RL model fitting

Beren Millidge

September 24, 2022

Outline

- Fitted RL models to the behavioural data maximizing log likelihood to find paramaters
- Collapsed all data across sessions (performance was similar on all sessions) but not across subjects
- Q learning based models fit very well (predict choice with 98% accuracy)
- Mice do not appear to exhibit significant side biases, or Q learning decay
- But do find substantial differences between free and forced choice and some choice stickiness

Fitting procedure

- Computed Log likelihood for RL model on the behavioural choice data across all sessions
- Fitted parameters of model to maximize log likelihood, picked best of 5 random initializations
- Computed log likelihood only on free choices but still updated model on forced
- Since only 2 ports available in a trial, set Q value of other to large negative value

Q learning accuracy

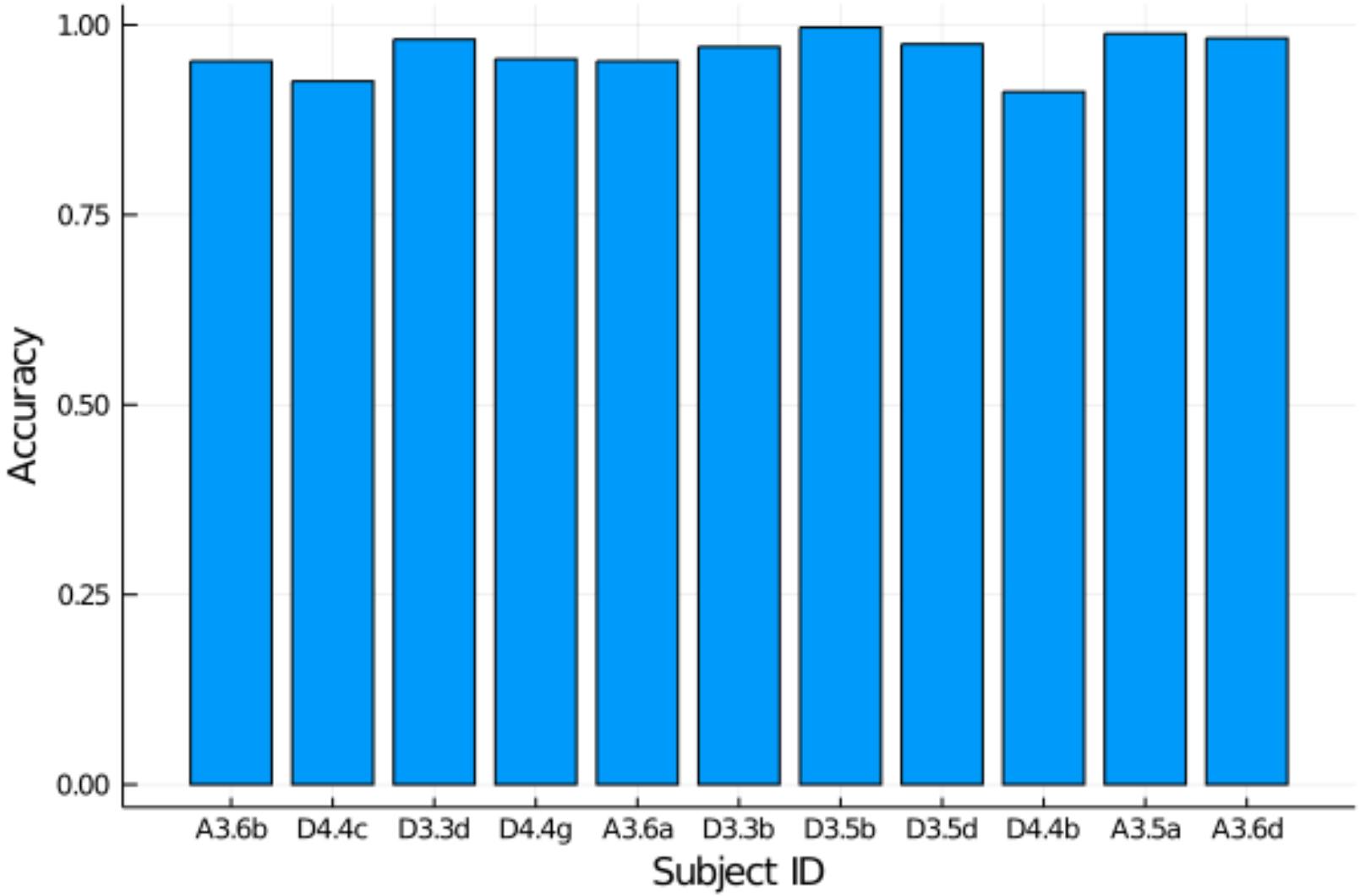

Figure: Q learning models for all subjects reach a high accuracy

Models

- Standard Q learning updates a Q function of 3 values (one for each port) with standard Q learning so Q values are updated $Q(x) += \alpha\delta(x)$ and actions probabilities are computed by $p(a|x) = \sigma(\beta Q(x))$
- Q learning decay introduces an additional decay parameter $Q(x) += \alpha\delta(x) - \gamma Q(x)$ where γ is the decay parameter
- Q learning with port bias introduces 3 bias terms (one for each port) so that action probabilities are $p(a|x) = \sigma(\beta Q(x) + b)$
- Q learning with stickiness adds a bias in action selection towards previous choice $p(a|x) = \sigma(\beta Q(x) + \eta p(a_{t-1}|x_{t-1}))$
- Q learning with forced and free learning rates updates the Q values with different learning rates α_{force} and α_{free} for forced and free choice trials

Log likelihoods

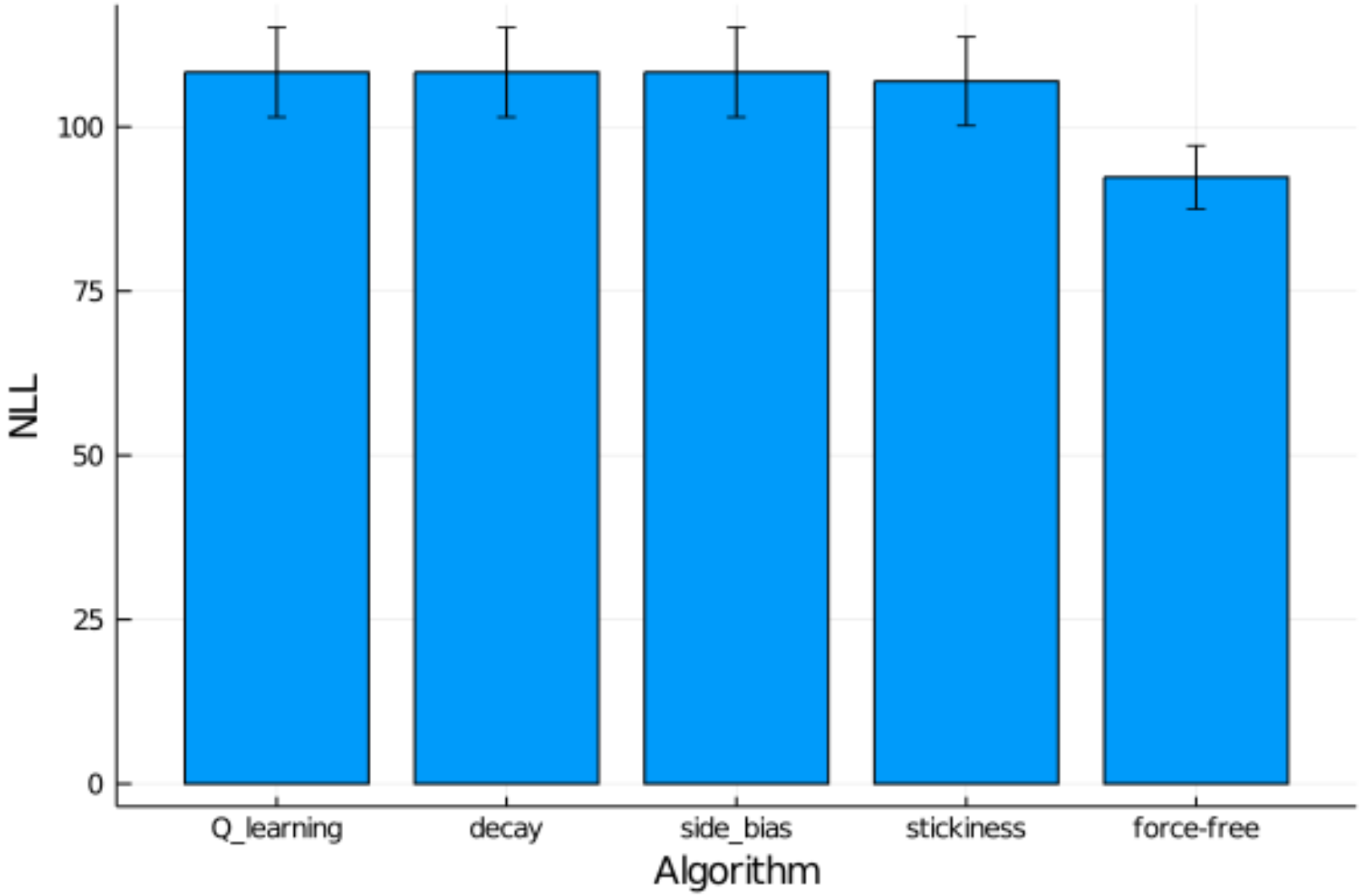

Figure: Most models reach the same log likelihood and are not significant. There is small evidence of choice stickiness and large evidence of differing learning rates

Next steps

- Fit Go/No-go models (I think this will also have very high accuracy)
- Regress Q values and Go and No Go values against the neural photometry data
- Comparison of fits in D1 and D2 to try to find evidence for the non-monotonicity predicted