Does reward positivity encode trial-by-trial reward prediction error? A model-based EEG analysis

Ka Chun Wu, Isaac Ip, Fiona Ching, Heytou Chiu, Rosa Chan,
Bolton K. H. Chau, Savio W. H. Wong, Yetta Kwailing Wong

Contact: kachunwu7-s@link.cuhk.edu.hk

Background

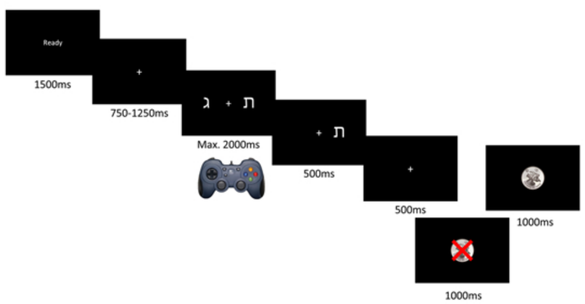
- Human compute the Reward Prediction Error (RPE), i.e. the discrepancy between the prediction and actual outcome, and adjust future behaviour accordingly
- Reward Positivity, a feedback-locked ERP sourced from anterior cingulate cortex (ACC), was thought to encode RPE [1]
- Current study used a model-based approach to explore the effect of RPE on RewP

Subject

- Thirty-seven healthy adults aged between 21 and 40
- 19 male, 18 female
- M_{age} = 26.97

Design

- Probabilistic Reversal Learning (PRL) Task using Hebrew characters
- Each character has either 85% or 15% chance of getting reward
  (a picture of a coin)
- 4 blocks (with different character pairs), each has 120 trials



Hierarchical Bayesian Modelling

- RPE is estimated by by hBayesDM package (based on Stan) in R [2]

Reward-Punishment Fictitious Update Model

- $\alpha+$ = *learning rate of reward*
- $\alpha-$ = *learning rate of punishment*
- τ = inverse temperature
- β = Indecision point

$$RPE_t = R_t - V_t^{Chosen}$$

$$V_{t+1}^{Chosen} = V_t^{Chosen} + \alpha^{+/-} \cdot RPE_t$$

$$V_{t+1}^{Unchosen} = V_t^{Unchosen} - \alpha^{+/-} \cdot RPE_t$$

$$P(Chosen) = \frac{1}{1 + e^{\tau \cdot [\beta - (V^{chosen} - V^{Unchosen})]}}$$

$$P_{t+1}(Unchosen) = 1 - P_{t+1}(Chosen)$$

EEG Analysis

- Hierarchical LInear Modeling by the LIMO-EEG toolbox (EEGLAB plugin) [3]
- 1st level analysis:
  - Channel-based GLM estimated by Ordinary Least Square method
  - Feedback types, blocks, and trial-by-trial RPE as regressors
- 2nd level analysis:
  - One-sample t-test of condition contrast
  - Multiple comparison correction using spatiotemporal clustering
  - Feedback contrast (reward vs. nonreward)
  - RPE contrast (positive RPE vs. negative RPE)

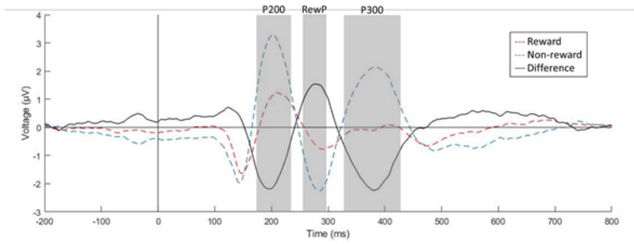
Results

Model Fit:

- LOOIC = 5918.53 (compare to 14794.54 of WSLS model)
- Posterior predictive check shows good resemblance to real data

Grand Average ERP:

- Significant amplitude differences at RewP time window (270-300ms)
- Enhanced P200 (183-233ms) and P300 (326-426ms) for non-reward feedback



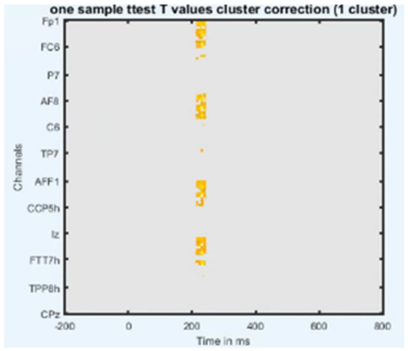
- Replicating results from previous studies [4]

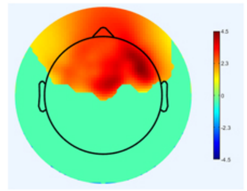
Feedback contrast (reward vs. non-reward):

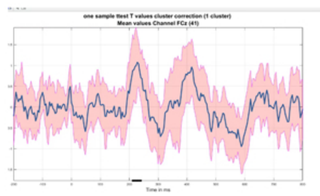
- No significant cluster is found

RPE contrast (+RPE vs. -RPE):

- A significant cluster started at 212-268 (mean beta value = 0.75, 95% CI [-.001 1.55], p-value = .04)
- Encompassing frontocentral electrodes







Discussion

- A more positive RPE predicts a more positive EEG response at frontocentral region
- However, the corresponding time is earlier than the typical RewP time window, which could influence RewP
- The results reveal time windows crucial to RPE computation, which lays between overlapped P200 and RewP
- **Conclusion:** RewP itself did not encode RPE, but is influenced by the computation preceded it

[1] Proudfit, G. H. (2015). The reward positivity: From basic research on reward to a biomarker for depression. *Psychophysiology, 52*(4), 449-459.
[2] Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational Psychiatry, 1*, 24-57.
[3] Pernet, C. R., Chauveau, N., Gaspar, C., & Rousselet, G. A. (2011). LIMO EEG: a toolbox for hierarchical LInear MOdeling of ElectroEncephaloGraphic data. *Computational intelligence and neuroscience, 2011*, 1–11. https://doi.org/10.1155/2011/831409
[4] Martínez-Selva, J. M., Muñoz, M. A., Sánchez-Navarro, J. P., Walteros, C., & Montoya, P. (2019). Time course of the neural activity related to behavioral Decision-Making as revealed by event-related potentials. *Frontiers in behavioral neuroscience, 13*, 191.