The Power Curve of the Brain: Reward Prediction Errors Follow Learning Curves

C.C. Williams and O.E. Krigolson

University of Victoria | Theoretical and Applied Neuroscience Laboratory

Centre for Biomedical Research, University of Victoria

Poster: F103

INTRODUCTION

- It is well established that performance during the early stages of learning follows a non-linear trend.
- These learning curves demonstrate rapid performance changes in the initial stages of learning that depreciate with succeeding trials (e.g., the power law of practice[1]).
- Although theoretical predictions of reinforcement learning suggest trial-by-trial changes in behavioural and neural measures follow learning curves[2,3], it has never been investigated whether neural signals of learning (i.e., the reward positivity[4]) follow these same properties.
- Here, we sought to provide a complete investigation of trial-by-trial changes in behavioural and neural data across a computational model and empirical data.

METHODS

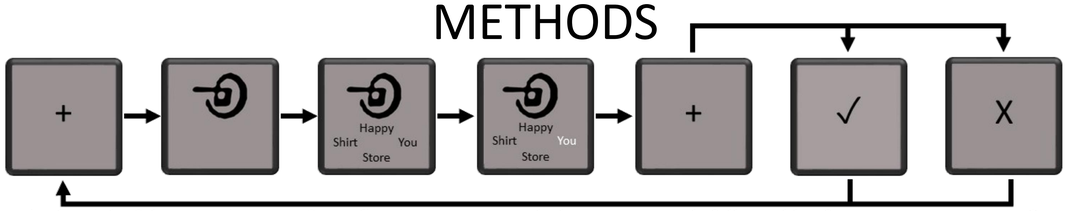

Figure 1. A single trial of the experiment. After a fixation cross a symbol was presented, followed by four response options. Once a response was selected, it was highlighted in white. After a fixation cross, simple correct and incorrect feedback was presented.

- Both the computational model and participants (n = 29) learned sixty words of a novel language through trial-and-error (see Figure 1).
- Trial-by-trial analyses of behaviour and neural signals were conducted and compared between simulated and empirical data.
- To determine learning curves of all data, linear, exponential, and power trends were compared using generalized linear mixed effects models.

RESULTS

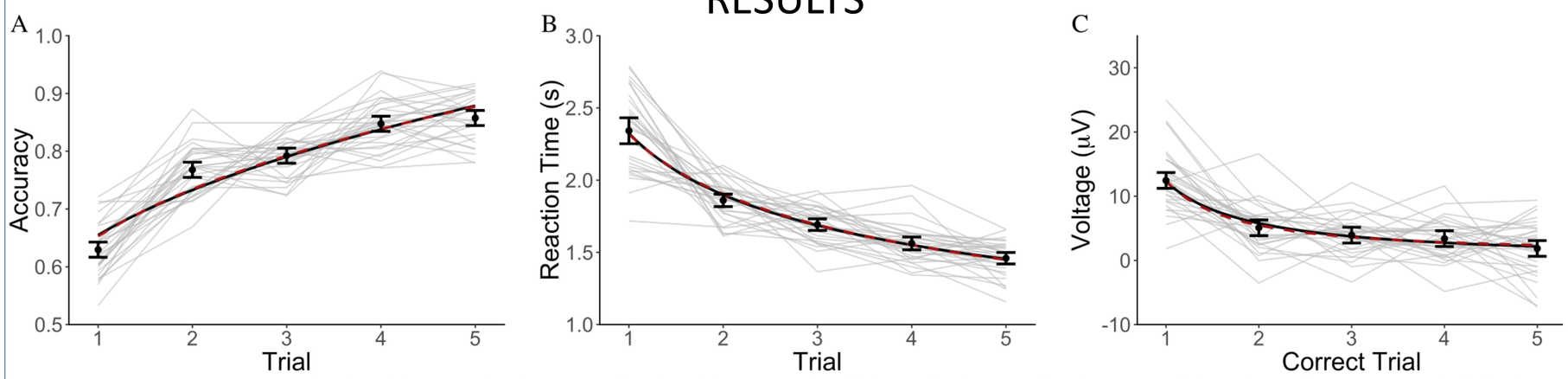

Figure 2. A: accuracy rates across first five trials, B: reaction times across first five trials, C: reward positivity amplitudes across first five correct trials. Each grey line represents an individual participant's data corrected for between subject variability. Black points reflect grand averaged participant data with 95% within-subject credible intervals. Black lines are the best fit power trend line. Red dashed lines are best fit power law functions of reinforcement learning model simulation data.

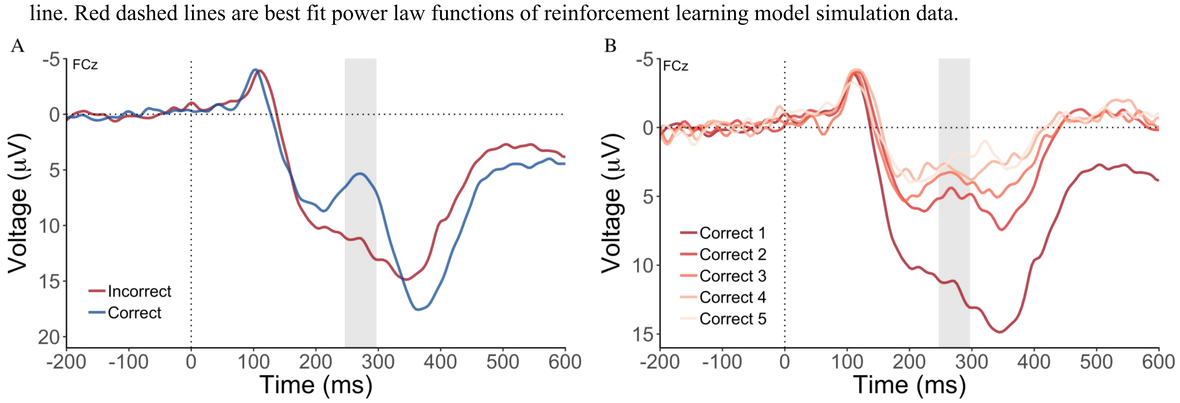

Figure 3. A: Conditional waveforms of incorrect feedback and first correct trial feedback. B: Conditional correct feedback waveforms across correct trials. Grey bar indicates range in which mean peaks were computed.

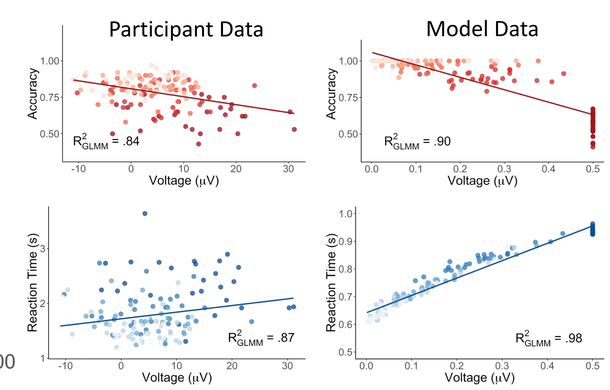

Figure 4. Each participant and simulation contributed data points for each of the first five trials. Colors scale to trial number.

CONCLUSIONS

Reinforcement learning model predictions suggest that behavioural and neural measures both adhere to power learning curves (see Figure 2).

Indeed, in line with these predictions, learning-related behavioural and neural changes were all best fit by power learning curves rather than linear or exponential trends (see Figure 2).

This included trial-by-trial changes in the reward positivity (see Figure 3).

Behavioural and neural measures of learning were strongly associated in both theoretical predictions and empirical findings (see Figure 4).

1. Newell, A., & Rosenbloom, P. S. (1981). *Cognitive skills and their acquisition.* **2.** FitzGerald, T. H., Dolan, R. J., & Friston, K. (2015). *Frontiers in computational neuroscience.* **3.** Schultz, W., Dayan, P., & Montague, P. R. (1997). *Science*. **4.** Krigolson, O. E. (2018). *International Journal of Psychophysiology*.

www.chadcwilliams.com Download PDF of poster ->
@WilliamsNeuro Find supplemental material ->

Chad C. Williams, Theoretical and Applied Neuroscience Laboratory, University of Victoria, www.krigolsonlab.com, ccwillia@uvic.ca

NSERC CRSNG