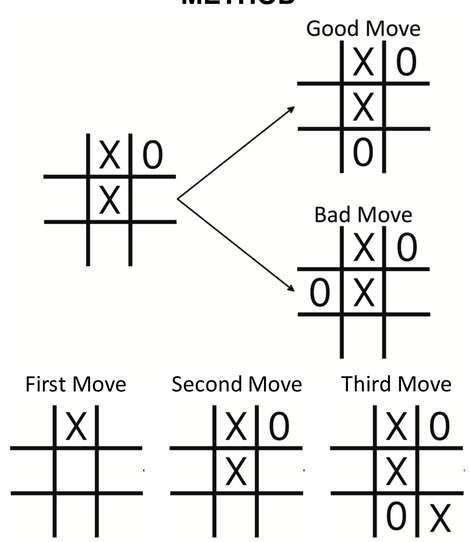
# Reward Processing when Evaluating Goals: Insight into Hierarchical Reinforcement Learning

Chad Williams, Clay Holroyd, and Olave Krigolson

## INTRODUCTION

- Past research has demonstrated reward processing when observing the feedback of others
- Whether it occurs when observing the actions of another is still poorly understood
- Furthermore, it is unclear whether it is affected by goal proximity
- We hypothesized to see reward processing when perceiving the actions of a computer
- Furthermore, we hypothesized that it would scale to goal proximity

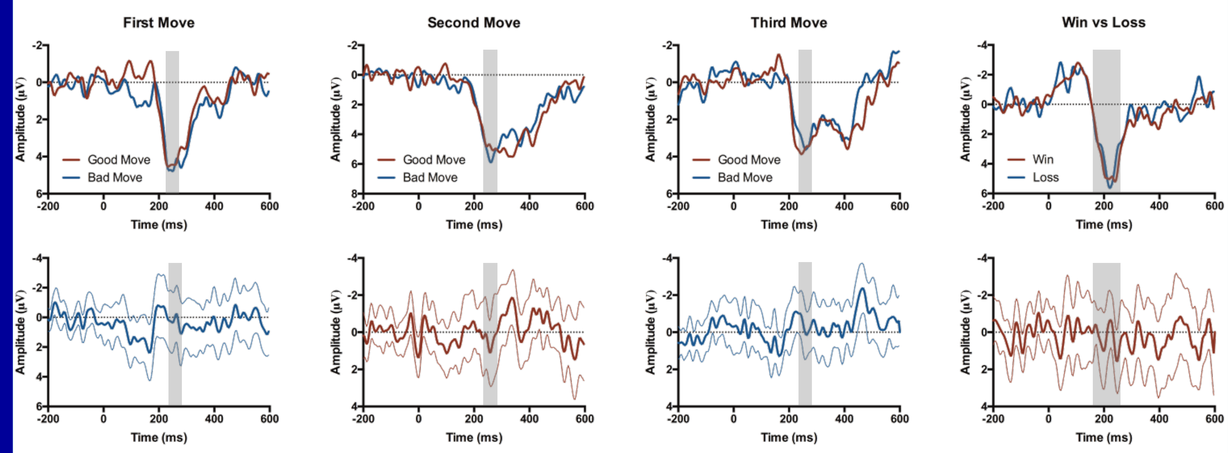
## METHOD




- 18 university students (M: 20 yrs, 14 Females)
- Played Tic Tac Toe against a computer
- Each game began within an ongoing game where it was either the computers first, second, or third move
- The computer played a good or bad move
- Neural underpinnings (e.g., reward positivity) were analyzed in response to these initial moves
- Good and bad moves were determined using a computational model that learned the values of game board states

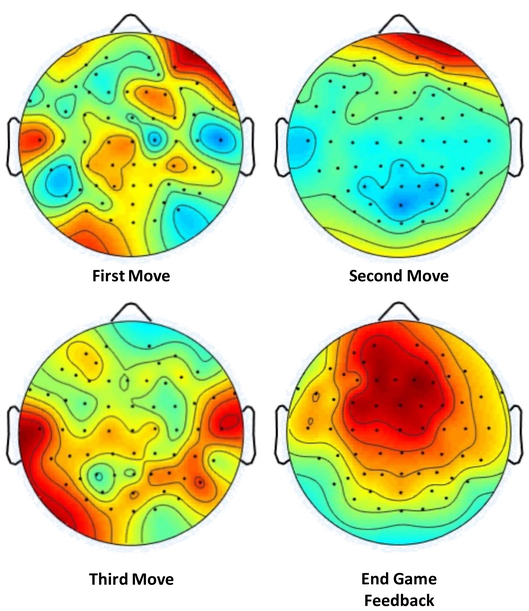
## RESULTS

### Conditional and Difference Waveforms (FCz)

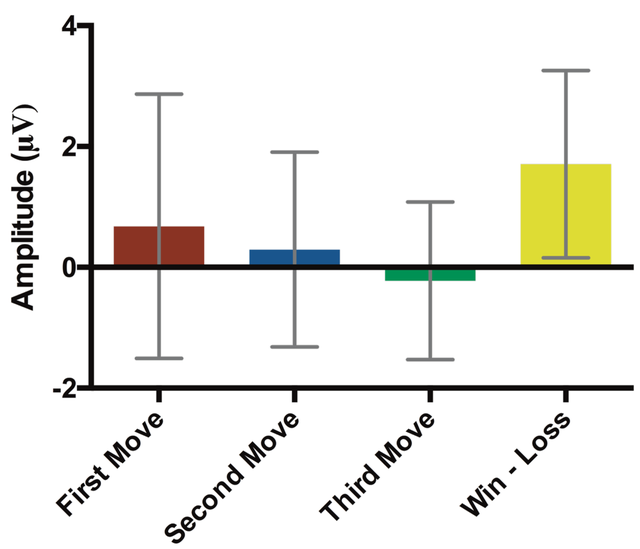



### Reward Positivity

#### Topographic Maps of Difference Waveforms




#### Difference Waveform Amplitudes




## CONCLUSIONS

- Prediction errors, as measured by the reward positivity, were not present in response to a computers actions
- Consequently, we were unable to examine whether reward processing scaled to goal proximity
- Interestingly, we found a reward positivity to the outcome of games
- This indicated that higher level reward processing occurred to sequences of behaviours that led to task goals
- This research is in support of hierarchical reinforcement learning

## REFERENCES

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*(4), 679–709. http://doi.org/10.1037/0033-295X.109.4.679

Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in cognitive sciences, 16*(2), 122-128.

Yeung, N., & Sanfey, A. G. (2004). Independent Coding of Reward Magnitude and Valence in the Human Brain. *The Journal of Neuroscience*, *24*(28), 6258–6264. http://doi.org/10.1523/JNEUROSCI.4537-03.2004

## CONTACT

Chad Williams
The Neuroeconomics Laboratory
University of Victoria
www.neuroeconlab.com
ccwillia@uvic.ca

INNOVATION.CA
CANADA FOUNDATION FOR INNOVATION | FONDATION CANADIENNE POUR L'INNOVATION
NSERC CRSNG

University of Victoria | Neuroeconomics Laboratory