

FORUM

Accumulated gain in a Prisoner's Dilemma: which game is carried out by the players?

SONIA KÉFI*, OLIVIER BONNET† & ETIENNE DANCHIN‡

*Department of Environmental Sciences, Utrecht University

†Centre d'Etudes Biologiques de Chizé, UPR CNRS 1934

‡Laboratoire d'Ecologie, CNRS-UMR 7625, Université Pierre et Marie Curie

(Received 4 July 2006; initial acceptance 31 August 2006;

final acceptance 6 November 2006; published online 4 September 2007; MS. number: SC-1298R)

Keywords: altruism; blue jay; cooperation; *Cyanocitta cristata*; mutualism

Explaining cooperation has remained one of the main challenges of evolutionary biology. Following Mesterton-Gibbons & Dugatkin (1992), we define cooperation in terms of its economic consequences as a joint action for mutual benefit. Two kinds of cooperation between two individuals may be distinguished (Stevens & Hauser 2004). Selfish cooperation or mutualism occurs when both individuals receive immediate selfish benefits when they cooperate. Such selfish cooperation has been shown many times in experimental situations (e.g. Clements & Stephens 1995 with blue jays, *Cyanocitta cristata*). On the other hand, altruistic cooperation is costly to the cooperator and beneficial to the recipient in the short term. In this latter case, the temptation to cheat is high because a defector receives immediate benefits. So why perform altruistic cooperation?

A solution is to perform altruistic cooperation with relatives only (i.e. kin selection) as it increases the indirect component of inclusive fitness (Hamilton 1964, 1972). Among nonrelatives, the Prisoner's Dilemma game (Axelrod & Hamilton 1981) has been a powerful framework to investigate the conditions under which altruistic cooperation can arise (Boyd 1988). In this game, two players can choose between cooperating and not cooperating (i.e. defecting). Mutual cooperation results in the payoff r for

both players, whereas mutual defection leads to the payoff $p < r$. If one player defects and the other cooperates, the defector receives the benefit of altruism without providing any help, and obtains the payoff $t > r$, whereas the cooperator obtains the minimum gain $s < p$ (Fig. 1a). Clearly, in a single trial, the best strategy is always to defect regardless of what the other player does. Two rational players should thus defect, and would both end up with the payoff p , whereas they could get $r > p$ in the case of mutual cooperation. However, cooperation is theoretically possible in a Prisoner's Dilemma if the game is repeated several times (Iterated Prisoner's Dilemma, IPD) and if the players reciprocate (Axelrod & Hamilton 1981). Indeed, time and reciprocity can combine to eliminate the temptation to defect. In these conditions, by cooperating, the players accept a smaller immediate payoff that becomes profitable only in the long term, rather than a larger immediate benefit.

Despite important theoretical interest (Axelrod & Hamilton 1981; Boyd & Lorberbaum 1987; Boyd 1989; Nowak & Sigmund 1992, 1993; Dugatkin 1997; Pusey & Packer 1997; Wahl & Nowak 1999) little empirical evidence supports the existence of altruistic cooperation among unrelated nonhuman animals in an IPD (Clements & Stephens 1995; Stephens et al. 1995, 2002; Mesterton-Gibbons & Adams 2002; Stevens & Hauser 2004; Noë 2006; but see Milinski et al. 1997). Why such lack of empirical evidence? Stephens et al. (1995, 2002, 2006) and Stephens (2000) realized that the relation between timing of payoffs and timing of action (cooperate or defect) could affect behaviour in Prisoner's Dilemmas dramatically. More precisely, they showed that animals strongly discount the value of delayed rewards, meaning that they strongly prefer a small immediate reward rather than a larger but more

Correspondence: S. Kéfi, Department of Environmental Sciences, Utrecht University, PO Box 80115, 3508 TC Utrecht, The Netherlands (email: kefi@geo.uu.nl). O. Bonnet is now at the Bioemco ENS, 46 rue d'Ulm, 75230 Paris, France. E. Danchin is now at the Evolution et Diversité Biologique (EDB), Université Paul Sabatier, Toulouse III, CNRS-UMR5174, 118 Route de Narbonne, 31062 Toulouse Cedex 9, France.

S. Kéfi and O. Bonnet have contributed equally to the study.

(a)		Stooge	
		C	D
Subject	C	$r = 4$	$s = 0$
	D	$t = 6$	$p = 2$

(b)		Stooge	
		C	D
Subject	C	$r_{acc} = 13.75$	$s_{acc} = 11.25$
	D	$t_{acc} = 12.75$	$p_{acc} = 10.25$

Figure 1. Reward matrix (number of food pellets) for the subject. (a) Prisoner’s Dilemma reward matrix of a one trial game-unit used by Stephens et al. (2002). (b) Accumulated reward matrix of a four trial game-unit obtained using the calculations explained in Methods, and based on the values of the Prisoner’s Dilemma matrix (a) (C: cooperate; D: defect).

delayed one. This is the so-called impulsiveness hypothesis (Stephens et al. 2006). Stephens (2000) and Stephens et al. (2002, 2006) proposed that this could limit the emergence of cooperation in the IPD. They wrote a theoretical paper showing that accumulating payoffs over several trials before it is delivered to the player decreases the discounting of delayed payoffs, and therefore can favour the emergence of cooperation in an IPD (Stephens 2000). They later tested this idea on experiments with blue jays (Stephens et al. 2002).

The experiment carried out by Stephens et al. (2002) is as follows: when payoffs are accumulated, players do not have access to their payoffs at the end of each trial, but must complete a full sequence of trials before collecting the accumulated gain (Stephens 2000). In a preliminary experiment, Stephens et al. (2002) showed that accumulation of the payoffs during several trials can switch the preference of players from small immediate to large delayed payoff, that is, reduces discounting. They then tested whether they could obtain cooperation when the payoff was accumulated. In their experiments, blue jays played in pairs. One bird was designated as the ‘subject’ and the other as the ‘stooge’. The subject was free to cooperate or defect. The stooge was constrained to play a fixed strategy: unilateral defection (All-D) or Tit for Tat (TFT: cooperate initially and then copy its opponent’s previous choice in all following trials). At each trial, payoffs were delivered according to the Prisoner’s Dilemma matrix (Fig. 1a). It was either immediately accessible or accumulated in a transparent box over four trials (payoffs were thus visible to the subject, but not accessible before the end of the fourth trial). Stephens et al.’s (2002) results show that the cooperation of the subject was maintained at a high and stable level only when payoff was accumulated (reduced discounting) and when the opponent played TFT (reciprocity).

In the accumulated treatment, there is by design a timescale of action (cooperation or defection according to a Prisoner’s Dilemma at each trial) and a timescale of consequences (rewards delivery every four trials). It is a well-known result that when individuals reciprocate, after a sufficient number of interactions, the IPD changes the Prisoner’s Dilemma reward matrix into another one, where mutual cooperation becomes theoretically stable (Axelrod & Hamilton 1981; Mesterton-Gibbons & Dugatkin 1992; Dugatkin 1997). Considering the design of Stephens et al.’s (2002) experiment, we question whether the way the subject plays the game maps on to the way the experiments assume the subject is playing it. The subject plays and makes decision at each trial (cooperates or defects). However, does it base its decision on the one trial reward matrix (unit of action) or on the four trial reward matrix

(unit of reward delivery)? In other words, is the subject still in the framework of the Prisoner’s Dilemma when payoffs are accumulated? Similar questions have also been raised by Noë (2006), for whom it is unlikely that the subject experiences a payoff configuration as dictated by an IPD design in Stephens et al. (2002). In this paper we discuss the implications of accumulating payoffs in the IPD using computations and re-examinations of Stephens et al.’s (2002) experiment.

Methods

We write a simple model to address our questions. We focus on the case where the stooge plays TFT, which is a required condition to maintain the subject’s cooperation at a high and stable level in Stephens et al.’s (2002) experiment. Two main variables describe the situation of a one trial game-unit: the reward of the subject in a given combination (given by the Prisoner’s Dilemma reward matrix; Fig. 1a), and the gain of the subject cooperating after a given combination (a combination is defined as a given arrangement of what the stooge and the subject played at one trial). We calculate two similar variables for a four trial game-unit: an accumulated reward matrix, and an expected accumulated gain of the subject cooperating after a given combination. This allows us to test which of these variables best explains the subject’s behaviour as observed by Stephens et al. (2002).

Calculation of the accumulated reward matrix

The standard way of analysing a repeated game is to abandon C and D and substitute them by possible strategies such as All-D, All-C or TFT (Axelrod & Hamilton 1981; Stephens 2000). Using this approach, a reward matrix can be calculated, showing the relative merit of the strategies against each other. However, in the case of Stephens et al.’s (2002) experiment, although the stooge has a fixed strategy, the subject has a free choice and can decide to play C or D at each trial. For this reason, we decide to calculate the elements of the accumulated reward matrix as explained in the following paragraphs, taking into account all the possible sequences of four combinations.

To describe a four trial game-unit, let us consider the trees of all possible sequences of four trials and the associated scores for the subject in the case where cooperation is maintained in Stephens et al.’s (2002) experiment (i.e. the stooge plays TFT and payoff is accumulated). Figure 2 displays one of the four possible trees describing the interactions between the stooge and the

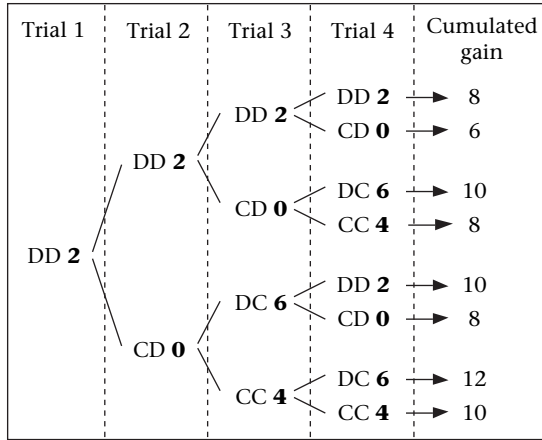


Figure 2. One of the four possible trees describing the interaction subject (left)—stooge (right) if they both begin with defection and if the stooge plays TFT (C: cooperate; D: defect). At each step, the score of the subject for the trial is given in bold. It corresponds to a number of food pellets depending on the combination subject/stooge corresponding to a Prisoner's Dilemma matrix (CC = 4, CD = 6, DD = 2, DC = 0; Stephens et al. 2002). The accumulated score of the subject for the whole sequence is given at the end of the line. Three other trees can be similarly obtained by beginning with CD, DC and CC.

subject, in the case where they both begin with defection. This tree has eight branches in total. The accumulated gain of the subject for each branch can be calculated by summing the rewards corresponding to the four combinations XY composing the branch (X, subject, and Y, stooge, can be C or D). Three other trees can be similarly obtained by beginning with the other possible combinations CD, DC and CC. There are thus 32 branches, or sequences, in total.

Now, we have the accumulated gain for the subject for all possible combinations of four trials when the stooge plays TFT. Let us imagine that the subject is at a given trial in a given combination XY. We would like to know the accumulated gain that the subject can expect at the end of the sequence, knowing only that the current combination is XY. Therefore, we calculate the mean accumulated gain of all the possible sequences of four trials, weighting the participation of each sequence by the number of times that the combination XY occurs in that sequence. A sequence where the combination XY appears two times weights two times more in the calculation of the mean than a sequence where it only appears once. We do this for the four possible combinations XY (CC, CD, DC, DD). This provides us with the elements of the accumulated reward matrix. Let r_{acc} be the element of the accumulated reward matrix corresponding to the combination CC. r_{acc} is given by:

$$r_{acc} = \frac{\sum_{i=1}^{32} X_i n_{ri}}{\sum_{i=1}^{32} n_{ri}}$$

where i is the index of the sequence ($i = \{1, \dots, 32\}$), X_i is the accumulated gain of the whole sequence i , and n_{ri} is the number of combinations CC in sequence i . The elements t_{acc} , s_{acc} and p_{acc} of the reward matrix (corresponding

to combination CD, DC and DD, respectively), can be calculated in the same way. Figure 1b displays the accumulated reward matrix.

Calculation of the mean expected accumulated gain of the subject cooperating after a given combination

Let us imagine that the subject just played a trial in which the combination was XY. We would like to know the accumulated gain that the subject can expect at the end of the sequence if it cooperates at the next trial. Therefore, we use the same method as above. Let m_{ri} (m_{ti} , m_{si} and m_{pi} , respectively) be the number of times that the event 'the subject cooperates after the combination CC' (CD, DC and DD, respectively) occurs in the sequence i , and G_{racc} (G_{tacc} , G_{sacc} and G_{pacc} , respectively) the mean expected accumulated gain of the subject when it cooperates after the combination CC (CD, DC and DD, respectively). Then:

$$G_{racc} = \frac{\sum_{i=1}^{32} X_i m_{ri}}{\sum_{i=1}^{32} m_{ri}}$$

This leads to the following accumulated gains: $G_{racc} = 22.0$, $G_{tacc} = 13.8$, $G_{sacc} = 12.2$, $G_{pacc} = 10.0$. For comparison, in a one trial game-unit, the expected gains of the subject cooperating after a given combination are: $G_r = 4$, $G_t = 0$, $G_s = 4$, $G_p = 0$.

Statistical analyses

Stephens et al. (2002) calculated the probability of the birds cooperating after a given combination. We want to investigate how well each of the two reward matrices: the Prisoner's Dilemma matrix and the accumulated reward matrix, explains the variation in the observed behaviour of the birds. As previously explained, for each game-unit (one trial or four trials), we have two types of variables: (1) the gain that the subject can expect from a given combination (variable 1a with values r , s , t , p , for a one trial game-unit, and variable 1b with values r_{acc} , s_{acc} , t_{acc} , p_{acc} for a four trial game-unit); and (2) the gain that the subject can expect when cooperating after a given combination (variable 2a with values G_r , G_s , G_t , G_p , for a one trial game-unit, and variable 2b with values G_{racc} , G_{sacc} , G_{tacc} , G_{pacc} for a four trial game-unit).

If the behaviour of the subject depends on one of the reward variables (1a, 1b, 2a or 2b), we expect the relationship between the observed probability of cooperating, and this reward variable to have a slope significantly different from zero. On the contrary, if the subject plays independently from the reward variable, we expect there to be no significant relationship. To test whether one or several of the reward variables explain the observed behaviour of the blue jays, we carried out linear regression analyses between each of the four variables, and the probability of the birds cooperating after a given combination as observed by Stephens et al. (2002). If the regression is significant for a particular variable, we can argue that this variable explains at least partly (depending of the R^2 value) the subject's behaviour. If the one trial reward variables (1a, 2a) can explain the birds behaviour whereas the accumulated

gain (1*b*, 2*b*) cannot, the birds are making decisions based on the one trial reward, and they are facing a Prisoner's Dilemma. Otherwise, if the accumulated gain explains the behaviour of the birds but the one trial gain does not, then the birds are making decisions based on the four trial reward and are playing a mutualism game. Intermediate situations are conceivable and would reveal that the birds take both types of variables into account.

Results

The accumulated reward matrix over four trials (Fig. 1*b*) is no longer a Prisoner's Dilemma matrix but a mutualism one (cooperation is always the best choice for the subject). This matrix corresponds to another game sometimes called 'the stag hunt' (Boyd 1988) and describes how mutualism (and not altruism) can arise in a population. In this context, the question of the game-unit on which the players base their choice is relevant.

Linear regression analysis between the gain (accumulated and not accumulated) and the behaviour of the blue jays as observed by Stephens et al. (2002) (Fig. 3) shows that the only variable able to predict the observed probability of the subject cooperating after a given combination (Table 1 in Stephens et al. 2002) is the variable 1*b*, that is, the expected accumulated reward after four trials (Fig. 3*a*). More precisely, the mean expected accumulated gain over four trials explains 92% ($P = 0.043$) of the variation in the observed probability (Table 1 in Stephens et al. 2002) of the subject cooperating after a given combination (Fig. 3*a*), whereas the regressions with the mean expected accumulated gain of the subject cooperating after a given combination (variable 2*b*, $P = 0.12$; Fig. 3*b*), the gain of the single trial Prisoner's Dilemma matrix (variable 1*a*, $P = 0.55$; Fig. 3*c*), and the gain of the subject cooperating after a given combination in a one trial game-unit (variable 1*b*, $P = 0.42$; Fig. 3*d*) are not significant.

So, our results suggest that the birds may be making decisions on the basis of the expected reward after four trials.

Discussion

Stephens (2000) and Stephens et al. (2002) suggest that the accumulation of payoffs (i.e. delaying the reward from a sequence of trials) could favour the emergence of cooperation in an Iterated Prisoner's Dilemma. Our calculations show that the Prisoner's Dilemma reward matrix of Stephens et al. (2002) is transformed into a mutualism reward matrix, that is, a matrix describing another game, when payoffs are accumulated during four trials. Obtaining a mutualism reward matrix over four trials is not an unexpected result, as pointed out by Stephens (2000). The reason why strategies like TFT do well in an IPD is that the reward for mutual cooperation becomes greater than the temptation to defect. In models, cooperation in Prisoner's Dilemmas has always been achieved by removing the dilemma (Mesterton-Gibbons & Adams 2002 illustrate this point nicely in their figure). The two important characteristics of accumulation as designed by Stephens et al. (2002) are: (1) the

repetition of the game over four trials; and (2) the decoupling between the timescale of action and the timescale of reward delivery by not giving access to the reward before the fourth trial. Here, the repetition of the game is responsible for the change in the type of reward matrix. The decoupling between the timescale of action and the timescale of reward delivery makes accumulation differ from a simple repeated game and raises questions. Do the birds play according to the unit of action or according to the unit of reward delivery? In other terms, do they play the game as four repetitions of the Prisoner's Dilemma or as one mutualism game?

Our analyses of the experiment of Stephens et al. (2002) show that the subject's strategy is better explained by the accumulated reward matrix than by the single trial Prisoner's Dilemma one. This suggests that blue jays can process the accumulated gain information over four trials, and that they base their decision at each trial on the four trial (mutualism) reward matrix rather than on the one trial (Prisoner's Dilemma) reward matrix, as already proposed by Noë (2006). This interpretation is supported by the preliminary experiment of Stephens et al. (2002), where individuals are capable of choosing a larger but more delayed benefit rather than a small immediate one when the trials are clumped in time and the gain accumulates in a transparent box. In case of a mutualism reward matrix, cooperating is always the best short-term choice for a four trial game-unit. So, if a four trial sequence is the game-unit for the birds, the specific conditions of the Prisoner's Dilemma are not satisfied in the experimental situation with accumulated gain, especially the short-term cost supported by the cooperator.

However, is the accumulation of the payoffs only switching the type of reward matrix and, therefore, the type of game, or does it allow the players to realize the long-term advantage of cooperating? In other words, why do blue jays cooperate in the accumulated case and not without accumulation? As already mentioned, accumulation presents the interesting property of decoupling the timescale of action and the timescale of consequence: the player makes a decision at each trial, but receives a reward after a sequence of trials. This implies that the player could be punished for a noncooperative action before he receives any payoff. In this context, defection rapidly becomes a high-cost strategy. Consequently, as pointed out by Stephens (2000), accumulation could remove the temptation to cheat, and therefore allow cooperation in a Prisoner's Dilemma context. It is noteworthy that our expected accumulated gain hypothesis is not a mutually exclusive alternative to the impulsiveness hypothesis of Stephens et al. (2006). Accumulation of gain over several trials enhances cooperation precisely because it reduces the impulsiveness (which is the rule in nonaccumulated situations).

In the case of Stephens et al.'s (2002) experiments, because the accumulated reward matrix is a mutualism one, it is difficult to conclude whether the birds cooperate because they play the mutualism game or because of the decoupling of timescales due to the accumulation. A test of why accumulation favours cooperation in Stephens experiment would be to select values for t , r , p and s , which keep a Prisoner's Dilemma relationship

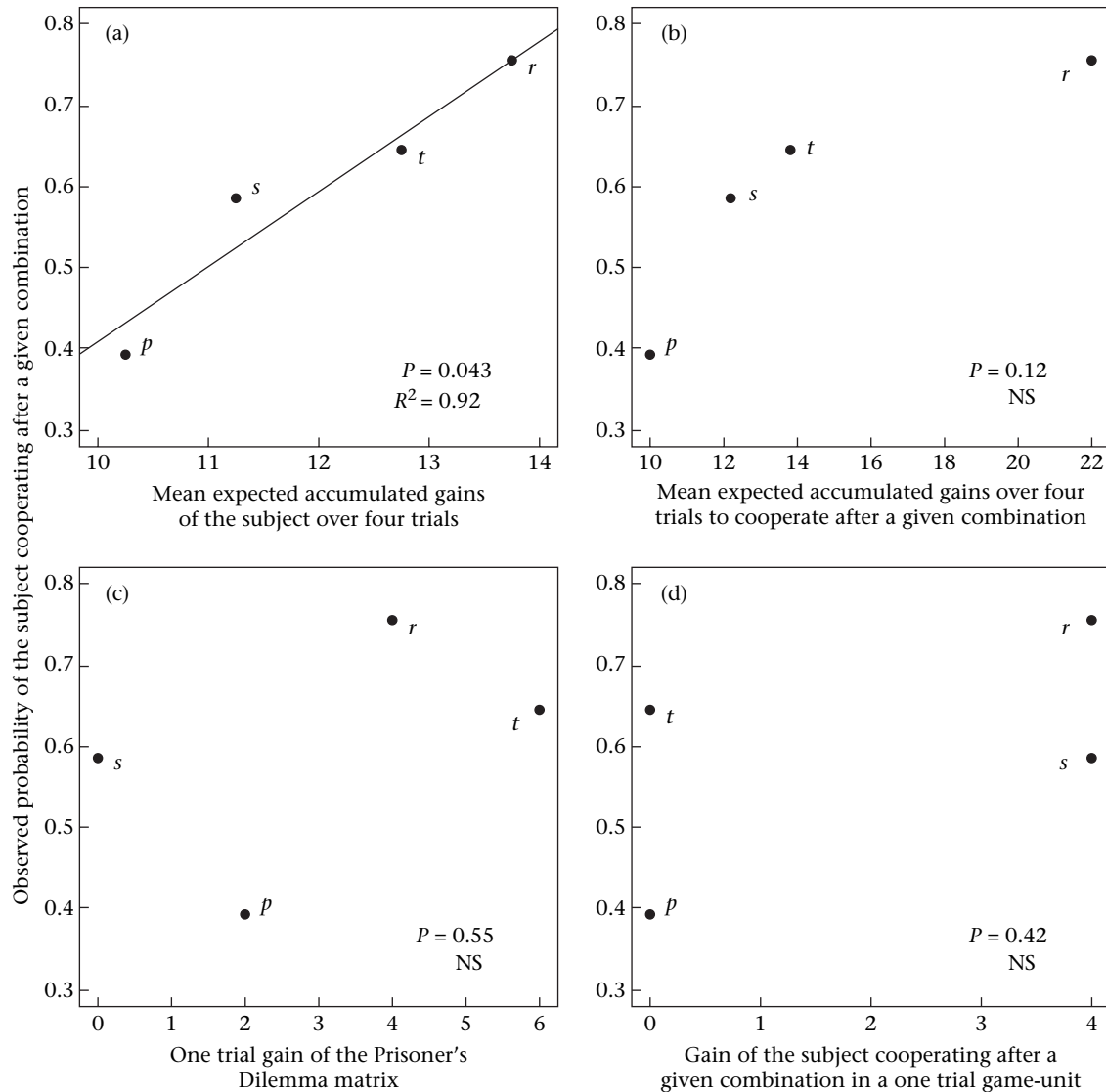


Figure 3. Linear regressions of the observed probability of the subject cooperating after a given combination (*p*, *s*, *t* or *r*, see Table 1 in Stephens et al. 2002) according to (a) the mean expected accumulated gain over four trials ($R^2 = 0.92$, $P = 0.043$, $df = 1$), (b) the mean expected accumulated gain over four trials of the subject cooperating after a given combination ($P = 0.12$, $df = 1$), (c) the actual gain of one trial according to the Prisoner's Dilemma matrix (Fig. 1a) used in Stephens et al.'s (2002) experiments ($P = 0.55$, $df = 1$), and (d) the gain of the subject cooperating after a given combination in a one trial game-unit (according to the Prisoner's Dilemma matrix) ($P = 0.42$, $df = 1$).

when accumulated over four trials. For example, $t = 9$, $r = 5$, $p = 4$ and $s = 0$ lead to the accumulated rewards (over four trials): $t_{\text{acc}} = 19.125$, $r_{\text{acc}} = 18.875$, $p_{\text{acc}} = 17.125$ and $s_{\text{acc}} = 16.875$. Can cooperation still emerge in this case? Again, it would not be totally irrational to cooperate in such a situation because of the potential punishment that could occur before the delivery of the reward. If cooperation emerges with a Prisoner's Dilemma accumulated reward matrix, it would mean that the decoupling between the timescale of action and the timescale of consequence favours cooperation because the players can react with each other's actions before the reward is released. Otherwise, it would mean that cooperation arose in the Stephens et al.'s (2002) experiment with accumulation because of the values of the accumulated reward matrix (i.e. because it changes the type of game), and that

accumulation did not lead to cooperation in a Prisoner's Dilemma context.

Cooperation in case of accumulation requires an ability to assess long-term benefit. Our analyses of Stephens et al.'s (2002) experiment suggest that blue jays can make unitary decisions on the basis of accumulated gain over four trials, which corresponds to a long-term benefit relative to the timescale of action. Actually, this is not too surprising, as several experiments with jays have showed fascinating cognitive capacities (Emery & Clayton 2001). An issue that remains unclear in Stephens et al.'s (2002) experiment is the role of the transparency of the box where the food pellets accumulate. Conducting similar experiments but with an opaque box (as in previous experiments, Clements & Stephens 1995) would allow us to test how the transparency of the box

influences the subject's behaviour. If cooperation is maintained at a high level with an opaque box, the subject clearly does not base its unitary decisions on the one trial reward, and therefore is not in a Prisoner's Dilemma situation. On the other hand, if cooperation is not maintained with an opaque box, it would mean that the information provided by the transparency of the box is essential to the subject's decision process. This information might be of two forms: either the subjects are capable of estimating the one trial reward through the transparent box, or (as suggested by Noë 2006) the transparency of the box simply allows them to make the link between their action and the delayed reward.

Stephens et al.'s (2002) experiments suggest that blue jays can not only process the accumulated gain information, but also take their opponent's former choice into account at each trial. The probability of the subject cooperating varies indeed greatly with the combination of the previous trial (Table 1 in Stephens et al. 2002) and increases with the expected accumulated gain for cooperating after a given combination. Blue jays may thus play in a combination-dependent way, which is one prerequisite for cooperation.

Stephens et al. (2002) succeeded in modifying the Prisoner's Dilemma to allow the emergence of cooperation in experimental situations. However, we suggest that the situation that they created may not be a Prisoner's Dilemma any more. Until now, cooperation in models of Prisoner's Dilemmas has always been achieved by finding ways of getting rid of the dilemma. Our calculations suggest that this may also be the case in the experiment of Stephens et al. (2002). This means that general conditions for nonhuman cooperation in empirical situations of Prisoner's Dilemmas still have to be obtained, and that the question remains of how cooperation emerges when players are confronted with the temptation to cheat. However, accumulation seems to be a probable case where cooperation could emerge in a Prisoner's Dilemma, because the players interact several times before receiving their reward. Therefore, it would be extremely interesting to investigate if accumulation can indeed favour the emergence of cooperation in an experimental Prisoner's Dilemma situation. The blue jay design developed by Stephens and collaborators would be an excellent system for the further testing of such questions.

We thank D. W. Stephens for very helpful and constructive comments and criticisms on earlier versions of this manuscript. We also thank Max Rietkerk, Walter de Back and David Craslake for helping in improving the clarity of the manuscript.

References

- Axelrod, R. & Hamilton, W. D. 1981. The evolution of cooperation. *Science*, **211**, 1390–1396.
- Boyd, R. 1988. Is the repeated Prisoner's Dilemma a good model of reciprocal altruism? *Ethology and Sociobiology*, **9**, 211–222.
- Boyd, R. 1989. Mistakes allow evolutionary stability in the repeated Prisoner's Dilemma game. *Journal of Theoretical Biology*, **136**, 47–56.
- Boyd, R. & Lorberbaum, J. P. 1987. No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game. *Nature*, **327**, 58–59.
- Clements, C. C. & Stephens, D. W. 1995. Testing models of non-kin cooperation: mutualism and the Prisoner's Dilemma. *Animal Behaviour*, **50**, 527–535.
- Dugatkin, L. A. 1997. *Cooperation among Animals: an Evolutionary Perspective*. Oxford: Oxford University Press.
- Emery, N. J. & Clayton, N. S. 2001. Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, **414**, 443–446.
- Hamilton, W. D. 1964. The genetical evolution of social behavior I and II. *Journal of Theoretical Biology*, **7**, 1–52.
- Hamilton, W. D. 1972. Altruism and related phenomena mainly in the social insects. *Annual Review of Ecology and Systematics*, **3**, 193–232.
- Mesterton-Gibbons, M. & Adams, E. S. 2002. The economics of animal cooperation. *Science*, **298**, 2146–2147.
- Mesterton-Gibbons, M. & Dugatkin, L. H. 1992. Cooperation among unrelated individuals: evolutionary factors. *Quarterly Review of Biology*, **67**, 267–281.
- Milinski, M., Luthi, J. H., Eggler, R. & Parker, G. A. 1997. Cooperation under predation risk: experiments on costs and benefits. *Proceedings of the Royal Society of London, Series B*, **264**, 831–837.
- Noë, R. 2006. Cooperation experiments: coordination through communication versus acting apart together. *Animal Behaviour*, **71**, 1–18.
- Nowak, M. A. & Sigmund, K. 1992. Tit for tat in heterogeneous populations. *Nature*, **355**, 250–252.
- Nowak, M. & Sigmund, K. 1993. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. *Nature*, **364**, 56–58.
- Pusey, A. E. & Packer, C. 1997. The ecology of relationships. In: *Behavioural Ecology: an Evolutionary Approach* (Ed. by J. R. Krebs & N. B. Davies), pp. 254–283. Oxford: Blackwell Scientific.
- Stephens, D. W. 2000. Cumulative benefit games: achieving cooperation when players discount the future. *Journal of Theoretical Biology*, **205**, 1–16.
- Stephens, D. W., Nishimura, K. & Toyer, K. B. 1995. Error and discounting in the Iterated Prisoner's Dilemma. *Journal of Theoretical Biology*, **176**, 457–469.
- Stephens, D. W., McLinn, C. M. & Stevens, J. R. 2002. Discounting and reciprocity in an Iterated Prisoner's Dilemma. *Science*, **298**, 2216–2218.
- Stephens, D. W., McLinn, M. C. & Stevens, J. R. 2006. Effects of temporal clumping and payoff accumulation on impulsiveness and cooperation. *Behavioural Processes*, **71**, 29–40.
- Stevens, J. R. & Hauser, M. D. 2004. Why be nice? Psychological constraints on the evolution of cooperation. *Trends in Cognitive Sciences*, **8**, 60–65.
- Wahl, L. M. & Nowak, M. A. 1999. The continuous Prisoner's Dilemma: I. Linear reactive strategies. *Journal of Theoretical Biology*, **200**, 307–321.