies that cooperate in the two-person Prisoner’s Dilemma (PD). In the two-person PD, two individuals have the opportunity to exchange some favor but face the possibility of exploitation by their partner. When the individual benefit of cooperating (Let’s say B = 200) is greater than the cost (assume C = 100), then the overall payoff when both cooperate (2*(B-C) = 200) is greater than if neither does (zero). An evolutionary challenge arises because there is always a possibility that one partner will not cooperate, thereby taking the benefit (B) and leaving the partner with less (-C) then if he or she hadn’t given anything (Again, zero). In the purely one-shot game among non-relatives the only evolutionarily stable solution is to always defect.

Trivers (1971) suggested that individuals might achieve cooperation if they have the opportunity to cooperate over a sequence of moves. This idea was first formalized and modeled by Axelrod and Hamilton (Axelrod, 1984; Axelrod and Hamilton, 1981). From this early work, the iterated prisoner’s dilemma (IPD) emerged as the canonical model for studying the evolution of cooperation via reciprocity in repeated interaction. To model long-term relationships, the evolutionary IPD randomly pairs individuals and then forces them to repeatedly play a PD, where the probability of play continuing to the next interaction (after the first) is fixed. In such a setup, reciprocally cooperative strategies can resist invasion by pure defectors—if initially common—and can often maintain high levels of cooperation. This early work suggests that successful reciprocating strategies should be NICE (always cooperate initially) and PROVOKABLE (withdraw cooperation at the first sign of defection), qualities associated with the oft-cited strategy of tit-for-tat (TFT) (Bendor and Swistak, 1997; Gotts et al., 2003).

However, later work revealed both theoretical and empirical challenges to the early success of reciprocity-based strategies—especially TFT—in the IPD. First, various forms of errors or noise can devastate reciprocal cooperation. Both errors of implementation, which occur when individuals mistakenly apply their strategies (e.g., defect when they meant to cooperate), and errors of perception, which occur when individuals misperceive their partner’s actions (e.g., believe their partner defected when he actually cooperated), reduce the ability of reciprocal strategies to maintain cooperation (Bendor et al., 1996; Boerlijst et
al., 1997). Second, the success of particular strategies depends critically on the distribution or ‘ecology’ of other strategies in the population which may change through mutation and selection (Boyd and Lorberbaum, 1987; Lorberbaum, 1994; Lorberbaum et al., 2002). And finally, there remains the long-recognized problem that, reciprocal strategies generally cannot spread when rare without some form of ordered spatial or interactional distribution (Axelrod, 1984).

Particularly in the human case, a number of empirical concerns accompany these theoretical shortcomings. First, by contrast with the canonical IPD, individuals in many real situations are often neither locked into a sequence of interactions with unwanted partners nor forced to randomly switch to a new partner after some number of iterations. Numerous examples from humans, and other animals, show that individuals often leave relationships with one partner to find others with whom they might have better payoffs (Kelley and Thibaut, 1978; Redouan and Noe, 2003). Analytical IPD models typically assume that strategies meet with a probability equal to their current representation in the population. This would not be the case, however, if partners who achieve successful cooperation tend to stick together longer than partners who don’t. As a corollary, individuals who manage to achieve successful long-term cooperation are also likely to live longer due to the fitness benefits of long-term cooperation. Thus, unlike the IPD where all strategy-pairings have the identical expected number of future interactions, the shadow of the future should be longer—at least in some situations—for strategy-pairings that achieve successful cooperation. Second, while interactions among strangers seem empirically consistent with a TFT-like strategies (maintain close accounting and remain highly provokable), interactions within long-term human relationships, such as friendships, do not. Empirical evidence suggests that friends do not closely track particular exchanges, and are not easily provoked into defection (Silk, 2003). Moreover, humans are more willing to engage in costly cooperation with trusted partners than with strangers, and thus are not NICE in the way that successful strategies often are in the IPD (Markovits et al., 2003).

An alternative formulation for the evolution of pairwise cooperation—the explicit partner choice model—allows us to address several of the theoretical challenges described above for the IPD. It also yields results that are substantially more consistent with existing empirical work on humans than those that have emerged from the IPD. In explicit partner
choice models, players are able to influence their pairings (and thereby the pairing of others) using past experience and are not stuck in a sequence of interactions with particular individuals. This means the interactions and pairings of some individuals depend on the pairing of other individuals. Existing work on such partner-choice models, which has emerged in fragments from a variety of fields (Batali and Kitcher, 1995; de Vos et al., 2001; Hayashi and Yamagishi, 1998; Sherratt and Roberts, 1998; Skyrms and Pemantle, 2000; Stanley et al., 1994), suggests that they have a number of advantages over the canonical formulation. First, when partner choice is permitted, strategies can often rapidly achieve high levels of cooperation, even when the cooperative strategies are not initially common (Ashlock et al., 1996). Second, partner choice can resolve the problem of cyclical invasions in a changing ecology of strategies (de Vos et al., 2001; Sherratt and Roberts, 1998). Third, the basic setup that allows individuals to select and/or reject their partners has an intuitive plausibility for humans that the canonical ‘lock-in’ models of reciprocal altruism lack. Consistent with this approach, non-evolutionary tournament simulations suggest that partner choice algorithms may be more important in facilitating cooperation than are decisions to cooperate once partners have been paired (Yamagishi et al., 1994).

Despite the promise of the explicit partner choice approach, existing work has some shortcomings. First, previous research has ignored the potential effects of errors, which have been shown to be so important in the canonical model (Bendor et al., 1996). Second, and more importantly, previous approaches assume that players maintain complexly integrated memories of past interactions with every other individual in a group, making the models implausible for populations of more than 5-10 individuals (Dunbar, 1998; Milardo, 1992). Finally, the existing simulations have been uniformly limited to fewer than 60 players (often using only 20), restricting the generalizability of findings in more realistic populations and raising concerns about the effects of random drift.

In light of this, our simulation model synthesizes and extends existing work on partner choice models. We have synthesized prior strategies by capturing a substantial portion of the strategy-space with 4 evolving parameters (1 dichotomous and 3 continuous) and by permitting mutation to change these parameters. Our model addresses the above-discussed shortcomings in six ways. First, we have included two types of error, one in which individuals mistakenly defect when they mean to cooperate and the other in which
they are unable to meet with a preferred partner. Second, we have restricted individuals’ memories to one number for K different partners. K is set to three for our work below, as this is cognitively plausible for many primates, and is in a range observed for the number of close, supportive ties observed in humans (Dunbar, 1998; Milardo, 1992; Sugawara, 1984). Third, as a natural consequence of our partner-choice setup, the shadow of the future is longer for pairings that achieve successful cooperation than it is for pairs of defectors that are forced to interact—because greater fitness leads to longer lives and more reproductive opportunity, on-average. Fourth, we consider the dynamics of strategies in much larger populations than those considered in past simulations (N=100 to 2000), which more accurately reflects common group sizes among humans (Dunbar, 1998). Fifth, our model is a fully evolutionary model. We have avoided non-evolutionary pairwise tournaments and included mutation on all of our strategy parameters in order to avoid the restricted ecologies that have clouded prior work. Finally, the model permits us to examine prior assumptions about relationship decay in the IPD environment and show the assumption of geometric decay is only valid in the long-run and not during initial interactions.

THE MODEL AND SIMULATION

To explore the evolution of cooperation, we constructed a computer simulation in which individuals are paired according to their mutual preferences to play a one-round PD game. Payoffs from each round contribute to an individual’s fitness. After each round, players update their preferences based on these payoffs, and these preferences are used in matching players for the next round of interaction.

We consider a population of fixed size (N=100, 400, 800, or 2000) in which each individual maintains a strategy that specifies when to cooperate in the PD and how to calculate partner preferences. In each round, individuals apply this strategy to a PD game with a fixed payoff matrix. We explored PD payoff matrices with benefit-to-cost ratios (B/C) varying from 1 to 16. Table 1 represents two examples of these payoff matrices, one for a high ratio (B/C = 8) and one with a lower ratio (B/C = 2) when C = 100. At the beginning of each game, players are assumed to acquire baseline fitness equal to the cost of cooperating in the PD game.

(Insert Table 1)
**Memory, Partner Preferences and Matching**

For a limited number of potential partners, individuals can remember a single number that summarizes past interactions and represents the strength of preference. As noted, this number is updated after each round of play. Here, we allow memory for up to three partners, with memory being a real number in the interval $[0,1)$ rounded to three decimal places; this restricts the information available in each memory slot. Each strategy determines how this index is recalculated and when to add and remove old partners from its list.

Partner matching works as follows: Before each round, individuals deploy partner preferences, which are based on summaries of past interactions (described below). Matching occurs in two stages: (I) the preference stage and (II) the random assignment stage. In the preference stage, an individual (referred to as ego) is chosen randomly, and the matching algorithm:

1) Considers ego’s most preferred partner (of the three that are potentially stored in ego’s memory). If that partner also prefers ego, and that partner is not yet paired, then ego and the partner are paired. If no partner exists in ego’s memory, then ego is moved to stage II (placed in the pool for random pairing).

2) If the first preferred partner is not available then the algorithm considers the next preference, and so on.

3) If none of the K preferred partners are available, then ego moves to stage II.

4) Another unpaired individual is randomly chosen and the matching routine repeats, starting at step 1.

After all individuals preferences have been considered, stage II randomly pairs the remaining individuals.

**Life, Death and Reproduction**

Unlike most implementations of the evolutionary IPD, fitness-biased reproduction occurs after each play of a one-shot game. Fitness for an individual is based on payoffs from the most recent game plus some carry-over from past games:

$$F_i = \pi_i + \alpha F_{(i-1)i}$$
Here, $F_{ij}$ is the fitness of individual $j$ after round $i$. The payoff, $\pi_{ij}$, is the payoff that individual $j$ received from play in his $i$th game. The quantity $\alpha F_{(i-1)j}$ in the fitness carried over from the previous round, so $\alpha$ is the fraction of fitness not used up between rounds. In this paper, we present results for $\alpha = 0.6$, although basic results do not differ for $\alpha = 0$ (no carryover) or 0.30.

This form of fitness accounting captures two key aspects of fitness. First, it creates an exponential decay that prevents individuals from accumulating unrealistically large fitness levels relative to others. This is consistent with the observation that there are biological and social upper limits on one's ability to "store" the bases for fitness: food storage, rank achievement, wife accumulation, realistic fat storage. Our form sets a maximum limit, $\max(F_j)$, on fitness:

$$\max(F_j) = \max(\pi_j)(\frac{1}{1-\alpha})$$

Second, as one approaches this ceiling, there is generally decreasing marginal capacity for the bases of fitness (e.g., fat stores) to produce survival and reproduction. This fitness formulation guarantees a monotonically decreasing slope in the curves linking (1) accumulated fitness and survival and (2) accumulated fitness and reproduction.

Natural selection occurs after each round with the elimination of one individual based on a fitness-weighted probability. An individual’s probability of elimination ($Pd_j$) is

$$Pd_{ij} = \frac{G_{ij}}{\sum G_i}, \text{ where}$$

$$G_{ij} = 1 - \frac{(F_{ij} - \min(F_i))}{(\max(F_i) - \min(F_i))}$$

Upon selection, knowledge of this individual is also erased from the memories of other individuals.

Similarly, on the reproduction side, one of the remaining individuals is selected to reproduce and fill the empty slot based on a fitness-weighted probability ($Pr$):

$$Pr_j = \frac{F_{ij}}{\sum F_i}$$

A generation is defined as the number of rounds it takes to reproduce the number of organisms found in the population. When $N = 100$, there are 100 rounds per generation.4
Errors of Execution and Matching

We examine the effects of two kinds of errors. Individuals in our model execute their intended behavior (based on their strategy) with probability \(1 - e_s\), and execute the opposite behavior with probability \(e_s\). Thus, if an individual plans to cooperate with a preferred partner, she will ‘accidentally’ defect a fraction \(e_s\) of the time. We explore these strategy errors from \(e_s = 0.00\) to \(0.20\). We also consider errors in partner choice or matching errors. Specifically, if one has identified a mutually preferring partner in the matching round, there is a probability \(e_m\) (ranging from \(0.00\) to \(0.20\) depending on the simulation) that one will not successfully pair with that player.

Parameterizing the Strategies

Four variables parameterize all possible strategies (that we studied):

1) **Always Defect, D**: A dichotomous parameter indicating whether a strategy unconditionally defects (\(D = 1\)) or follows an algorithm determined by the next three variables (\(D = 0\)).

2) **Past weighting, \(\delta\)** (\(0 \leq \delta \leq 1\)): The relative weight of past experience in determining partner preferences. If \(\delta\) is near one, past preference is heavily weighted relative to the present payoff. If \(\delta = 0\), only the payoff from the most recent round matters, as is the case for TFT.

3) **Threshold of acceptable cooperation** (\(0 \leq \tau \leq 1\)): The threshold used to judge the acceptability of a partner’s past cooperation. If a player’s preference for a partner drops below \(\tau\), the partner is removed from the player’s memory.

4) **Cliquishness** \(\chi\) (\(0 \leq \chi \leq 1\)): The probability that a player using this strategy will defect with a new partner given that she already maintains \(K\) acceptable partners.

A strategy calculates future preferences for a specific partner by re-adjusting past preferences with outcomes from the most recent round:

\[
p_{jk} = \delta \cdot p_{i-1,jk} + (1 - \delta) \cdot M_{ijk}
\]

Here, \(p_{jk}\) is the preference that player \(j\) has for player \(k\) on the \(i\)th round (\(p_{0jk} = 0\)). \(M_{ijk}\) is the payoff from the \(i\)th round between the two players (if they paired) divided by the maximum payoff in the game. If no pairing occurred in a particular round between \(j\) and \(k\), \(p\) is not updated. This accounting method has several useful properties:
1. It is simple and requires keeping only one number in long-term memory.
2. It integrates past and present experience.
3. If $\delta$ is non-zero, it weights the long-term expected value of $M_{jk}$ by the number of times the two players have interacted together ($m$). Specifically, under conditions where the payoffs from a partner are temporally homogenous, $p_{mjk}$ is equal to $E(M_{jk})(1 - \delta^m)$, the product of the expected value of past payoffs $E(M_{jk})$ and an index of the duration of the relationship ($1 - \delta^m$).

This accounting scheme is similar to the weighted average of past interactions proposed by Bendor (Bendor, 1987). Unlike Bendor’s approach, however, the strategies in this paper do not base their actions on this index. Rather, they use the index to choose partners for play. Their subsequent decision to cooperate or defect depends only on $D$ and $\chi$. If they are matched with one of their preferred partners (and $D = 0$), they attempt to cooperate (and succeed with probability $1-e_q$). If ego is not able to pair with any of its preferred strategies, then its choice to cooperate is contingent on whether it already has $K = 3$ preferred partners. If it does, then it defects with its randomly assigned partner with probability $\chi$, waiting until it can cooperate with a player it already knows. However, if the strategy still has free memory available, then it always cooperates (unless $D = 1$).

If ego and a new partner cooperate, ego may decide to include the partner in its memory. This occurs if the preference value for that player after the mutual cooperation is greater than that of any of ego’s partners. If so, ego replaces the partner having the lowest preference score with the new partner’s information.

Partners whose value $p_i$ drops below ego’s threshold $\tau$ are removed from ego’s memory, while those with $p_i > \tau$ are ranked accordingly—keep in mind that players maintain a maximum of $K = 3$ preference values in mind at any one time.

**Mutation**

When a strategy is selected for reproduction, its four parameters are reproduced with mutational change. The offspring’s values for $\delta$, $\tau$, $\chi$ are mutated from the parent’s value based on a normal distribution with mean zero and standard deviation, $\sigma$ (typically set to 0.05), and the value for $D$ is reversed with a probability of $p_{switch}$ (typically set to 0.05)\(^5\). In cases where mutations produce a value outside defined conditions (i.e. $\tau < 0$) the value is set to the nearest boundary condition (i.e. $\tau = 0$).
Different specifications of the four variables (δ, τ, χ, D) correspond to the commonly known strategies from the canonical IPD—see Table 2. While Defectors are exactly the same as the analogous strategies used in canonical models, pure cooperator and TFT strategies require some comment. While pure cooperators here always intend to cooperate, they can still exercise partner preferences, and thus some cooperators are more fit than others. TFTers have partner preferences also, but those preferences are based entirely on what happened the last time those two strategies met. Also, TFTers are always NICE (initially cooperative) to individuals not contained in their K memory slots. Table 2 includes a novel strategy, clique, which we derived based on empirical studies of human friendships.

Cliquer strategies have a number of characteristics worth noting, some of which contrast with TFT.

1) A high δ means that Cliquers place value on the duration of the relationship. They prefer partners with whom they have developed a long series of beneficial interactions.

2) A high δ combined with a low τ makes them insensitive to short-term aberrations, thus making both cooperation and an individual’s preferences robust against any single defection (or even a few defections) after a long, cooperative relationship. However, cliquers remain sensitive to a series of defections—as do humans (Silk 2003).

3) Conversely, these same preferences make individuals more sensitive to a partner’s behavior early in the relationship—as in humans.

4) Unlike TFT, Cliquers are not generally NICE. Cliquers cooperate with preferred partners, but cooperate with novel partners only if their memory slots are not full. Faced with a new partner when it has enough acceptable partners, it defects.

We will refer to the broad class of strategies where D = 0 (Cliquer, Cooperator, and TFT) as cooperative strategies (CS) in contrast to the specific sub-class of pure cooperators (χ = 0 and D = 0).
**Analytical Strategy**

In using simulations to examine the emergence and persistence of cooperation, we focus on how key aspects of the environment (i.e. B/C, e_s, e_m, N) influence the: (1) invasion and long-term persistence of cooperative strategies (D = 0) in a population of defectors, and (2) long-term values of the three continuous strategy parameters (δ, τ, χ). For the first purpose, we begin simulations with D = 1 for all strategies. When a defector mutates into a CS, the three strategy parameters, δ, τ, and χ were randomly assigned (uniform distribution) to values in the interval [0,1).

For N = 100 and N = 400, we examined the sensitivity of strategy parameters (D, δ, τ, χ) to changes in environmental conditions (B/C ~ 1-16 & errors ~ 0.00-0.20). For each scenario, we ran 5 trials of 200 generations to examine the consistency of results. To explore the effect of population size (and drift) on the simulations we conducted single runs at N=800 and N=2000 for 100 generations in four environmental conditions (B/C = 4, 16; e_s= 0.05, 0.15, e_m= 0.05).

We determined the success of cooperative strategies (D = 0) by the extent to which their population prevalence deviated from that expected under neutral selection. Specifically, at reproduction, D has an equal probability of switching from 0 to 1 as it does from 1 to 0. Thus, if the two variants of D (0 and 1) experience equal selective pressures, then we would expect values of this parameter to reach equal proportions in the population. However, if D is consistently biased towards 0, then we should conclude that cooperative strategies are being selected for. In most simulations, the proportion of cooperative strategies quickly settles into a confined area of the parameter space. Consequently, we used the mean proportion of cooperative strategies (D = 0) from between 100 to 200 generations (200 generations = 200*N rounds) to assess the selective advantage of cooperative strategies versus defector strategies. The same rationale was applied to population means for the other parameters (δ, τ, χ).

**SIMULATION RESULTS**

Across a wide range of the environments, CS—with different values of the evolving parameters—emerge and sustain mean proportions in the total population well above 50% (the mean of null model simulations of neutral selection) with much smaller temporal
variation than observed in the null model. Figure 1A illustrates the temporal dynamics of a
typical run, while Figure 1B shows that same run without selection. The curve marked D =
0 gives the proportion of CS in the current population across 200 generations (200*400 =
80,000 rounds). As expected, parameters in the model without selection all exhibit a great
deal of temporal variation with means around 0.5.

The prevalence of cooperative strategies (CS) is sensitive to the benefit-to-cost
ratio, the population size, and strategy error, but not matching error. Increasing population
size and/or B/C ratios favors CS, while increasing strategy errors reduce the prevalence of
CS.

(Insert Figure 1)

Figure 2A plots the means and standard deviations (as error bars) in the proportion
of cooperative strategies across benefit to cost ratios for N = 100 and 400 (e_s = 0.05). Two
points are worthy of note. First, there is a precipitous decline in the frequency of
cooperaive strategies between B/C = 4 and B/C = 1, which is consistent with the benefi-
to-cost thresholds derived from analytical models. Second, increasing the population size
from 100 to 400 favors cooperation and reduces the standard deviation in the mean.

(Insert Figure 2)

As expected in Figure 2B, increasing the strategy error (e_s) both decreases the mean
proportion of cooperative strategies in the population and increases the standard deviation
in the mean. Again, the larger population resists the effects of increasing error more
effectively than the smaller population. Cooperative strategies, with their three evolving
parameters, are able to withstand quite high (e_s) rates.

Figures 2A and 2B both indicate that larger population sizes facilitate the success of
cooperative strategies (CS) against defectors. The prevalence of CS in larger populations
(N=800 and N=2000) is nearly identical to that in N=400, indicating that the effect of
population size is most pronounced when moving from N =100 to N=400 and decreases
rapidly toward zero beyond N=400. This pattern of effects suggests increasing N reduces
the influence of random drift, which acts against CS in these simulations.

Errors in matching with a preferred partner had no effect on cooperative invasions
(for e_m = 0.00 to 0.20). This is probably due to the possibility that an individual could turn
to another of its three potential partners if a specific match failed. With 3 partners and e_m =
0.1, the chances of failing to match with any of them and ending up in the random pairing pool is one in 1000. We suspect that at some higher value of $e_m$ the buffering effect of three partners will not be enough to counteract matching errors and the prevalence of CS will reduce with increasing rates of matching error.

The long-run continuous parameters ($\delta$, $\tau$, $\chi$) for cooperative strategies were each distributed, usually rather tightly, around one central tendency within a population. To ascertain the effect of selection on the variance in these parameter distributions, we compared the standard deviations, calculated by averaging over the last 100 generations of a single run, for each parameter under selection against those derived from a neutral model (the same simulation with selection turned off). Our findings show that selection substantially reduced the standard deviation of the parameter distributions. For example, the mean within-population standard deviations for $\delta$, $\tau$, and $\chi$ over 100 generations were 0.17, 0.08, and 0.15 respectively (when $N = 400$, $e_s$ & $e_m = 0.05$, and $B/C = 4$), while under neutral selection the mean standard deviations were nearly twice as large taking values of 0.30, 0.28, 0.31, respectively. This suggests that the three continuous strategy parameters were under considerable selective pressure, and indicates that the parameters of CS can be accurately summarized using their population means.

The specific parameters of successful cooperative strategies depended on the environmental details:

*Cliquishness* ($\chi$): Figure 3A shows that with increasing $B/C$ ratios, the long-term mean values for cliquishness, $\chi$, decrease, while temporal variation in a population’s mean cliquishness increases. This suggests that at low $B/C$ ratios, more cliquish strategies succeed. However, as $B/C$ ratios increase, the benefits of cliquishness are outweighed by the windfalls to be captured by unconditionally cooperating with new partners. Neither errors ($e_s$ or $e_m$) nor population size ($N$) had substantial impacts on cliquishness.

(Insert Figure 3)

*Past weighting* ($\delta$): Figures 3A and 3B shows that the weight of past experience, $\delta$, increases as both $B/C$ and $e_s$ increase (independently). The increase is slight for $B/C$, rising from 0.51 at $B/C = 2$ to 0.60 at $B/C = 16$. For strategy error, there is a dramatic increase from $e_s = 0$ to $e_s = 0.01$, with a slower rise from $e_s = 0.01$ to $e_s = 0.20$. This substantial increase is accompanied by a tightening of the temporal variation in population means,
suggesting that strategies which place more weight on past interactions succeed in environments with higher rates of execution errors. Conversely, in environments characterized by no execution error, strategies with low $\delta$ (i.e. similar to TFT) survive.

Population size also influences the long-term mean of $\delta$, with increasing N leading to greater values of $\delta$. For example, in the base situation ($e_s = 0.05$, $e_m = 0.05$, B/C = 4) there is a steady increase from 0.44 to 0.64 in mean values of $\delta$ between N = 100 and N = 2000.9

Because $\delta$ values are in the vicinity of 0.50 (e.g. values also favored by neutral selection) we compared the standard deviations seen in Figure 3A and 3B with those from our neutral model simulations, and found that these parameters are clearly under strong selection. Not only are most $\delta$ values significantly different from 0.5, but the within-run standard deviation for $\delta$ under neutral selection (std dev = 0.07) was two to three time greater than that for $\delta$ under selection in any environmental condition (std dev = 0.02-0.03).

Threshold of acceptable cooperation ($\tau$): Figure 3A shows that with increasing values of strategy error the long-term mean values of $\tau$ decreased, with the most dramatic decline occurring between $e_s = 0$ and $e_s = 0.01$. This suggests that at larger values of strategy error, more forgiving strategies succeed. Neither matching error ($e_m$) nor benefit-to-cost ratio had an impact on long-term means of $\tau$.

With increasing population size, long-term means of $\tau$ also decreased. For example, in the base situation ($e_s = 0.05$, $e_m = 0.05$, B/C = 4) there is a steady decrease from 0.13 to 0.08 in mean values of $\delta$ between N = 100 and N = 800, which is statistically significant considering the small between-run standard deviations (std dev = 0.007-0.011) within environmental conditions. This effect is flat between N=800 and N=2000.

The relatively small within- and between-run variation under specific environmental conditions in population means of $\tau$ (within-run std dev = 0.01-0.06, between-run std dev = 0.001-0.03) suggests that this parameter was under strong selection. The primary advantage of high $\tau$ values is the ability to leave partners who start by cooperating but then begin to consistently defect after some period of cooperation. In this model, such latent defectors are not possible as they would require more memory slots per partner (see Discussion for future directions). On the other hand, there is a strong selective force
pushing τ towards 0—the need to forgive long-term cooperative partners who make a consecutive string of accidental defections.

**Distribution of relationship durations**

In contrast to many evolutionary IPD models, our partner selection approach does not fix the probability (ω) that a relationship between two partners will continue to the next round (Axelrod and Hamilton, 1981; Boyd and Richerson, 1992; Brown et al., 1982). Rather, mutually preferred interactions end ‘naturally’ as the result of changing partner preferences, the death of partners, and the end of the simulation. In addition, mutually preferred interactions may be temporarily interrupted by interactions with other partners. From these ‘naturally’ occurring relationships, it is possible to derive the actual distribution of relationship durations and to compare it to the geometric distribution assumed in most IPD models.

To illustrate, consider a situation where cliquers invades defectors (e_s = 0.05, e_m = 0.05, B/C = 2, N = 100). Over 200 generations (200*100 iterations), there were 49,038 relationships where two partners interacted for more than one round. Of these, 5298 ended after two interactions, while another 3780 ended after three interactions. From these dissolutions we can calculate the probability of ending a relationship after a given number of interactions. Figure 4 shows the evolution of ω during the first 40 interactions between two individuals. Although it settles around 0.96 after the first 10 interactions, it shows a divergence from the assumption of uniform ω in early interactions, with higher rates of dissolution at early stages of a relationship.

(Insert figure 4)

**DISCUSSION**

Our findings build on past work on the evolution of dyadic cooperation by studying an explicit partner choice model that (1) permits individuals to play the PD game with preferred partners at a greater probability, (2) effectively parameterizes the strategy space (rather than limiting the process to a finite number of discrete strategies), (2) allows mutation on all evolving parameters permitting a full range of strategic ecologies, (3) includes two kinds of behavioral noise, (3) limits partner memory, and (4) considers the effects of population size and drift. Our work confirms the most general finding from the
canonical model: benefit to cost ratio show a steeply non-linear threshold effect on the emergence of cooperation.

However, in contrast to the canonical findings, we have demonstrated that certain kinds of cooperative strategies can invade when rare and remain stable even in the face of substantial behavioral noise. The difficulty of invasion that plagued early work appears to be due to the inability of individuals to increase their probability of playing the PD game with preferred partners (a restriction we feel is largely unrealistic for humans). In our model, rare cliquers can keep searching for other sufficiently cooperative partners until they find them, and once found, they can continue cooperating for as long as acceptable cooperation is maintained.

Successful cliquer strategies differ from TFT-like strategies in two ways. First, cliquers integrate past behavior into partner preferences, with about half of the weight on the most recent interaction and about half on prior interactions (higher values of $e$ produced greater weight on the past)—note, these strategies are not remembering more information about past interactions; they are merely weighting the construction of their preferences differently. Interestingly, weighting of past outcomes was considerably weaker in environments with no strategy noise. This is consistent with past findings that TFT-like strategies are most successful in environments where there is no threat of execution errors.

The second point of difference between cliquers and TFT is that cliquers are not generally cooperative (NICE) with new interactants. If their $K$ partner slots are filled, they defect on non-partners most of the time. Interestingly, the tendency to cooperate with non-partners depends on the benefit to cost ratio, with low $B/C$ ratios favoring less cooperation with non-partners. Thus, the prediction is that individuals will tend to be more cooperative with non-partners in higher-$B/C$-ratio environments, not wanting to miss out on the big benefits to cooperation. Overall, a strategy’s power to invade and remain stable in the face of noise and a wide-ranging ecology of strategies rests on its ability to cultivate and to maintain a small set of local relationships with known individuals, while largely ignoring (and perhaps exploiting) the rest of the population.

We think this work offers both an intuitively more plausible evolutionary model for our species than the IPD, and yields results consistent with empirical patterns observed in human friendships (Silk, 2003). Specifically, successful strategies eschews tit-for-tat
accounting by tracking a level of preference in a partner that is slowly built up or eroded based on repeated interactions (due to a high $\delta$ value). Although this mode of accounting is slow to break with an old friend, it is sensitive to early interactions, which makes it an effective antidote to full-time defectors. As just noted, the strategy is not unconditionally NICE and does not typically cooperate with non-partners (due to a high $\chi$ value).

Not only does the model generate psychological and behavioral patterns consistent with human friendship, it also creates realistic-looking friendship networks. Over the long-run these networks accrete new cliquers as a crystal adds new matter. The ability to supplant populations and resist invasion is based in part on the stable, yet flexible network of mutual partnerships that arise due to Cliquer’s preference strategy.

**Evolutionary avenues to cooperation**

Various forms of cooperation in humans likely arose via several selective processes, including kinship, reputation and culture-gene coevolution (Barclay, 2004; Dasgupta, 1988), which in one fashion or another solve the same underlying problem (Henrich, 2004). Thus, the success of the cliquer strategy is a specific case of the general observation that cooperative strategies must have a greater than random chance of interacting with other cooperative strategies. Past examples of modeling this non-random matching have used spatial patterning (Eshel et al., 2000; Sella and Lachmann, 2000; Watanabe and Yamagishi, 1999) and tribal or dialectical assortment (Nettle and Dunbar, 1997). Our approach is different from these because it relies on the processes and resulting psychology of focusing effort on long-term enduring cooperative relations with specific partners, rather than determining whether to cooperate with a randomly chosen individual with whom one may have several repeated interactions. We argue that this specification of partner choice is resistant to the roving defectors that might exploit spatially configured populations (Watanabe and Yamagishi, 1999) and mimics that can exploit tribal or dialectical matching (McElreath et al., 2003; Nettle and Dunbar, 1997)—because it is much harder to mimic a specific person.

**Concerns: Tradeoffs and Diabolical Cliquers**

Solving the dilemma of dyadic cooperation involves two interrelated trade-offs. First, an actor must decide to cooperate and face potential exploitation, or avoid cooperation altogether (Bendor 1993). However, the possibility of cultivating a relationship
introduces a new tradeoff—that between the increased certainties gained through building a few relationships and the potential opportunities lost by remaining confined to specific partners. In environments with heterogeneous payoffs, if one does decide to cooperate and to cultivate a relationship, it is necessary to monitor whether one is getting an adequate payoff from a current partner relative to other potential partners. This paper has focused on resolving the first trade-off by allowing individuals to test and cultivate relationships. However, as social exchange theories have suggested, it is also important to understand how individuals make guesses about the opportunities they lose by engaging in one relationship versus another (Kelley and Thibaut, 1978). In the future, it will be important to model how individuals survey their environment and make decisions to move to new relationships based on their estimations.

Although cliquers are successful in emerging and persisting against traditional strategies in these models, it is already possible to imagine alternative strategies that might be able to infiltrate a population of cliquers. For example, a strategy which starts off cooperating, develops a relationship, and then begins defecting at a certain probability may well supercede a population of cliquer strategies. However, under such circumstances, cliquers may persist or re-invade by adjusting their long-term memory weightings ($\delta$) and cooperation thresholds ($\tau$) to counterbalance these diabolic latent defectors. Such invading strategies may also be dealt with by incorporating a second system of cultural inheritance, which would allow cliquers to adapt their parameters ($\delta$, $\tau$) to the current local distribution and strategic tactics of these diabolical cliquers (Henrich and Henrich, in press). Future simulation work will examine the effects of such latent defectors and the effects of the availability of a second system of cultural inheritance.

**Does reciprocal altruism explains prosocial behavior in one-shot economic experiments?**

Findings from this model bear on a current debate at the interstices of evolutionary biology and behavioral economics. Results from behavioral economics showing prosocial behavior in anonymous one-shot experimental games, such as the Ultimatum Game, Prisoner’s Dilemma, and Public Goods Games, have been interpreted as inconsistent with the predictions of reciprocal altruism (derived from IPD models), and in need of an alternative evolutionary explanation (Fehr and Fischbacher, 2003). Critics of this have
countered by suggesting that reciprocal altruism can explain people’s behavior once one accounts for the fact that human social psychology evolved in smaller-scale societies with stable long-term membership. In these environments, some have reasoned (presumably from the results of the canonical IPD models) that individuals should have evolved to be \textit{NICE} (i.e., always cooperate in ‘round 1’ of a repeated game) to everyone, because anyone they encounter is likely to be a long-term future interactant. Carrying this evolved psychology to modern social environments and behavioral experiments, the argument suggests that people cooperate in one-shot games (are \textit{NICE}) because they are acting ‘as if’ they are in a repeat sequence of interactions—even though they are not (Johnson et al., 2003).

Testing the theoretical logic underlying this argument, our model—which is based on the logic of repeated interaction—predicts that people should \textit{not} typically cooperate with just anyone from their group (i.e., they should not generally be \textit{NICE} to anyone). If the above proposal were sound, \(\chi\) in our model would tend to stabilize near a value of zero. Instead, \(\chi\) remains well above 0.5, even for extremely high value of \(B/C\). In fact, we did not find any conditions favorable to the \(\chi = 0\) supposition. Given this, we tentatively submit that claim of \textit{niceness} is an artifact of the independence assumption of interactions in the IPD. And thus, the behavior in one-shot experiments cannot likely be explained by the evolutionary logic of repeated interaction.\textsuperscript{11}

\textit{Thinking about relationships}

Theory in the evolution of cooperation has been influenced by the notion that ‘altruism’, ‘trustworthiness’, ‘reciprocity’ and ‘reputation’ are properties of the individual. However, in daily life we frequently see individuals extend a helping hand to specific companions, while neglecting to aid others. Some individuals may have a bad reputation in general, but still be a reliable friend and cooperator with a select few. Our model begins to capture this by producing individuals that are ‘cooperative’ and ‘altruistic’ with some people, but not with most. In a population of cliquers, for example, no cliquer is more or less altruistic than the others, yet there is a great deal of heterogeneity in who is altruistic to whom. The possibility for altruistic interactions exists precisely because individuals can cultivate \textit{cooperative relationships}. This is not to say that there are not trait-like differences between individuals, but rather to shift the emphasis to the relationship-building aspects of
our species. Further exploration of how individuals cultivate cooperative local environments would be a useful addition to the current emphasis on individual traits.
References Cited


Figure Legends

Fig. 1. Evolution of strategy parameters (N=400, e_s & e_m = 0.05, B/C = 2). (A) Mean parameter values under selection: Mean $D = 0.81$ (SD = 0.05), $\delta = 0.51$ (SD = 0.03), $\tau = 0.10$ (SD = 0.01), $\chi = 0.85$ (SD = 0.02). (B) Under neutral selection mean parameter values approximate 0.5 (0.48-0.54) with larger standard deviation (SD = 0.06-0.09). SD is an index of temporal variation in population means, based on the variance in population mean values over time.

Fig. 2. Prevalence of cooperative strategies. (A) By B/C ratio ($e_s & e_m = 0.05$). (B) By strategy error ($B/C = 4, e_m = 0.05$). Data points represent 5 simulations of 200 generations, and means are taken from generations 100-200. Error bars represent standard deviations of the mean of population means taken from 5 runs. Values for N=800 and N=2000 were too close to N=400 to represent on the graph. For example, in (A), the values for N=800 at B/C = 2,4 & 8 were 0.84, 0.95 & 0.97 respectively.

Fig. 3. Mean strategy parameters ($\chi, \delta, \text{and } \tau$). (N=400). (A) By B/C ratio ($e_s & e_m = 0.05$). (B) By strategy error ($B/C = 4, e_m = 0.05$). Data points represent 5 simulations of 200 generations, and means are taken from generations 100-200. Error bars represent standard deviations of the mean of population means taken from 5 runs.

Fig. 4. Rate of relationship (tie) dissolution (w) during first 40 interactions between two partners. (N=100, $e_s = 0.05, e_m = 0.15, 200$ generations).
assumption of constant $\omega$

B/C = 8
R/C = 2
Table 1: Payoffs to partner A

<table>
<thead>
<tr>
<th></th>
<th>B/C = 2</th>
<th></th>
<th>B/C = 8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B cooperates</td>
<td>B defects</td>
<td>B cooperates</td>
<td>B defects</td>
</tr>
<tr>
<td>A cooperates</td>
<td>200</td>
<td>0</td>
<td>800</td>
<td>0</td>
</tr>
<tr>
<td>A defects</td>
<td>300</td>
<td>100</td>
<td>900</td>
<td>100</td>
</tr>
</tbody>
</table>

Table 2. Common strategies mapped onto parameters specifications

<table>
<thead>
<tr>
<th>Strategy</th>
<th>$\delta$</th>
<th>$\tau$</th>
<th>$\chi$</th>
<th>$D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Defector</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>1</td>
</tr>
<tr>
<td>Cooperator</td>
<td>any</td>
<td>any</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Tit-for-Tat</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cliquer</td>
<td>high</td>
<td>low</td>
<td>high</td>
<td>0</td>
</tr>
</tbody>
</table>
Footnotes

1 Based on the two partner’s choices, there are four possible payoffs: \( R = B - C \) if both cooperate, \( T = B \) if your partner cooperates but you defect, \( S = -C \) if you cooperate and your partner defects, and \( P = 0 \) if neither cooperates. Given that \( B/C \) is greater than 1, we know that \( T > R > P > S \), and \( R > 2*(T+S) \).

2 Another interpretation of the iterated IPD is that individuals play the PD with every other individual in the population for a specified number of rounds. In this sense, the IPD allows individuals a kind of partner choice by allowing them to cooperate with some individuals but not with others (if the payoff to mutual defection is equivalent to the payoff of non-interaction). This, however, forces each player to expend time in an interaction with each member of the population, rather than permitting players to focus their efforts on a few known, mutually cooperative partners. Conversely, in our model, strategies that develop cooperative relationships effectively remove themselves from the mix of strategies that defectors, for example, might encounter.

3 By using one population (of size N) we have conflated the effects of (a) drift with (b) the number of potential partners to which one has access. While we agree this is a theoretical possibility, our actual results alleviate these concerns, since the effects of varying N look remarkably like the expected drift effects.

4 While non-overlapping generations would allow a more direct comparison with the canonical IPD, our approach, because we are interested in the emergence of partner networks that outlast any given individual, demands overlapping generations.

5 We also examined simulations at lower values of \( \sigma \) and \( p_{\text{switch}} \) (0.01 and 0.005), and found that cooperative strategies achieved even higher proportions in the population (though after a much greater number of generations). For practical considerations we selected these higher mutation rates, but they do not substantially change our findings.

6 For \( N = 100 \) under the same conditions, \text{std dev} \( \delta = 0.20 \), \text{std dev} \( \tau = 0.10 \), \text{std dev} \( \chi = 0.19 \) under selection and \text{std dev} \( \delta = 0.26 \), \text{std dev} \( \tau = 0.24 \), \text{std dev} \( \chi = 0.24 \) under neutral selection.

7 In the course of our analysis we consider three types of parameter variance: 1) the variance of parameter values within a population, which we characterize by averaging the standard deviations of individual parameter values across the last 100 generations, 2) the temporal variance in population means over a series of generations, and 3) the variance between the overall run-level means. Tighter population variance indicates greater selection pressure on the parameter, tighter temporal variance indicates that a population’s parameter mean has reached a temporal point-equilibrium (with some error), and tighter between-run variance indicates that equilibria are similar across different runs.

8 Long-term means taken from generations 100-200 under the following condition: \( B/C = 1,2,3,4,8,16; \) strategy or matching error rates = 0.00, 0.01, 0.05, 0.10, 0.15, 0.20; \( N=100,400 \).

9 At \( N = 100, 400, 800, 2000 \), the long-term means of \( \delta \) were 0.44, 0.53, 0.59, and 0.64 respectively (with between run standard deviations between 0.02 and 0.05).

10 The ecology of possible strategies has been shown to drastically affect the success of particular strategies.
There are other empirical and theoretical problems with using reciprocal altruism to explain the prosocial behavior observed in one-shot games. (Fehr and Henrich, 2003)