Material and Method

All experimental procedures were approved by the ethics committee of the IByME-CONICET and were conducted according to the NIH Guide for Care and Use of Laboratory Animals.2.1 Subjects and Housing.

**Subject.** Two groups of male Long-Evans rats (300–330 g) of two months of age were provided by the IBYME-CONICET. Twelves males were subjects and six (Nstooge=6) identical rats were opponents. At weaning time all subjects were housed in groups of two rats per cage to allow social interaction, whereas each stooge was housed in single individual cages. We had 12 stainless-steel cages altogether with sawdust as bedding and metal lids. All rats were food restricted to maintain animals at between 90-95% or 80-85% of free feeding body weight for subjects and stooges respectively. All animals were kept in a well-ventilated, temperature-controlled room (22ºC±2ºC) with a 12/12 h light/dark cycle (lights on at 8 am). Tap water was available ad libitum. To ensure that animals obtained sufficient food to survive, we provided supplementary food (at 22hs) for any rats that obtained the average amount of pellet to maintain body weight.

**Housing.** All behavioral procedures were performed during the light phase of the light/dark cycle, using a standard operant chambers (MED associates Inc., USA). White noise with a flat power spectral density was used to reduce sensibility to ambient noise. The experiment was conducted in special ad hoc dual chamber, see fig. 1A. We placed two Med association standard operant chambers facing each other in such manner that each rat could make olfactory and eye contact through metal mesh windows. Each standard chamber was equipped with: two not retractable levers and two stimulus light placed one over each other respectability and in the center put a illuminated feeder. The window's height were such that the lever's height was 80% of maximum height of the forepaws while rearing (F. Cabrera et al., 2013).

**Procedures.** During the whole duration of experiment, every actor were trained for 2 session per day regardless of treatment (handing, habituation, training and iPD). Each typical trial began in the darkness during 5 second inter-trial interval(ITI) and then the stimulus light were illuminated for either 45 second or until a lever was press. Before deliver any reward the feeder light was 1 second turn on. The standard experiment session had 30 trials.

**Pre-experimental training.** At weaning time when the animals were moved to housing room started a handing procedure to decrease the stress by experimenter manipulation and finished when animals had 60 days old. The pressing lever training on rats was done through a successive approximations procedure, “Shaping” (Mazur, 1994, page 122) on a single chamber. The chamber was equipped as described above and was added a external lever. It training finished when the rat made the task for at least two session. The standard trial procedure: each trial began when one or two light stimulus was turned on and the lever below was activated, then if levers wasn't pressed and 45 second had elapsed, the trials finished and not reward was given. But nevertheless, if levers was pressed, the light turn-off and the positive or negative reinforcement was dispensed, when 5 second elapsed all light were turn off. Finally, 5 second ITI started. The subject rats was exposed to a basic toggle task training in dual iPD chamber. The purposed of the training was to incorporate two lever and balance their lever's preference. All rats reached the 80% good trial on two consecutive sessions criteria. The opponent learned to develop a mimic tit for tat strategy by to learn a basic operant task. The opponent rat group was exposed to a fixed ratio (FR) schedule where the side of pair light-lever that was turn on was randomly assigned in each trials along the session. The light-lever pair that worked was pseudo-randomly assigned. Pseudo-randomly means that the assigned was random but not allowed the same assigned for more than four time. The training was developed during five days and the FR schedule was increased step by step: day 1º the FR=2 and day 2º to 3º the FR=3, then day 4º the FR=4 and finally day 5º the FR=5. If the rat presses the lever on ITI time, it received 2 second as punishment that was added to ITI.

**Experiment.**

The main experiment was designed to study the reciprocal altruism in a iterated Prisoner's Dilemma game(iPD). We used a pay-off matrix in which the iPD states(outcomes) meets that T>R>0>P>S

(Temptation, Reward, Punishment and Sucker respectability). The T and R outcomes were positive reinforcement and P and S was negative reinforcement(timeout). The pay-off matrices are shown in figure 1A. Thus, if both choose cooperation option (C) or not cooperation option (D), the subject pay-off was 1 pellets or 4 second timeout, respectively. If the subject choose D and the opponent choose C, the subject received 2 pellet and in reverse situation, the subject received 8 second timeout and no pellet in this trial. The opponent's pay-off was alway 1 pellets whether it pressed the correct lighted lever. The amount of pellets preference was previously tested on discrimination treatment and the rat shown preferences for 2 pellet than 1 pellet.

The main experiment was conducted for at least 10 days with 2 session per day of 30 trials. The experiment was stopped, if the average behavior didn't change along 5 consecutive sessions. Each subject was tested with the same opponent while during the experiment.

The single trial procedure was: (1) The stimulus light was turned on stating that trail was started. The focal rats had both lights turned on (this was done throughout the iPD experiment) and the opponent only one for the tit for tat strategy. (2) After both rats done their response, the light was turned off and the rewards computed. (3) The feeder's light was lighting one second before a reward was delivered. If the opponent pressed the lever, it always received one pellet irrespective the focal rat's choice. If timeout was assigned, the focal rat feeder's light remained off and the delay time start, but the opponent's feed was lighted and its one pellet reward was delivered. (5) After either five second eating time expired or timeout complete, all light was turned off and fixed ITI time started.

The second experiment was designed to evaluated how the change on positive reward modify the rats behaviors. We used the same procedure as in the main experiment with six naive rats. After to exposed all rats to a pay-off matrix where the temptation was increase one pellet respect to main experiment pay-off, we divide rats in two groups. A cooperative group and a defect group. The cooperative groups was expose to a pay-off matrix with the temptation (T) biggest than the reward (R). The defect group was exposed to a matrix that empower the cooperative behavior, but it was low contrast among positive reward.

**Statistic.** All statistical analysis were performed using statistic library from octave open source software. To analysis the choice stability of rats in the iterated game, we used the last 5 session in which the rat's behaviors had more stable. The mean of cooperation per rat was obtained by count the number of times a rat chose the cooperation lever per session. To compare probability to cooperate after each outcomes was assess using a chi square test. The probability of cooperation given specific outcomes was compared from chance, 0.5. We performed a Chi-square goodness of fit test with Bonferroni corrected value of 0.05/n.

To compare mean rate of the different outcomes for each game we performed a Friedman's ANOVA test with the Bonferroni corrected a value of 0.0125. We used Wilcoxon rank-sum test to evaluated the mean between animals and between treatment.

The individual's decision rules can be described by a transition vector t, r, p, s and markov chain diagram. The t, r, p, s value represent the probability of cooperation when the previous trials resulted in outcomes of R(reward), T(temptation), S(sucker) or P(punishment), respectively (Stevens and Stephens,2002). If every component of this vector is 0.5, the agent's decision rule is random irrespective of the last outcome. With the diagram of Markov chain we computed and gave a graphic representation of the complete making decision rule for each rats.