Supporting Online Material for Stephens RE 1078498, sched. for 13 December issue

Materials and Methods

Here we describe materials and methods of two experiments, as well as present SOM text with brief results. The first was a preliminary experiment to explore the effects of payoff accumulation on discounting. We then provide detail on the specific methods for the Iterated Prisoner's Dilemma experiment.

General Methods

Apparatus

The apparatus used in both experiments is shown in Fig. 1A. Each V-shaped compartment was equipped with one rear perch and two front perches, all connected to microswitches so that the experimental program could detect which perch a bird occupied. Each perch was positioned immediately below a stimulus light (Med Associates ENV-123) that could display any of several colors.

The front panel of each compartment was also equipped with a food cup, an associated pellet dispenser (dispensing 20-mg pellets, Med Associates ENV-203-20), and a custom-made small transparent box in which pellets accumulated (Fig. 1B). At experimentally determined times, a flap on the bottom of this box could be opened (by energizing a solenoid), delivering the pellets into the food cup. This device allowed us to create accumulated treatments in which food accumulated, seen but unavailable, over several trials, and unaccumulated treatments in which food was dispensed immediately at the end of a trial.

Training

Before being tested in an experiment, subjects were trained to wait on the rear perch for a fixed time and then hop forward to an illuminated front perch to obtain food. This training used conventional shaping (or successive approximation) techniques and, typically, took 3 to 6 weeks to complete.

Closed Economy

We ran both experiments as a closed economy: Subjects lived in the apparatus and obtained all of their food from it. Experimental contingencies were in effect from 0700 until 1100 and from 1200 to 1600 every day. There was a 1-hour break for apparatus maintenance and animal care from 1100 to 1200 daily; otherwise, the animals stayed in the apparatus.

All stimulus lights were dark during the periods when contingencies were not in effect. To ensure that subjects obtained sufficient food to survive, we provided supplementary food (at 1600 daily) for any bird that obtained <7 g during the day. A white noise generator provided masking noise whenever the experimental contingencies were in effect.

Preliminary Experiment

The goal of the preliminary experiment was to determine whether and under what conditions accumulation and the temporal arrangement of trials influence blue jay preferences for immediacy (or discounting). This experiment followed the self-control procedure typically used in discounting studies [see (*S1*) for an example]. In self-control studies, subjects must choose between a small immediate benefit and a benefit that is larger but more delayed. Animals with strong preferences for the small immediate benefit are said to exhibit strong discounting. This experiment considered the effects of temporal clumping of trials and the accumulation of benefits across trials.

Subjects

The subjects in the preliminary experiment were six adult blue jays of unknown sex and mixed experimental histories: Band numbers were b70, b85, b108, b223, b229, and b239.

Overview of a Trial

The sequence of events within a trial was as follows. (i) The subjects waited for a fixed time (the intertrial interval). (ii) The apparatus offered a choice between small immediate and large delayed options, by switching on lights of different colors at the front of the apparatus. (iii) The subject chose one of the options by hopping forward to the associated perch. Once the choice was made, the unchosen light was switched off, and the experimentally programmed delay to food began. (iv) When the programmed delay expired, food was dispensed, and the process began again at step 1.

Trials were organized into blocks of 32. The first 8 trials within each block were forced (or "no choice") trials in which the subjects were offered only one option (i.e., either the small immediate option or the large delayed option, but not both); the remaining 24 trials were free-choice trials. We randomly selected the light colors associated with small immediate and large delayed options for each subject, and this association was maintained throughout the experiment.

Treatments

Clumping treatments. Trials were arranged into groups of four (within the blocks of 32 discussed above). In temporally "clumped" treatments, the subject waited 345 s between clumps and then was presented with a quick succession of four trials, with a 5-s gap between each trial, so that it experienced four trials every 360 s. In unclumped trials, the subjected waited 90 s between each trial, again experiencing four trials every 360 s.

Accumulation treatments. In accumulated treatments, food was held in the accumulator—visible to the subject, but unavailable—for four trials and was delivered immediately after the fourth trial was completed. In unaccumulated treatments, the accumulator dispensed food immediately.

Delay treatments. To assess whether our manipulations influenced the blue jays' sensitivity to delay, we tested three levels of delay-to-small (one 20-mg food pellet after 5, 15, or 30 s) and two levels of delay-to-large (three 20-mg food pellets after 45 or 75 s). Each subject, therefore, experienced 24 distinct treatments (two levels of accumulation, two levels of clumping, two levels of delay-to-large, and three levels of delay-to-small). We first randomized the order in which each subject experienced the four accumulation/clumping treatments. We then randomized the order in which they received the six delay treatments within each accumulation/clumping treatment. Each delay treatment ran for 3 days, yielding ~430 free trials per treatment.

Baseline treatments. To minimize order effects, we set up the experiment so that the subjects experienced a baseline treatment before testing in each accumulation/clumping treatment. We designed the baseline treatment to be intermediate between the actual experimental treatments. The delay-to-large was 60 s, and the delay-to-small was 25 s. At the beginning of each set of four trials, the baseline program randomly determined whether the next four trials would be (i) clumped or not and (ii) accumulated or not. In clumped trials, the subject waited 285 s before being presented with a succession of four trials, one 5 s after the other. In unclumped trials, the subject waited 75 s between each of the four trials. Subjects experienced 4 days of these baseline trials before starting each new accumulation/clumping combination.

Brief Results

Repeated measures of analysis of variance (ANOVA) of the arcsine-transformed data found a significant accumulation/clumping interaction ($F_{1,5} = 9.9683$, $P = 0.0252$), suggesting that temporal clumping enhanced the effect of accumulation on discounting. Specifically, payoff accumulation shifted individual jays' preference toward larger, more delayed options when trials were clumped in time (univariate test of significance for post hoc comparison, $F_{1,5} = 13.3131$, $P = 0.0148$).

<u>Main Experiment</u>

The main experiment was designed to study the interaction between strategy and discounting in an Iterated Prisoner's Dilemma. In this experiment, pairs of blue jays repeatedly chose between two alternatives representing cooperation (C) and defection (D). To manipulate strategies, we assigned one individual in each pair to act as a stooge. The stooge was trained to simply follow lights, and in doing so it could be made to follow an experimentally determined strategy [either tit-for-tat (TFT) or All-D]. We used payoff accumulation to manipulate discounting.

Subjects

The subjects were eight pairs of adult blue jays (16 birds total). In an effort to maintain similar levels of motivation, we chose birds with similar body weights for each pair. We randomly designated one member of the pair as the subject and another as the stooge. The pairs in this experiment (subjects are listed first in each pair) were b22 and b18, b24 and b84, b70 and b1, b85 and b140, b122 and b3, b223 and b106, b229 and b130, and b239 and b208. In referring to the pairs, we cite only the subject's band number.

Overview of a Trial

The sequence of events within a single trial (or play of the game) was as follows. (i) The computer switched on the rear lights at the beginning of each trial, signaling that a new trial had started. When both subjects occupied the corresponding rear perches, the rear lights were "washed out" (by switching on an additional white light) to indicate that beginning of the intertrial interval. (ii) When the intertrial interval had expired, the subject's front lights were switched on, indicating that a choice (or play) had become available. The appropriate front light (as determined by the programmed strategy) for the stooge was switched on at the same time. The trial only proceeded to this choice phase if both individuals were on their rear perches. (iii) Next, both birds hopped forward to one of the front perches, the unchosen light was extinguished, and the chosen light was washed out as described above. (iv) Once both birds occupied their front perches, food was dispensed according to experimentally determined game matrices.

Treatments

Following the results of the preliminary experiment, all trials were arranged into clumps of four to enhance the effect of accumulation on discounting. The birds waited 345 s and then played four times in quick succession with 5 s between plays. In addition, trials were arranged in blocks of 40 (10 clumps of 4), with 8 forced or no-choice trials followed by 32 free-choice trials. During the forced choice trials, the apparatus presented only one option to the subjects (either C or D), whereas the stooge continued to play its programmed strategy (TFT or All-D).

The C and D choices were defined by their positions (see Fig. 1A). In addition, we randomly assigned the color of stimulus light associated with the subject's C and D for each treatment. The stooge's lights were both the same color.

Stooge strategy treatments. As explained above, the stooge implemented a strategy that we determined (either TFT or All-D). We simply programmed the apparatus so that the appropriate choice (C or D) was the only one available to the stooge. To program TFT, for example, we programmed the apparatus to show the inside C light, if the subject cooperated on the previous trial, and to show the outside D light, if the subject defected on the previous trial. The appropriate payoff matrix (baseline or PD) determined the subject's payoffs, just as if the stooge had chosen freely. When there were long programmed gaps between plays, such as from one day to the next or over the midday break, the TFT player began by cooperating.

Accumulation treatments. In accumulated treatments, food was held in the transparent accumulator for a series of four trials, whereas in unaccumulated treatments, the flap was opened at the end of each trial, immediately dispensing food.

Game matrices used in treatments. In describing payoff matrices, we use the conventional Prisoner's Dilemma notation. That is, a focal player obtains R (the reward for cooperation) if both players cooperate, obtains S (the sucker's payoff) if the focal player cooperates and his opponent defects, obtains T (the temptation to cheat) if the focal player defects and his opponent cooperates, and obtains P (the punishment payoff) if both players defect.

Game theoretical equilibria are stability concepts. To test the stability of cooperation in a Prisoner's Dilemma treatment, we sought to first create high levels of cooperation. To achieve this, we presented each subject with a baseline matrix ($R = 4$ pellets, $S = 2$ pellets, $T = 0$ pellets, $P = 0$ pellets) that quickly led to high levels of the C response. Subjects experienced this baseline treatment before each Prisoner's Dilemma test, for a minimum of 3 days, terminating only when the subject showed 80% C or higher for two consecutive days. All treatment variables (accumulation, stooge strategy) were in effect during this baseline phase.

During tests, the subject experienced the Prisoner's Dilemma payoff matrix $R = 4$ pellets, $S = 0$ pellets, $T = 6$ pellets, $P = 2$ pellets. The stooge obtained two pellets on all trials. Thus, the stooge received slightly less food on average than the subject did. We did this to maintain the stooge's overall hunger level and to make sure that the stooge was equally motivated to choose either the C or D alternative.

Summary of general treatment procedures. The general plan of a single treatment was as follows. (i) Stimulus colors, stooge strategy, and accumulation treatment were randomly assigned. (ii) The subject experienced the baseline matrix (with the chosen accumulation and stooge-strategy treatments in effect) until the cooperation criterion was satisfied. (iii) The subject's payoff matrix was changed to the PD matrix, and this was in effect until the subject completed 1000 free-choice trials (5 to 7 days). We repeated this cycle until each subject had completed all four accumulation/strategy combinations. In one case, a subject completed less than 1000 free trials (never less than 800 free trials), because of an error in the program that managed the transitions from one treatment to the next.

Brief Results

A repeated measures ANOVA of arcsine-transformed choice proportions found the following effects: main effect of time period $F_{2,14} = 59.61$, $P \approx 0.0$; main effect of opponent strategy $F_{1,7} = 63.72$, $P = 0.000092$; main effect of accumulation $F_{1,7} = 4.20$, $P = 0.08$; two-way interaction of time period and opponent strategy $F_{2,14} = 25.12$, $P = 0.000023$; two-way interaction of time period and accumulation $F_{2,14} = 1.28$, $P = 0.31$; two-way interaction of opponent strategy and accumulation $F_{2,14} = 0.42$, $P = 0.54$; and three-way interaction of time period, opponent strategy, and accumulation $F_{2,14} = 4.31$, $P = 0.034$.

References and Notes

S1. D. W. Stephens, D. Anderson. *Behav. Ecol.* **12**, 330 (2001).