

Modeling reciprocation and cooperation in primates: evidence for a punishing strategy

M. Keith Chen^{a,*}, Marc Hauser^b

^a*Department of Economics, Yale School of Management, 135 Prospect Street, P.O. Box 208200, New Haven, CT 06520, USA*

^b*Harvard Department of Psychology, 33 Kirkland Street, Cambridge, MA 02138, USA*

Received 22 June 2004; received in revised form 24 November 2004; accepted 13 December 2004

Available online 25 January 2005

Abstract

Experiments in which animals strategically interact with one another or search over some controlled domain are becoming common. While these experiments often promise to illuminate sophisticated animal behavior, the analyses brought to bear on these data are often quite coarse. For example, many papers simply tally the number of observations consistent with a behavioral theory. This analysis is simple, but ignores a potentially rich source of information by failing to take into account patterns and systematic variation among observations inconsistent with the theory. Using a new data set generated by cotton-top tamarin monkeys playing a repeated food-exchange game, we apply a maximum-likelihood estimation technique (more commonly used to study human economic behavior) which utilizes much more of the information in these data, and which uncovers unexpectedly sophisticated cooperative behavior from our subjects. Tamarin cooperation remains stable as long as both actors consistently cooperate, but requires at least two consecutive unexpected acts of cooperation to restart cooperation after it has collapsed, a strategy that resembles two-tits for a tat. We conclude by enumerating the benefits of a maximum-likelihood approach in experimental settings such as ours, and suggest other areas in which these techniques may be fruitful.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Behavioral strategies; Reciprocal altruism; Cooperation; Primates; Games

1. Introduction

Animals are often presented with situations where there are alternative strategies for solving a problem, whether they are engaged in finding a mate, fighting with a competitor or, in the experiment reported here, cooperating with an unrelated animal in a repeated interaction. One question of interest is not only what strategy the animal uses, but whether this strategy is stable, and whether or not it is actually the specific manifestation of a more general behavioral rule. In strategic settings we are also naturally interested in how

one individual's strategy interacts with another's; for example, whether they stabilize or destabilize a cooperative norm.

Traditionally, studies, which sought to test game-theoretic claims or to test the validity of complex search and foraging strategies did one of two things. Often, a simple experiment or field observation was designed to test a novel prediction of a particular theory. More generally, theories have been evaluated by comparing the proportion of observed behavior, which was consistent with either the theory or an ancillary prediction of it. For example, there has been a large literature on war of attrition models as applied to animal contests. This work often tested implications of the theory such as: the length of fights should be negatively correlated with differences in the size of combatants (see, for example, Parker and Thompson,

*Corresponding author. Tel.: +1 203 432 6049;
fax: +1 203 432 6974.

E-mail addresses: keith.chen@yale.edu (M.K. Chen),
mdhauser@wjh.harvard.edu (M. Hauser).

1980; Riechert and Hammerstein, 1983; Austad, 1983; Leimar et al., 1991). These analyses, however, fail to utilize all the information present in the data collected; in specific they ignore *systematic* variation in those observations inconsistent with a given theory. For instance, if a simple war of attrition models predicts that an animal of small enough size will never continue fighting past a point, it is informative if those instances when fighting drags on (in violation of the prediction,) differ systematically from instances when fighting stopped on time. Key of course, is that a model which at least partially explains these non-predicted violations must be available; latter we will discuss the steps we took to assure that a reasonably large set of models were applied in our analysis.

Here we propose to explore a new set of methodologies for analysing strategic behavior in animals. Borrowing from techniques typically used to study strategic human behavior, we use a maximum-likelihood estimation method to uncover game-theoretic patterns of behavior in the data (for an excellent discussion of the economic analysis of strategies, see Houser et al., 2004). Specifically, we fit a comprehensive suite of strategic behavior rules to observations of animals playing a cooperative food-exchange game. This creates a posterior-likelihood statistic, which has several desirable features. First, it exploits far more of the available patterns in these data, providing an objective mechanism for rejecting hypotheses that fail to explain systematic variation inconsistent with the theory. Secondly, a posterior-likelihood statistic can be easily aggregated across studies. This greatly simplifies the process of meta-analysis, and aids the theorist who wishes to judge behavioral theories on the basis of parsimony. We apply these techniques to games of cooperation between unrelated tamarins, taking advantage of prior studies of this species. Because of the nature of the experiments conducted and the focus on cooperation, these techniques yield new results that shed light on the nature of cooperation, its stability, and the possible underlying psychological mechanisms; the latter is particularly important as it sets up new hypotheses that can be tested with additional experimentation. We further show that the behavior of our subjects is remarkably similar to behavior of humans in similar settings, suggesting that the psychological mechanism by which humans sustain repeated cooperation can best be understood as evolved behavior. We begin by a brief review of our earlier experimental data, and then turn to our game-theoretic analyses.

2. Cotton-top tamarin cooperation

Trivers' theoretical paper on reciprocal altruism launched a cascade of studies on cooperation in wild

and captive animals (reviewed in Dugatkin, 1997). Though there is abundant evidence for cooperation among kin as well as for mutualistic interactions among non-kin, the evidence for reciprocity among non-kin is still slim and disputed (Stevens and Hauser, 2004; Hammerstein, 2003). Some have argued that the lack of decisive evidence is due to lack of proper experimental approaches or appropriate target species, while others argue that it is a reflection of cognitive constraints. Independently of the cause, we can nonetheless profit from those cases that appear (Milinski et al., 2002; de Waal, 1996) to provide some evidence for reciprocity to model the strategies underlying their behavior.

In the tamarin experiments, Hauser and colleagues placed two genetically unrelated adults in a situation in which each player had the opportunity, on alternating trials, to provide food to its opponent. The task involved pulling a tool that on a given trial brought food to the receiver but no food to the actor (see Fig. 1 in Hauser et al., 2003); definitionally, therefore, the act is altruistic as it entails a cost to the actor and a benefit to the recipient. To understand the conditions mediating food exchange, four experiments were conducted. Of these four, the first three were intended to explore the degree to which a tamarin's behavior in a repeated game was contingent on both the behavior and the perceived intentions of their opponent, and the fourth was intended to illuminate exactly what strategies tamarins play against each other in that same game.

The first experiment was designed to explore the possible effects of contingency and reputation. Several untrained individuals played consecutive games against either a trained unilateral altruist or a trained unilateral defector; these trained animals were both adult females and both genetically unrelated to the untrained animals, some of which were adult males and some adult females. The trained altruist always pulled the tool (pulling delivers food to its opponent, see Fig. 1 in Hauser et al., 2003), while the defector never pulled. Results showed that the untrained animals pulled significantly more for the altruist than for the defector and consistently maintained this pattern from the first session to the fourth; when playing the altruist individuals pulled on approximately 40% of all trials. These data show either that tamarin food giving depends upon prior, intentional acts of food giving by the opponent, or that tamarins are simply more likely to give food when they have been given to and are sated, independently of the particular details of the food-giving agent. To test between these alternatives, we ran experiments two and three.

In the second experiment, we reasoned that if simple reinforcement causes an increase in food exchange, then cooperation should change independently of how an individual receives food. In these experiments, two

untrained animals first played three sessions [24 trials each] of alternating, altruistic food exchange; on each trial, one player had the opportunity to give food to the other player without any immediate benefits; pilot studies revealed that under these conditions, individuals start out pulling about 40% of the time and then drop 5–10% before stabilizing. Following these three sessions, the pair then played a fourth session in which each pull resulted in a piece of food for the actor and a piece for the recipient; the fifth and final session involved a round of alternating altruistic food exchange. If cooperation is enhanced by simple reinforcement, then following session 4 in which each player is expected to pull on every trial due to selfish reasons alone, both players should pull more on session 5 than they did on session 3. That is, if reinforcement drives cooperation, we predicted a decrease in pulling from session 1 to 3, an increase to 100% in session 4, and a higher rate of pulling in session 5 as compared with session 3. In contrast, if cooperation depends on prior acts of altruistic food giving, then session 4 will have no impact, and the rate of pulling in session 5 will be equal to or lower than in session 3. Results argue against the reinforcement hypothesis, showing no significant change in session 5.

Experiment 3 provides further evidence against the reinforcement hypothesis. Two untrained actors played repeated games in which actor 1's pull resulted in one piece of food for self and three pieces for actor 2; actor 2's pull resulted in no food for self and two pieces for actor 1. Thus, if both actors pulled, each would receive three pieces of food after one round. The central problem, however, is how actor 2 interprets actor 1's pull. If actor 2 perceives actor 1's pull as cooperative (i.e. as motivated to give food), then actor 2 should often pull. In contrast, if actor 2 perceives actor 1's pull as selfish, with the three pieces of food delivered as a byproduct of actor 1's motivation to get one piece, then actor 2 should rarely pull. Results support this latter interpretation, with actor 1 pulling on almost every trial and actor 2 almost never pulling.

In a final experiment, we showed that the patterns of food giving appear to be contingent upon playing with a conspecific, as opposed to any reinforcing agent. One tamarin played the role of an active player, capable of altruistically giving food to an opponent. The second tamarin played the role of passive receiver. When the tool used to deliver food to the other actor was placed on the passive tamarin's side, the handle was flipped over toward the human experimenter who pushed the tool, delivering food to the active player. Consequently, this game mirrored the payoff structure of experiment 1 in which one individual obtains food on every other trial. The only difference between these two experiments was that in Experiment 1, a tamarin played the role of unilateral altruist, whereas in Experiment 4 a human

substituted for the tamarin and played the role of unilateral altruist. If tamarins disregard the kind of agent delivering food, attending to payoffs alone, then they should cooperate at the same level in Experiment 1 as in Experiment 4. Results reveal highly significant differences between these experiments, with the active tamarin pulling at a rate of approximately 10% of all trials, a rate that is almost identical to that observed when tamarins played against the unilateral defector. Together, these results indicate that tamarin food giving is based on prior, intentional acts of altruistic food giving by a conspecific; simple reinforcement is insufficient.

3. Strategic analyses

To illustrate how a maximum likelihood approach can be applied to strategic data, let us look at how subjects behave in our repeated food-sharing experiment, and how the data they generate can be used to test the relative fit of behavior rules to that behavior. Specifically, we will analyse all sessions from the repeated food-exchange game in which both actors were untrained, and for which pulling resulted in the other actor receiving a single piece of food. This type of analysis has begun to be extensively used in economics to study the behavioral rules used by human decision-making agents; see, for example, Houser et al. (2004), Camerer and Ho (1999), El-Gamal and Grether (1995) and Harless and Camerer (1994). Unlike the economic analysis of the above papers however, we will restrict our study to identifying a core strategy that is shared by all the tamarin actors; in many human studies (differentiated by richer action sets and thicker, more contingent actions,) behavioral heterogeneity is explicitly allowed and tested for.

For simplicity, let us look at two simple rules, a tit-for-tat type rule and a simple one-stage learning rule that is sometimes called "Pavlov". Given that our actors are only allowed to choose from two actions in each stage—pull or not pull—these two rules can be operationalized in our setting in the following ways:

- (1) Tit for Tat (hereafter written TFT): If the first chance to act is yours, then pull the tool (giving your partner food); thereafter copy what your partner did the previous round.
- (2) One-stage learning, or "Pavlov": If the first chance to act is yours, then pull the tool (giving your partner food); if what you did two rounds ago resulted in your partner pulling for you on the last round then repeat your previous action, otherwise switch.

Since any rule is most likely a simplification of an actor's true (possibly stochastic) decision process, with

anything but the shortest string of observations no rule would perfectly predict an actor’s behavior; to fit any particular rule to the data we must allow for each rule to be followed with error. The simplest way to do this and the method that we adopt in this paper, is to fit each rule to the data by taking the history of play at each stage as given, and asking what each behavior rule predicts will happen in the next trial, while allowing for an error rate in this prediction.

Suppose, for example, that we observed a string of data produced by two actors, with zeros indicating no-pulls and ones indicating pulls. One actor is associated with the even trials, the other actor with the odd trials. Take the hypothetical data below in Table 1. TFT predicts pulls in trials 1–4, 8 and 9, and non-pulls for every other trial. Our Pavlov rule would predict pulls in trials 1–4, 6, 7 and 9, and non-pulls for every other trial.

When TFT predicted the actor would pull (in trials 1–4, 8 and 9) the actor pulled in four of six trials. In contrast, when TFT predicted the actor would not pull (in trials 5–7 and 10) it was correct in only two of four trials. The same exercise for our simple learning rule has a success rate of four out of seven when it predicted pulling, and one out of three when it predicted non-pulls. We can therefore think of TFT as having a hit rate of two-thirds when it predicts pulling, and of one-half when it predicts non-pulling. Likewise, our Pavlov rule has a hit rate of four-sevenths when it predicts pulling and of one-third when it predicts non-pulling. We can now ask the question: how well does either of our models fit the data, given the observed hit rate for each action? This is equivalent to asking: using the error rates as free parameters, how well does either model fit the data?

We address this question by assuming that the error rates for each model are independent and stationary across time, allowing us to aggregate into a posterior probability measure of how well each model predicts the data. If there are two potential actions and when the model predicts action 1 (A_1 times,) the data of N trials conforms p_1 percent of the time (and similarly for action 2 matches the model p_2 percent of A_2 times) then the posterior likelihood of the model

generating the data is

$$p_1^{A_1 p_1} (1 - p_1)^{A_1 (1 - p_1)} p_2^{A_2 p_2} (1 - p_2)^{A_2 (1 - p_2)}.$$

Under this framework, tit-for-tat scores

$$(2/3)^4 (1/3)^2 (1/2)^2 (1/2)^2 = 1/729$$

or approximately 0.00137.

Likewise, our simple learning rule scores

$$(4/7)^4 (3/7)^3 (1/3)^1 (2/3)^2 = 1024/823543$$

or approximately 0.00124.

These values can be thought of as the posterior probability that each model generated the data given that each model’s *observed* errors are indicative of as each model’s “*true*” error rate.

Perhaps the most natural way to compare these two models then is to ask: what is the relative likelihood of either model generating the observed data? A natural first step would be to compare each model’s fit to the null hypothesis that the data were simply generated by a stochastic, independent and identically distributed pulling process (from here on out called the IID model). That is, that each actor’s actions were a weighted random draw that is independent of both time and history. The best such a process could fit the data is by matching a pull rate of six-tenths. This generates a posterior probability of

$$(6/10)^6 (4/10)^4 = 11664/9765625$$

or approximately 0.00119.

Hence, both TFT and the one-stage learning beat out the IID model in terms of posterior probability. Comparing these three posterior likelihoods: TFT fits best, with a 36% likelihood; then one-stage learning fits with a 33% likelihood leaving 31% for the IID model. In other words, if you were told that only two models or an IID null could be responsible for generating the observed short string of data, you could be 36% confident that it was TFT. Consequently, our hypothetical data are slightly more consistent with TFT than they are with the learning rule, and both seem to have some predictive power.

Table 1
Hypothetical data and predictions for two models

Hypothetical data and model predictions		Trial number									
		1	2	3	4	5	6	7	8	9	10
Hypothetical data for	Actor 1:	1		1		0		1		0	
	Actor 2:		1		0		0		1		1
Tit-for-tat prediction for	Actor 1:	1		1		0		0		1	
	Actor 2:		1		1		0		1		0
Learning rule (Pavlov) prediction for	Actor 1:	1		1		0		1		1	
	Actor 2:		1		1		1		0		0

This approach provides a simple quantitative method of comparing the relative goodness-of-fit of two models, but leaves two questions unresolved:

- (1) How should we choose an adequate yet not redundantly large suite of models with which to try and describe the data?
- (2) How should we tradeoff fit vs. parsimony in determining whether a more complex model (larger number of free parameters) fits increasingly well enough to justify the added complexity?

On the second question, note that in our earlier thought experiment it was natural that the two models we test would fit the data better; they both had one additional free parameter over the null (IID) model. Indeed, since these hypothetical data are only ten points, we would expect the posterior probability fit of any model with ten or more parameters to reach one. That is, just as a best-fit line can perfectly fit two data points, we would expect a model with enough parameters to fit our hypothetical data with no prediction error.

A formal way of evaluating such problems is to apply a “minimal prior” posterior odds criterion, especially necessary in the general case where the models to be compared are non-nested (for example, when distinguishing between two models does not simply reduce to testing whether a parameter in the more complex model is non-zero.) Fortunately, the Bayesian statistics literature on this subject is relatively well developed. Klein and Brown (1984) show that as the informativeness¹ of the prior beliefs² of the experimenter goes to zero, the Bayesian posterior odds of model one vs. model two converge to the posterior odds ratio corrected by a factor of

$$n^{-(K_1-K_2)/2},$$

where n is the number of observations and K_1 and K_2 are the number of free parameters in each model. Intuitively, as the experimenter applies less and less prior knowledge to their interpretation of the experiment at hand, their beliefs after observing the results become easy to characterize; they differ from the straight posterior odds by a simple correction term for the number of parameters in each model. We apply this correction to the posterior odds we report in the paper.

¹The informativeness of beliefs is simply the information contained in them, or more specifically the negative of the entropy contained in the distribution, which represents those beliefs.

²The prior beliefs of the experimenter are what they believe the results of the experiment will be before observing the outcome. Specifically, it is the belief distribution over all possible outcomes of the experiment held by the experimenter, before observing these data.

With respect to the first question, if all models, which captured an essential pattern of behavior were left out, the analysis would fail to reject any of the remaining rules for failing to pick up the missing pattern. Our model set was chosen both for completeness and for ease of comparison with human-generated experimental data (for an excellent survey, see Kagel and Roth, 1995). We included every behavior strategy that could be generated by a finite automaton of four states or less, and included simplified versions of many rules commonly studied in experimental economics.

For example, the rules TFT and Punish are both widely studied in the repeated prisoners’ dilemma literature. Our TFT has a subject pull in trial 1, then copies their opponent’s last action. Our “Punish” rules are similar to a grim form of TFT, but vary in how hard it is to re-establish cooperation once cooperation has broken down. Punish 2 has a subject pull in trial 1, then pull if and only if the last 2 moves (their opponent’s last and their own last) were pulls; similarly Punish 3 is based on the last 3 moves. We also include several models, which explicitly predict that agents will mix (behave stochastically.) For example, many well-studied stimulus-response or noisy learning models have this property. These models are quite complicated but were included for completeness and comparability to human experiments; for a brief description of them see the appendix.

Finally, we included ‘Fictitious Play’ as the most widely studied example of semi-rational, myopic learning rules (for a good summary, see Fudenberg and Kreps, 1993). In fictitious play, an agent chooses the action they believe will maximize their payoff today, making this judgment by assuming that the past history of play is indicative of a stationary process in which their opponent’s actions are conditional on their own immediately preceding action. Fictitious play is myopic in the sense that the agent does not value learning, nor do they explicitly experiment, and is naïve in that it models their opponent as a simple contingent machine. It is sophisticated however, in that it updates its beliefs on exactly what this contingency is, using Bayes rule. More formally, our fictitious play is equivalent to having a subject with flat prior and myopic preferences playing a two-armed bandit problem from decision theory.

After fitting all of the above models to the data, we dropped those models, which were strictly dominated from our analysis. That is, if a model fit more poorly than another model in the sample, which also had fewer parameters, then the model was dropped from the sample. This can be thought of as providing a simple one-dimensional tradeoff; all of the models we present lie on the fit vs. simplicity frontier. We then computed our posterior odds measure for all remaining rules, and reported them in Table 2 below.

Table 2
Maximum-likelihood analysis of Tamarin strategic play

The model	A brief description of the model	dF/dx	Chi-square	Error rate	Posterior likelihood measure	# Free parameters	Adjusted conditional probability
Linear/IID model	Predictions of a linear fit.	—	37.66	NA	2.0295E-204	1	<0.001
Tit for tat	Pull then imitate.	0.172	10.82	41%	1.0452E-205	2	<0.001
Punish 2	Pull if last 2 trials were pull.	0.323	16.56	34%	9.1842E-200	2	83.168
Punish 3	Pull if last 3 trials were pull.	0.342	14.16	33%	1.0454E-200	2	9.467
Fictitious play	Play what has the highest historic payoff.	0.130	2.63	44%	1.5320E-207	2	<0.001
Stochastic fictitious play	Mix in proportion to historic payoffs.	0.176	3.00	41%	1.2272E-206	2	<0.001
Stimulus response	See discussion section for governing dynamic	0.429	6.41	29%	4.0656E-201	2	3.682
Optimal stimulus response	An optimally fit Beta stimulus-response model.	0.460	6.55	27%	4.2224E-201	3	0.134
Fictitious Mimic	Play what opponent has done most.	0.228	11.98	39%	4.3737E-203	2	0.040
Fictitious Mimic 3	Play what opponent did most over last 3 moves.	0.228	22.37	39%	1.9999E-202	2	0.181
Fictitious Mimic 5	Play what opponent did most over last 5 moves.	0.248	19.84	38%	3.4967E-201	2	3.166

All of the candidate models are clustered by color into types. We report the transformed coefficient on a Probit model and the results of a Chi-squared test of that coefficient being 0. This coefficient may be interpreted as how much more likely the subject is to pull the tool when the model predicts he will pull, then when the model predicts he will not. Half this difference is the average error rate implied by the model; too high an error rate indicates very low predictive power. After computing the posterior probability, we adjust for the number of free parameters in each model, and produce our adjusted goodness-of-fit measure. This correction for the number of fit parameters in the model by assuming the following: (1) our prior is flat over our initial suite of models, (2) errors are equal and independent, and (3) the informational content of our prior is “minimal,” i.e. allowed to be much smaller (asymptotically) than the expected amount of information contained in the data. We report the unique posterior probabilities that satisfy our three assumptions and are conditioned on the true model belonging to our set.

The first three sets of models are discussed at length in the paper, and the stimulus response models are described in the appendix. In addition, several ‘mimic’ rules survived deletion in our analysis, they are essentially more complex tit-for-tat-like rules that mimic a longer history of the opponent’s play, not just their last move. For example, Mimic 3 cooperates if and only if it’s opponent has cooperated in at least two out of their last three moves. We omit from the table several stochastic versions of this mimic rule, those rules cooperate at a rate which is the discounted history of their opponent’s past cooperation. These rules were included for completeness but had no effect on the overall analysis, so we dropped them from the table for simplicity. Finally, we ran the maximum-likelihood analysis on the first and second halves of these data separately, comparing them to test whether our results differed across those samples. No rule’s posterior fit changed by more than a percent, suggesting that the patterns we identify are stable across time.

4. Results and discussion

Looking at Table 2, the results of our posterior analysis strongly suggests that the monkeys' behavior is far from random, tending overwhelmingly towards our 'Punish 2' and 'Punish 3' strategies. A thoughtful reviewer called the first of these strategies two-tits-for-a-tat; it is essentially a form of tit-for-tat which requires more than one round of cooperation after a defection to re-start cooperation. The linear/IID model does very poorly on average, indicating that at minimum there are predictable patterns of behavior that are captured by the suite of models we apply. Generally, the rules, which explicitly predicted that actors behaved stochastically underperformed their discrete counterparts. That is to say, in times when subjects tended to pull there was very little remaining variation in the likelihood that they would pull which could be predicted by the history of play, at least with respect to our models' predictions.

Note that over 80% of the posterior weight falls on the Punish 2, or the two-tits-for-a-tat model, a model, which requires two consecutive rounds of cooperation before cooperation can re-start after a breakdown. Adding in the weights on the three-tits-for-a-tat model lifts the overall weight to over 90%, both with error rates around 30%. Overall, this strongly suggests that subjects played a relatively simple strategy, albeit with a considerable amount of noise. The two-tits-for-a-tat strategy (with an error rate of 30%) is most simply described as: begin by pulling (giving food) to your opponent at a high (70%) rate, and maintain this high rate of pulling as long as they do. Should they ever not pull the tool for you, drop the rate of pulling for them down to 30%, and only resume pulling at 70% if they ever reciprocate a pulling act of yours. This produces initial strings of cooperation followed by even longer strings of non-cooperation with moments that are occasionally punctuated by intermittent bouts of cooperation. Empirical support for this description is somewhat surprising given the large number of alternative models fitted which have more of a feel of learning and reinforced behavior.

The two-tits-for-a-tat strategy deployed by the tamarins has several interesting properties. First, it tends to result in a very high overall rate of pulling due to its relatively low fit-error rate. That is, in most of the games played, subjects began by pulling and then continued to do so as long as their opponent continued to pull. Any decision rule that follows this pattern will result in high rate of pulling so long as the error rate with which it is followed is low. In general, therefore, the lower the error rate the more likely the majority of each session will be dominated by cooperation, at least initially. We should note, however, that of course the more forgiving tit-for-tat type strategy would result in an even higher aggregate pulling level.

To understand the difficulty of "rebooting" cooperation in the dominant tamarin strategy, one only need notice that once cooperation has broken down, tit-for-tat requires only one error to resume cooperation, while a more punishing tit-for-tat requires several errors. In general, more punishing forms of tit-for-tat have proven more stable than their forgiving counterparts, both in simulated tournaments (Axelrod and Dion, 1998) and in analyses using several classes of plausible replicator dynamics (Ho, 1996.) These results also suggest that simple rules outperform far more complex and memory-intensive counterparts, so long as some minimal costs are imposed on that complexity. That is, the simple strategies we see emerge in the tamarin games are very similar to that which in human tournaments earns some of the highest long-term payoffs, largely by proving relatively robust to exploitation. This strategy also creates strong incentives for the maintenance of reciprocally cooperative play, and demonstrates a willingness to withhold cooperation from uncooperative partners in a more strategic way than would be suggested by a simple re-enforcement learning story.

A final strength of this maximum likelihood approach is its ease of aggregation across studies; this is done by simply treating any set of experiments as one large set of observations. This makes it easy to test a behavioral rule's generality over any set of games in which the same basic strategies and payoffs are present. More broadly, even in settings in which the available strategies in one setting have no obvious counterpart in another, simpler hypotheses can be tested; for example, it is easy to test the hypothesis that simple rules based only on the last two periods tend to dominate play.

Now as a general methodological problem, the incompleteness of the suite of models used in the analysis limits the possible usefulness of any maximum-likelihood approach. Unfortunately, for any reasonably experiment the problem is unavoidable; given data with N observations over two choices the set of all possible models is 2^N large, too large to be practically applied. Given this, care must be applied when choosing which models to include in the analysis. Recognizing this, we attempted to balance simplicity with completeness by including all strategies, which could be described by a four-state machine. We also included the most discussed strategies in the human experimental-economic literature, to allow direct comparison with human experimental results. This initial set is paired down (mechanically, so as not to introduce selection biases) by the dominance criterion we discussed earlier; any model which both fits worse and requires more parameters than another model, is removed.

In the other direction, there is the danger that your analysis includes too many models. It is a potential problem that with the inclusion of a new behavioral

model the data will fit the newly included strategy just by chance alone. However, as long as these data under consideration have many more observations than there are model parameters, the chance of a random model accidentally fitting the data is vanishingly small, falling exponentially with the number of observations.

In summary, we illustrate some of the potential benefits for students of animal behavior from exploiting the modeling and statistical tools that have been developed in economics, and from designing experiments that yield data suitable for testing. By providing quantitative tests of actual strategies, these techniques may more readily pinpoint why strategic play sometimes yields stable cooperation, and why this cooperation sometimes breaks down. Such analyses will help pinpoint specific psychological capacities and limitations in the evolution of cooperation.

Acknowledgements

The authors would like to thank seminar participants at Brown, Caltech, Harvard, the University of California San Diego, and Yale University for generous and insightful comments. We would also like to thank Alisia Eckert for tireless and invaluable research assistance on this project, as well as two anonymous referees for numerous detailed suggestions, all of which helped in both simplifying and clarifying the exposition of this work.

Appendix. : Description of stochastic rules used

Our stimulus-response model at time t is governed by the equation

$$\theta_{t+1}^i(s^i) = (1 - \alpha \tilde{u}_t^i(s^i))\theta_t^i(s^i) + I(S_t^i, s^i)\alpha \tilde{u}_t^i(s^i),$$

where u is the realized payoff function with range $\{0,1\}$, I is the indicator function, and θ^i is the probability with which strategy i is played. A subject is assumed to start out pulling 50% of the time ($\theta^1 = 0.5$) and to then increase their propensity to play what they did if it elicits pulling from the recipient. We fit a more general functional form that allowed negative reinforcement but found the optimal aspiration level to be 0, so we dropped this model from our analyses. Both weighted and finite Mimicry rules (mimicking a weighted average

of your opponents' last few moves) are included in our model set; subjects play a weighted average of their opponent's recent actions. TFT can be thought of as the limiting case of Mimic as the subject's memory becomes increasingly myopic.

References

- Austad, S.N., 1983. A game theoretical interpretation of male combat in the bowl and dolly spider (*Frontinella pyramitela*). *Anim. Behav.* 31, 59–73.
- Axelrod, R., Dion, D., 1998. The further evolution of cooperation. *Science* 242, 1385–1390.
- Camerer, C., Ho, T.-H., 1999. Experience-weighted attraction learning in normal form games. *Econometrica* 67, 827–874.
- de Waal, F.B.M., 1996. *Good Natured*. Harvard University Press, Cambridge.
- Dugatkin, L.A., 1997. *Cooperation Among Animals: An Evolutionary Perspective*. Oxford University Press, New York.
- El-Gamal, M.A., Grether, D.M., 1995. Are people Bayesian? Uncovering behavioral strategies. *J. Am. Stat. Assoc.* 90, 1137–1145.
- Fudenberg, D., Kreps, D., 1993. Learning mixed equilibria. *Games Econ. Behav.* 5, 320–367.
- Hammerstein, P., 2003. Why is Reciprocity So Rare in Social Animals? A Protestant Appeal. *Genetic and Cultural Evolution of Cooperation*. MIT Press, Cambridge, MA.
- Harless, D., Camerer, C., 1994. The predictive utility of generalized expected utility theories. *Econometrica* 62, 1251–1289.
- Hauser, M.D., Chen, M.K., Chen, F., Chuang, E., 2003. Give unto others: genetically unrelated cotton-top tamarin monkeys preferentially give food to those who altruistically give food back. *Proc. R. Soc. London B* 270, 2363–2370.
- Ho, T.-H., 1996. Finite automata play repeated prisoner's dilemma with information processing costs. *J. Econ. Dyn. Control* 20, 173–207.
- Houser, D., Keane, M., McCabe, K., 2004. Behaviour in a dynamic decision problem: an analysis of experimental evidence using a bayesian type classification algorithm. *Econometrica* 72 (3), 781–822.
- Kagel, J., Roth, A.E., 1995. *The Handbook of Experimental Economics*. Princeton University Press, Princeton, N.J.
- Klein, R.W., Brown, S.J., 1984. Model selection when there is 'minimal' prior information. *Econometrica* 52, 1291–1312.
- Leimar, O., Austad, S., Enquist, M., 1991. A test of the sequential assessment game: fighting in the bowl and doily spider *Frontinella pyramitela*. *Evolution* 45, 862–874.
- Milinski, M., Semmann, D., Krambeck, H.-J., 2002. Reputation helps solve the 'tragedy of the commons'. *Nature* 415, 424–426.
- Parker, G.A., Thompson, E.A., 1980. Dung fly struggles: a test of the war of attrition. *Behav. Ecol. Sociobiol.* 7, 37–44.
- Riechert, S.E., Hammerstein, P., 1983. Game theory in the ecological context. *Ann. Rev. Ecol. Syst.* 14, 377–409.
- Stevens, J.R., Hauser, M.D., 2004. Why be nice? Psychological constraints on the evolution of cooperation. *Trends Cogn. Sci.* 8, 60–65.