MSc in Computational Cognitive Neuroscience

Goldsmiths
UNIVERSITY OF LONDON

Computational Models for Reward Based Motor Learning

Osama Shah. Supervisor: Maria Herrojo Ruiz[1]

Background

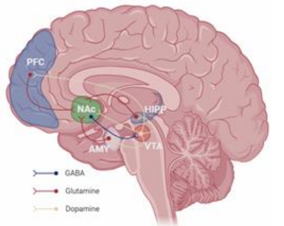


Fig 1: The areas of the brain involved in reward based learning.

Learning a novel motor skill, such as playing a musical instrument, requires the precise coordination of the motor control areas of the brain with the pathways involved in decision making & learning.

Previous neuroscientific studies[2] have implicated the prefrontal cortex, the brain's cognitive decision-making area[3], indicating that motor skill learning is a decision-making process. Other studies[4] have shown that the rate of learning is sensitive to extrinsic reward, highlighting a role for the brain's reward neurotransmitter, dopamine, and the basal ganglia structures involved with its regulation (fig 1).

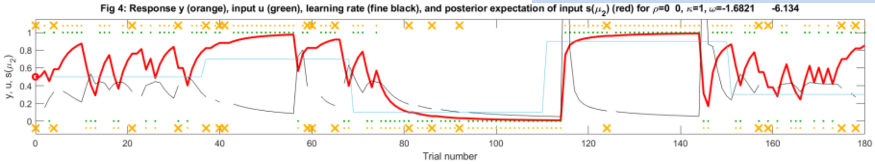
Research Goals

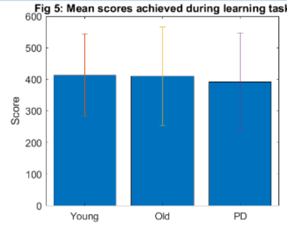
This study attempts to determine how reward prediction modulates the rate of motor-based learning. Computational models based on a Hierarchical Bayesian model, the Hierarchical Gaussian Filter (HGF)[5], would be fitted to the results of a motor decision task, and compared to other reinforcement learning models to estimate how well reward-based prediction errors serves as a predictor for motor learning rates.

Hypotheses

- Applying the HGF framework to the participant's Win data will output a set of per-round precision-weighted prediction errors, that serves as a metric measuring that participant's learning rate. An example output is shown in fig 4, with black indicating learning rate (red indicates reward expectations, blue indicates probability contingencies, orange indicating which sequence was played, green indicating the outcome of the round)
- If dopamine plays an important role in motor skill learning, then Parkinson's Disease patients (diagnosed with deficiencies in dopaminergic neurons) should exhibit impaired learning in the online task. Therefore, their prediction errors would be significantly larger than those in the young and old groups.
- The learning rates for all participants would be summed and compared between the three groups using linear mixed models. It is expected that the learning rates for the young group will be significantly higher than those in the PD group, and less significantly higher than the old group.
- The original and extended versions of the HGF framework will be tested, and compared with other reinforcement learning models (e.g. Rescorla-Wagner). Additional insights may be obtained by analysing timing data and accuracy of key-presses.

Fig 4: Response y (orange), input u (green), learning rate (fine black), and posterior expectation of input $s(\mu_2)$ (red) for ρ=0 0, κ=1, ω=-1.6821 -6.134





Details of the Task

79 adults took part in an online motor decision task, adapted from the two-armed bandit. The adults consisted of 37 young participants (aged between 18 and 39), 31 older participants (aged 60 and over) and 11 patients from the Neurological Clinic of University Hospital Padua diagnosed with Parkinson's Disease (PD).

The task was performed on a desktop web browser with a computer keyboard and involved participants familiarising themselves with a web-based piano keyboard, mapped to the second row of keys (keys 'a', 's', 'd', and 'f') on the computer keyboard (fig 2). After a practice phase where participants were trained to play 2 four-note sequences (associated with fractals fig 3a and 3b), participants played 180 rounds of a task based on the two-armed bandit.

In each round, the participant was required to play either sequence 1 or sequence 2, with a 5-point reward being hidden behind one of the sequences.

In order to receive the reward, the participants had to choose the correct sequence, and play it correctly. However, the probabilities of a sequence being rewarding was not always 50:50.

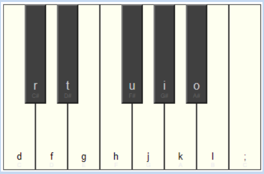


Fig 2: The web based piano keyboard presented to participants

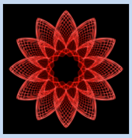


Fig 3a: sequence 1 fractal

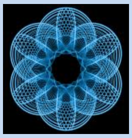


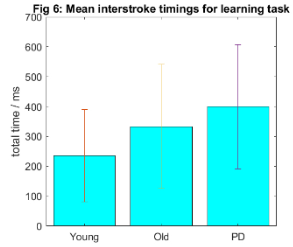
Fig 3b: sequence 2 fractal

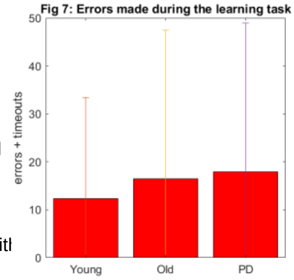
The probabilities were recalculated approximately every 36±6 rounds, with the probability of sequence 1 possessing the reward being as high as 90%, or as low as 10%.

The participant's goal was to maximise their reward based on their perceptions of the current probabilities of a sequence being rewarding. There was a total of 180 rounds for a theoretical maximum reward of 900 points.

Preliminary Results

- PD patients achieved slightly lower scores than other groups (fig 5). The difference is not statistically significant.
- PD patients played sequences, on average slower than other groups, with younger participants playing sequences faster than older participants (fig 6)
- Young participants made fewer errors and experienced fewer timeouts than other groups (fig 7)





Future work

- In order to perform a learning rate analysis, a HGF will need to be applied to the result data.
  This will require tuning of the hyper-parameters to deal with errors found with the default weightings.
- Analysis of the prediction error estimators (AIC, BIC, LME[6]) to estimate the quality of the model
- Further Insights from timing data
- Comparisons by contingency block
- Analysing learning rates from other computational models

[1] The author would like to thank Margherita Tecilla, Peter Holland & Angelo Antonini for their invaluable contributions to this project.

[2] Modulation of neural activity in frontopolar cortex drives reward-based motor learning, M Herrojo Ruiz et al, bioRxiv 2020, https://doi.org/10.1101/2020.05.05.077743
[3] Principles of Neural Science 5th Edition, McGraw Hill Medical, Kandel et al 2013, p403.
[4] Attenuation of dopamine-modulated prefrontal value signals underlies probabilistic reward learning deficits in old age, de Boer et al, eLife 2017;6:e26424, https://doi.org/10.7554/eLife.26424
[5] Uncertainty in perception and the Hierarchical Gaussian Filter, Mathys et al 2014, Front Hum NeuroSci, https://doi.org/10.3389/fnhum.2014.00825
[6] Akaike Information Criteria, Bayesian Information Criteria, Log Model Evidence respectively.