PONE-D-17-44749  
High mutual cooperation rates in rats learning reciprocal altruism: the role of payoff matrix  
PLOS ONE  
  
  
Please ensure that your manuscript meets PLOS ONE's style requirements, including those for file naming. The PLOS ONE style templates can be found at  
[http://www.journals.plos.org/](http://www.journals.plos.org/plosone/s/file?id=wjVg/PLOSOne_formatting_sample_main_body.pdf) and [http://www.journals.plos.org/](http://www.journals.plos.org/plosone/s/file?id=ba62/PLOSOne_formatting_sample_title_authors_affiliations.pdf)  
  
1. We suggest you thoroughly copyedit your manuscript for language usage, spelling, and grammar. If you do not know anyone who can help you do this, you may wish to consider employing a professional scientific editing service.   
  
Whilst you may use any professional scientific editing service of your choice, PLOS has partnered with both American Journal Experts (AJE) and Editage to provide discounted services to PLOS authors. Both organizations have experience helping authors meet PLOS guidelines and can provide language editing, translation, manuscript formatting, and figure formatting to ensure your manuscript meets our submission guidelines. To take advantage of our partnership with AJE, visit the AJE website (<http://learn.aje.com/plos/>) for a 15% discount off AJE services. To take advantage of our partnership with Editage, visit the Editage website ([www.editage.com](http://www.editage.com/)) and enter referral code PLOSEDIT for a 15% discount off Editage services.  If the PLOS editorial team finds any language issues in text that either AJE or Editage has edited, the service provider will re-edit the text for free.  
   
Upon resubmission, please provide the following:

* The name of the colleague or the details of the professional service that edited your manuscript
* A copy of your manuscript showing your changes by either highlighting them or using track changes (uploaded as a \*supporting information\* file)
* A clean copy of the edited manuscript (uploaded as the new \*manuscript\* file)

Reviewers' comments:  
Reviewer's Responses to Questions  
**Comments to the Author**  
  
5. Review Comments to the Author  
  
Please use the space provided to explain your answers to the questions above. You may also include additional comments for the author, including concerns about dual publication, research ethics, or publication ethics. (Please upload your review as an attachment if it exceeds 20,000 characters)

Reviewer #1 has suggested:

1. “I strongly suggest the authors to add some discussion on this point by citing following literatures that introduced universal scaled dilemma strength in 2 various by 2 games.”

**Based in your suggestion we modified our manuscript.** **we add in the MS discussion:**

**“In a standard pay-off matrix is possible to calculate the dominant cooperative strategy, through a standard mathematical analysis [Tanimoto and Sajara, 2007 and Wang et al., 2015], but it demand to mathematics operation among all components of payoff matrix and in our case means operating between positive (pellet) and negative (times) reinforcement . For this reason, we make a single analysis only a positive part of reinforcing matrix and then consider the aversive punishment from the accumulated timeout.”**

1. “A technical suggestion; In the final MS, the authors should provide more clear visual materials. The current figs seem too ugly to read.”

**When the figures were uploaded their resolution was degradated.**

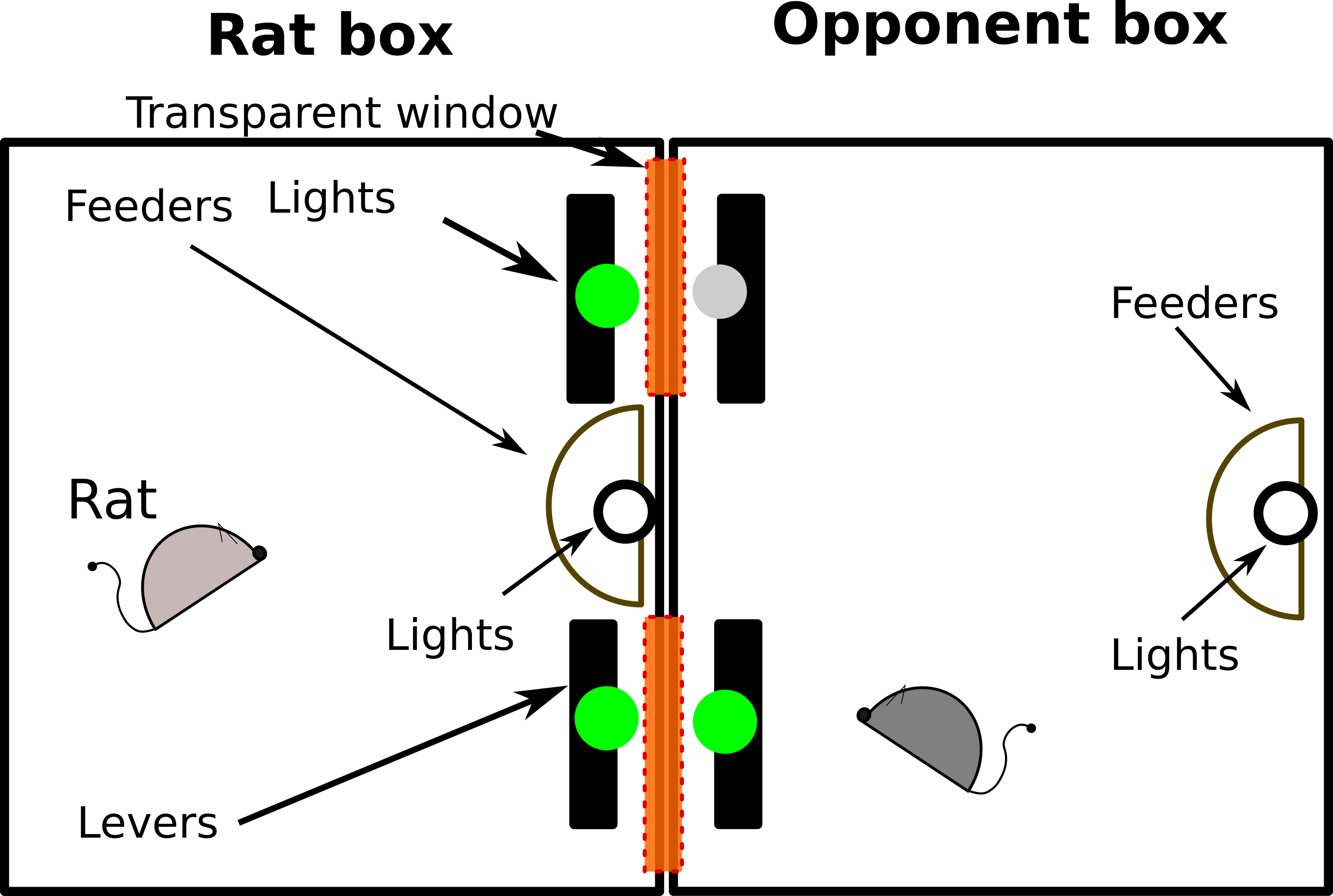
Reviewer #2 has suggested:

1. “The English would benefit from a more thorough inspection.” “the authors should look up the difference between 'defect' and 'defecate'.”

**Spelling was checked.**

1. “Photos, for instance, would likely make the description much clearer.”

**We had add some setup graphic.**

**The figure caption:**

**“S2 Fig. Description of dual operand conditioning chamber. Two operant box in front of each other in such as way that windows were aligned. In the front panel there are two stimulus lighting (green=on/gray=off) and two levers and windows(red shadow). In the**

**Opponent box only a light was lighting per trials and the subject had both light on.”**

1. why were these particular amounts of pellets chosen?

**Added to MS:**

**“Here, we design a iPD setup to test learning of reciprocal altruism based on a combination positive (pellet) and negative (timeout) reinforcement. Our goal was to choose the amounts of pellets in order to minimized the amount of positive reinforcement earned in long term and in conjunction with timeout maximize the profit's contrast in long term. Those maybe help rats to discriminate the amount of reward in long term.”**

1. “what the predicted behaviour of a rational player would be.”

**Line292:**

**“… In this case, ALLC strategy is the only one that maximizes positive reinforcement and minimizes the negative one (Pareto Optimum). Since negative reinforcement is timeout, ALLC strategy gives more food per unit of time. In this case, the role of the negative reinforcement appears. ...”**

1. “The notation is messy.” “letters R, T, P, S are used to describe game outcomes. In others, it seems to indicate the rewards given in those outcomes, then later they seem to be state occupancies in a Markow model.”

**We fix the notations. The states were named R, T, P and S. The reward given in those outcomes: PR, PT, PP, PS. And the probability of each state: p(R), p(T), p(P), p(S).**

1. “ l. 184-189 ... I suggest restructuring the text, or putting the numbers in a table”

**The text in line 184-189 (and also in other lines) was restructured and we elaborated a table as a data summary.**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Exp | Detalle | Coopertion | Probability State | | | | Transition Probabilities | | | | p<0.01 |
| p(T ) | p(R) | p(S) | p(P ) | p(c|T−1) | p(c|R−1) | p(c|S−1) | p(c|P−1) |  |
| 1 | coop | 0.86±0.05 | 0.1 | 0.76 | 0.1 | 0.04 | 0.76 | 0.85 | 0.93 | 0.87 | SI |
| No coop | 0.36±0.03 | 0.44 | 0.38 | 0.32 | 0.32 | 0.25 | 0.19 | 0.33 | 0.23 | NO |
| reverse | 0.87±0.04 |  |  |  |  |  |  |  |  |  |
| 2 | A | 0.88±0.01 | 0.09 | 0.80 | 0.03 | 0.08 | 0.65 | 0.90 | 0.87 | 0.94 | SI |
| B | 0.64±0.13 | 0.23 | 0.34 | 0.21 | 0.22 | 0.47 | 0.55 | 0.56 | 0.65 | NO |
| 3 | A | 0.61±0.10 | 0.18 | 0.44 | 0.214 | 0.17 | 0.45 | 0.64 | 0.62 | 0.78 | SI |
| B | 0.71±0.04 | 0.20 | 0.51 | 0.09 | 0.20 | 0.62 | 0.66 | 0.67 | 0.66 | SI |

1. “the results, I am not convinced that the results presented in fig 1D show "two well defined strategies"

**We expand our explanation about fig 1D. We assessed whether exist an intersection on set of level´s choice between animals and simulated subject:**

**“Due to the fact that many sequences of lever pressing can give the same amount of**

**reward and/or timeout, independently of the cooperation level. We analyzed the**

**relationship between total reward and timeout for each animals respect to a simulated**

**population. Thus, for each animal we compared those values with a simple regression**

**model fitted to a population of 100.000 simulated individuals when their cooperation**

**level was set to 60%, see fig 1D. Each simulated individual had a different complete**

**strategy and each strategy was a combination of thirty C and D choice (session length).**

**An individual that play a iPD game with a 60% of her choices in C will be near to the**

**line regardless of her strategies. As can be seen in the figure, the higher the cooperation**

**levels, the larger the total reward and the lower the total timeout. Accordingly the fig.**

**1D, the group of cooperator’s rats developed a complete different behavior respect both**

**line regression with 60% of cooperation and the no cooperator group. It means that not**

**exist any strategy with low level of cooperation that get high level of reward and small**

**timeout as in the cooperative group.”**

1. “At least, P© seems to span the entire color gamut**”**

**We zoomed the cooperation bar.**

Reviewer #3 has suggested:

1. I think it is better to say something like:  
   **"It is not obvious at the first sight how could Darwin's theory natural selection explain altruistic behaviour."**

**We chance the sentence.**

1. Sentences form the discussion should be in the introduction:  
   “In order to evaluate if animals developed ALLC strategy by place preference or by reward maximization, we applied a reversion treatment” **(ln. 261)**  
   or  
   “we wanted to evaluate if a change in the payoff matrix could modify their behaviour, to do so we studied the effect of modifying positive reinforcements” **(ln.264)**

**We modified the introduction and in the last paragraph we added your suggestion:**

**“Here, we design a iPD setup to test learning of reciprocal altruism based on a combination positive (food) and negative (timeout) reinforcement, where our goal was to choose the amounts of pellets in order both to maximization contrast among pay-off get by each iPD states and to minimized the amount of food that each rat gets in long term. These maybe help the rats to discriminate the amount of reward at long term. In order to evaluate if animals developed ALLC strategy by place preference or by reward maximization, we applied a reversion treatment. Finally, we evaluated how payoff matrix components promotes or disrupt altruistic behaviors.”**

1. There needs to be a table that sums up the experimental treatments.  
   **We made a summary table with all relevant treatment data.**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Exp | Detalle | Coopertion | Probability State | | | | Transition Probabilities | | | | p<0.01 |
| p(T ) | p(R) | p(S) | p(P ) | p(c|T−1) | p(c|R−1) | p(c|S−1) | p(c|P−1) |  |
| 1 | coop | 0.86±0.05 | 0.1 | 0.76 | 0.1 | 0.04 | 0.76 | 0.85 | 0.93 | 0.87 | SI |
| No coop | 0.36±0.03 | 0.44 | 0.38 | 0.32 | 0.32 | 0.25 | 0.19 | 0.33 | 0.23 | NO |
| reverse | 0.87±0.04 |  |  |  |  |  |  |  |  |  |
| 2 | A | 0.88±0.01 | 0.09 | 0.80 | 0.03 | 0.08 | 0.65 | 0.90 | 0.87 | 0.94 | SI |
| B | 0.64±0.13 | 0.23 | 0.34 | 0.21 | 0.22 | 0.47 | 0.55 | 0.56 | 0.65 | NO |
| 3 | A | 0.61±0.10 | 0.18 | 0.44 | 0.214 | 0.17 | 0.45 | 0.64 | 0.62 | 0.78 | SI |
| B | 0.71±0.04 | 0.20 | 0.51 | 0.09 | 0.20 | 0.62 | 0.66 | 0.67 | 0.66 | SI |

1. “There needs to be a table that sums up the various abbreviations.”
2. “P and S values are mixed up in the figures (Figs 2 and 3).”

**Yes, the matrices were wrong, we had fixed the mistakes changing P and S values.**

1. “The authors should explain what is a reversion treatment (ln.106).”

**At the end of line 106 we added the following explanation:**

**“The reversion consist on inter-change sides of C and D lever in both subject and opponent chamber. In that sense, if a animals has a place preference behavior, he will not learn the new side that maximize reward.”**

1. The authors should explain what is a 'normalized reward' (Fig 3D).

**Normalized reward explanation:**

**“A normalized reward was calculated as quotient between total reward obtained in the session and the maximum reward achieved using the best strategy. Considering that opponent always played Tit for tat strategy, the best strategy depend on the pay-off matrix values. When the matrix favor the cooperation ALLC was the best, but when matrix favor no cooperation alternate between C and D was the best strategy.”**

1. “explaining their unique success” " What is the crucial difference?

**We put effort into explaining main success and in line 235 we added follow paragraph:**

**“To our knowledge, results show for the first time high levels of cooperation (86,11%) and mutual cooperation (76,32%) in iPD, see Fig 1B. Several works have tested reciprocity using iPD game with similar version of standard matrix, showed that animals prefer short-term benefits or only improve a poor level of cooperation [4, 9, 20, 30, 31] or have had to use a treatment to enhance cooperation preference [10, 23, 29, 34]. The main differences with other research works are the levels of cooperation and mutual cooperation achieved. A possible explanation is that using a standard matrices, as [P T = 6, P R = 4, P P = 1, P S = 0], animals could not discriminate between amount of reinforcements obtained in long-term option respect to short-term option, mainly due to after several trials they got amounts of reward difficult to recognize [24]. For example, if a rat played in four sessions [C C C C] got 16 pellets and if played [C D D D] got 12 pellets. In our framework, a rats using the**

**same choices will earn 4 or 3 pellets plus 16 second timeout. The rats can able to recognize small amounts and timeout degrades the profit.”**

1. **I could not find any datafiles in the Supplementary Information.**
2. “Spelling: defecated, defecat. I assume it is “defect” and “defected”.

**Yes, the word is “defected”. We checked the lines.**

Due to the above deficiencies I cannot recommend the publication of the paper in this form. As it stands the manuscript needs a major revision and I highly encourage the authors to submit a revised version.

6. If you would like your identity to be revealed to the authors, please include your name here (optional).  
  
Your name and review will not be published with the manuscript.   
  
Reviewer #1: (No Response)  
  
Reviewer #2: (No Response)  
  
Reviewer #3: (No Response)