Computational mechanisms of structure learning: how humans update relational knowledge

UNIVERSITY OF OXFORD

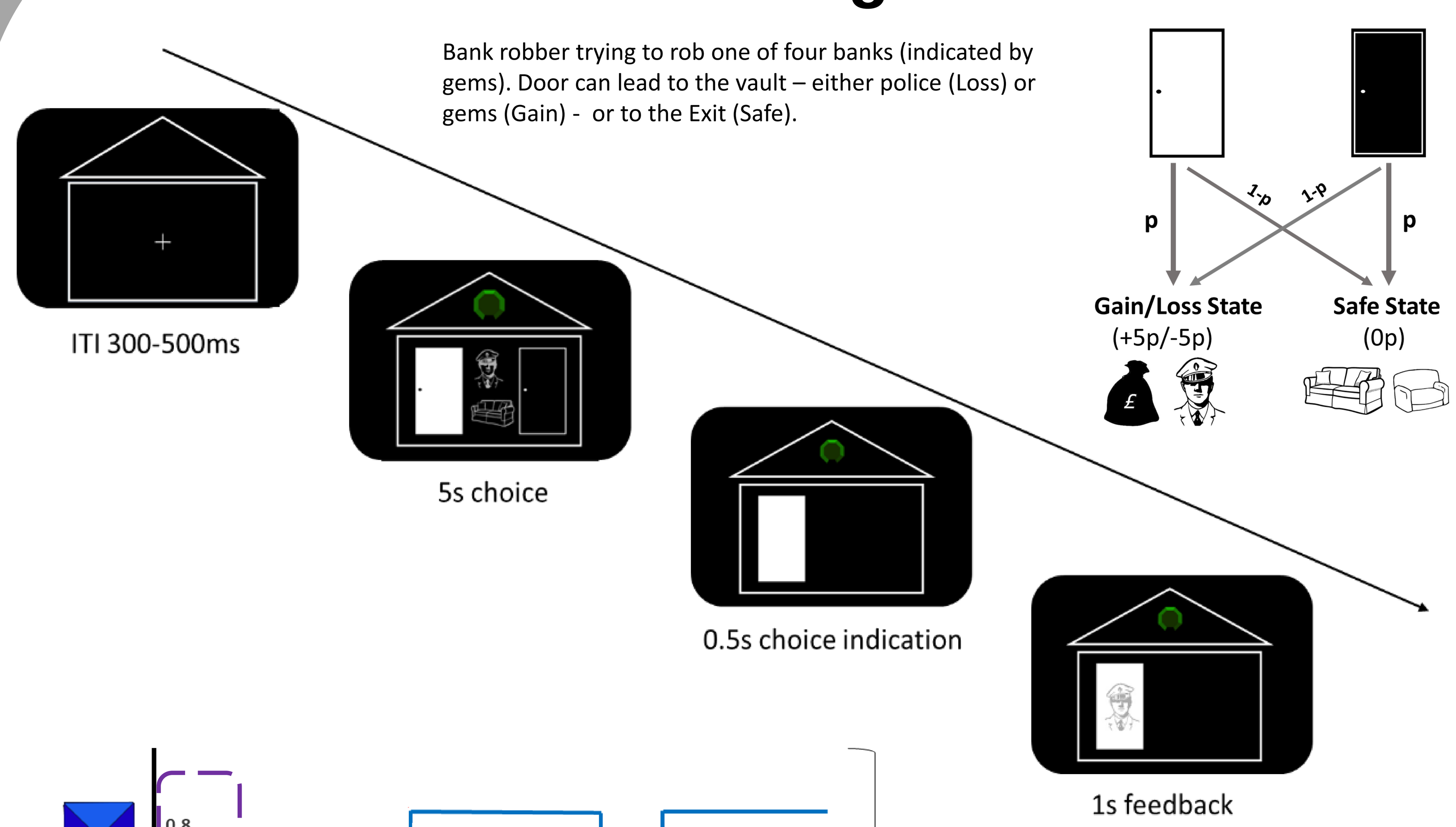
[1]Leonie Glitz, [1,2]Neil Garrett, [1]Keno Jüchems & [1,3]Christopher Summerfield

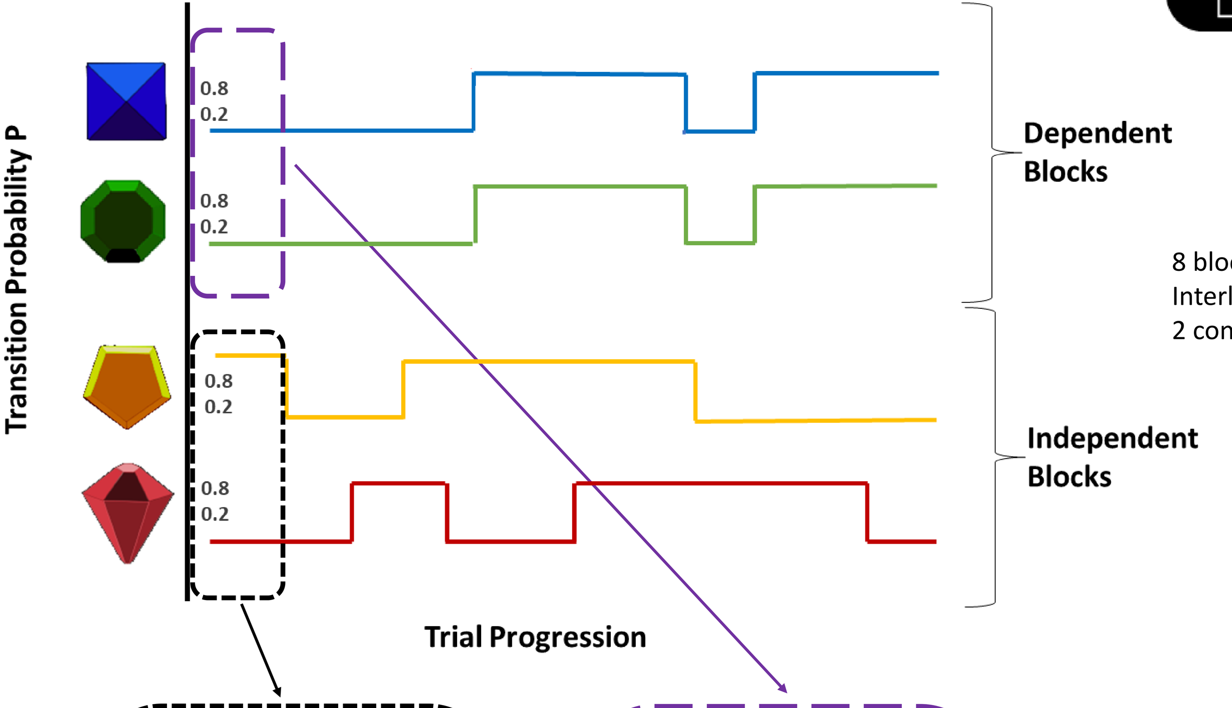
[1] Department of Experimental Psychology, University of Oxford, Oxford, UK [2] Neuroscience Institute, Princeton University, Princeton, NJ, USA [3] Google DeepMind, London, UK

Introduction

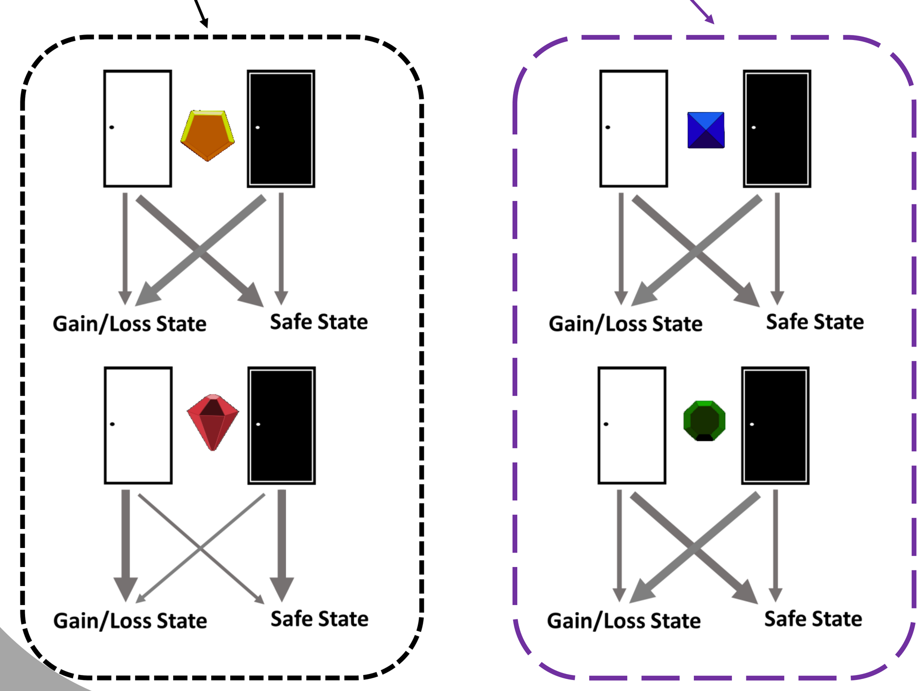
- To make inferences about the world, we need a model of it that includes transition probabilities between different states
- Having and updating a unique model for each context is computationally expensive
- **Can we learn to reuse models for different contexts when appropriate but not when inappropriate?**
- Current tasks investigating model-based & model-free learning are solvable using either of the two learning systems
- Here we use a novel decision-making task that coerces participants to use model-based learning of transition structures across four different contexts
- This allows us to test whether participants generalise new information about state transitions learnt in one context to other contexts when appropriate

Task Design





8 blocks of 60 trials
Interleaved dependent/independent blocks
2 contexts per block



2 trial types:

Gain Trial

Loss Trial

Transition probability **p** either **0.8 or 0.2**; **0.1 chance of probability reversal**;

Participants learn about **p** on every trial ➡ should use estimate of p to make optimal choices depending on trial type

Computational Models

Model Free Learner

$Q^{MF}_{t+1}(a, t) = (1-\alpha) * Q^{MF}_{t} + r_t$

Model Based Learner

$T(s, a, s') = T(s, a, s') + \alpha * \delta_{SPE}$

$Q^{MB}_{t}(s, a) = \sum_{s'} T(s, a, s') * r(s')$

T(s, a, s'): Transition matrix, stores the probability of transitioning from state s to s' given action a.

δ_{SPE} : state prediction error

Basic Model
single learning rate: α

Learning Asymmetry Model
two learning rates: α⁺, α⁻

Structure Learning Model
Q^{MB} from both states weighted by trial-by-trial covariance between state transitions

Combined Model
Integrates structure learning and learning asymmetry model; Q^{MB} & α⁺, α⁻

Choice

$$p(black) = \frac{\exp(\beta^{MF} * Q^{MF}_{Black} + \beta^{MB} * Q^{MB}_{Black})}{\exp(\beta^{MF} * Q^{MF}_{Black} + \beta^{MB} * Q^{MB}_{Black}) + \exp(\beta^{MF} * Q^{MF}_{White} + \beta^{MB} * Q^{MB}_{White})}$$

β^{MF}, β^{MB}: inverse temperature parameters for Model Free and Model Based value estimates

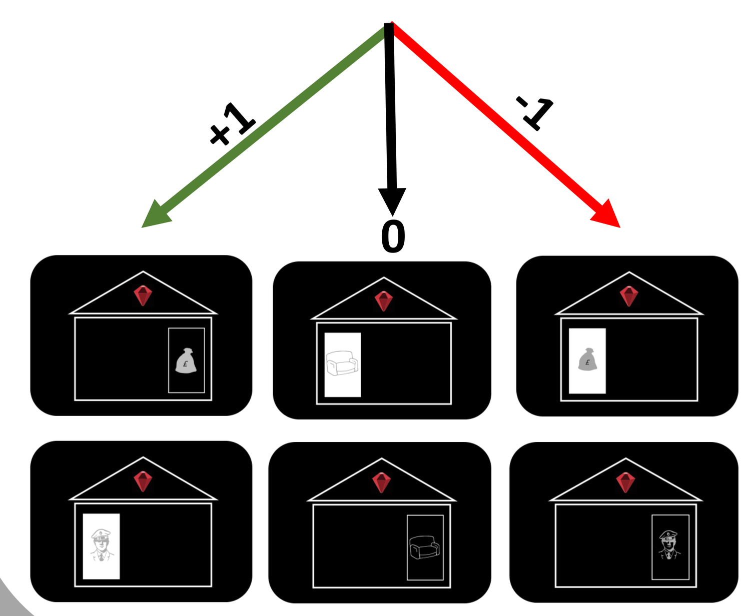
Hierarchical Logistic Regression Model

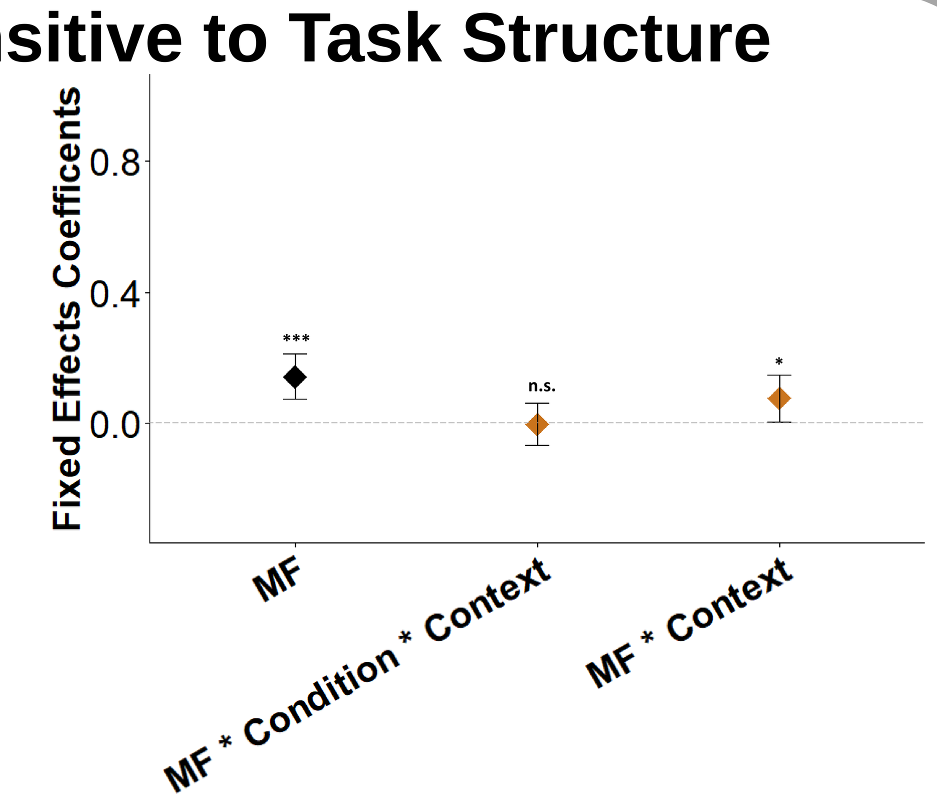
*P(Black Door) ~ MF*context*condition + MB*context*condition + MB*outcome + (1 + MF*context*condition + MB*context*condition + MB*outcome | Subject)*

Model-free Learning Insensitive to Task Structure

Dependent variable: Likelihood of choosing black door

Independent variable: Outcome on the last trial





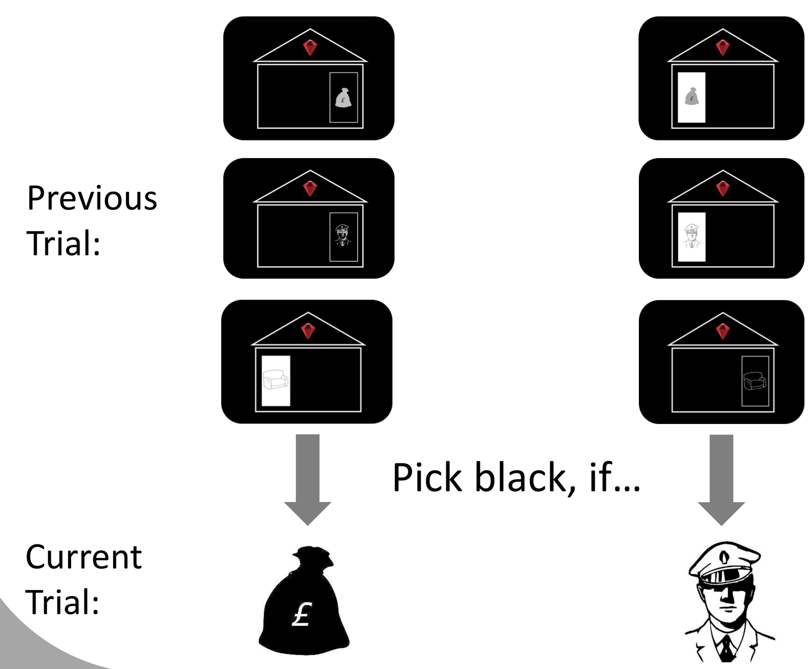
MF: model-free estimate
Condition: dependent/independent block (outcome similarity)
Context: gem presented on the trial

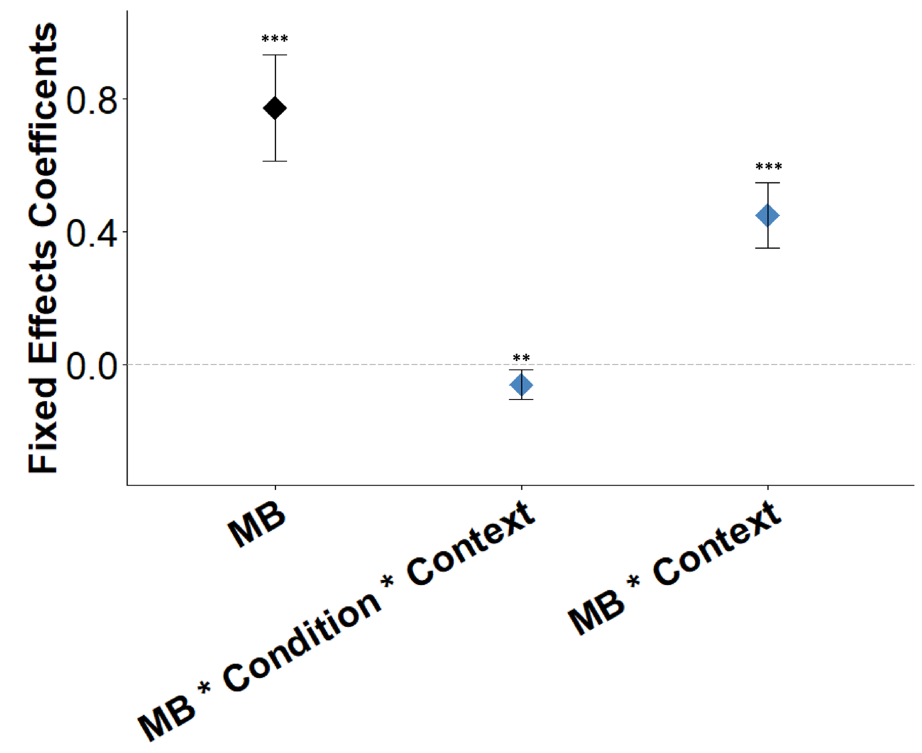
*** $p<.001$, ** $p<.01$, * $p<.05$

Model-based Learning is Sensitive to Task Structure

Dependent variable: Likelihood of choosing black door

Independent variable: State transition on the last trial combined with trial type (gain/loss) of the current trial.

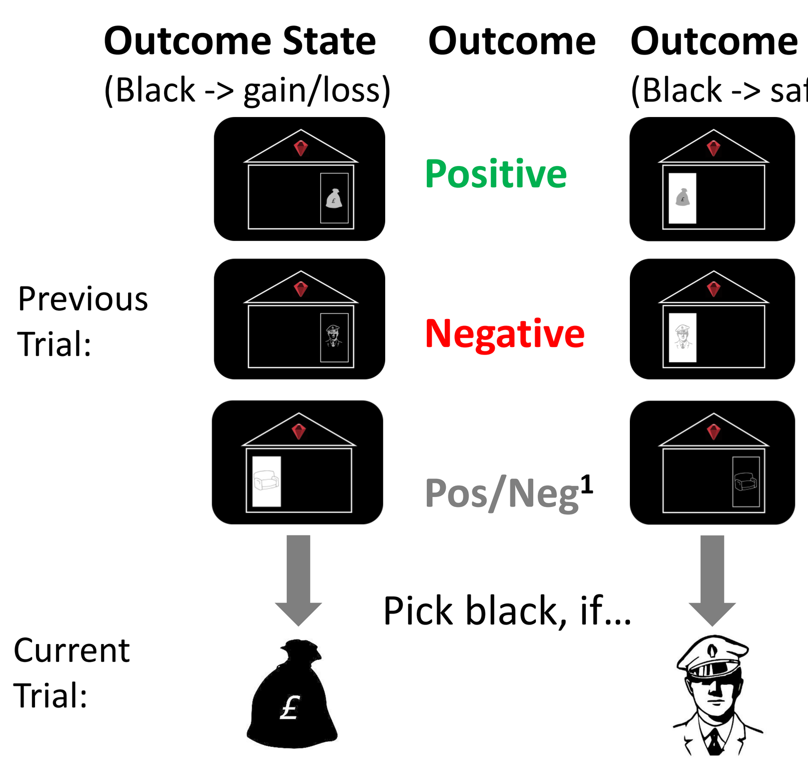


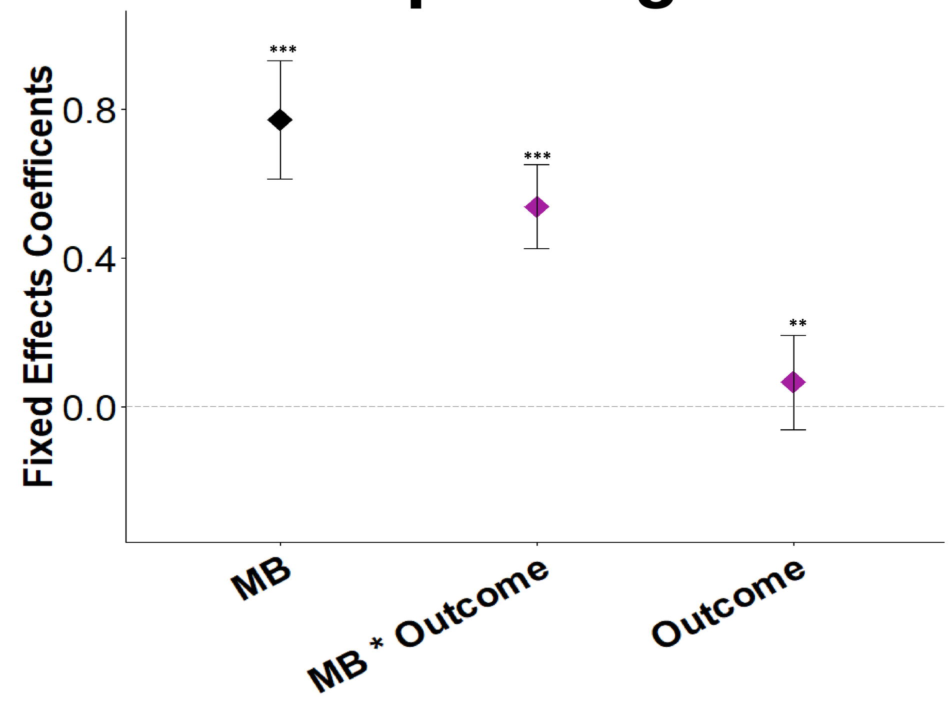


MB: model-based estimate
Condition: dependent/independent block (outcome similarity)
Context: gem presented on the trial

*** $p<.001$, ** $p<.01$, * $p<.05$

Asymmetries in Model-based Updating

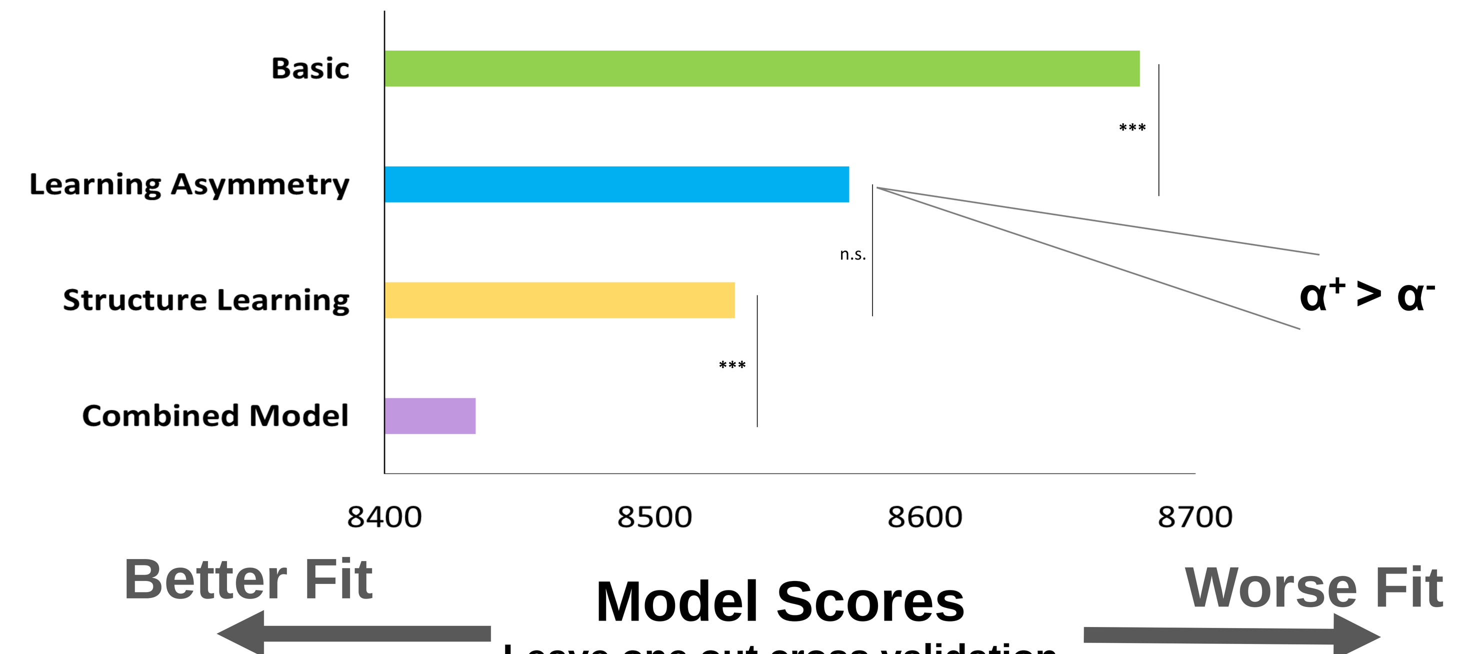




[1]Coded as positive or negative depending on whether previous trial was a loss or gain trial (i.e. the alternative outcome)

*** $p<.001$, ** $p<.01$, * $p<.05$

Structure learning and learning asymmetry a better fit to choice data



*** $p<.001$, paired samples t-test

Conclusion

- Participants integrate information about transition structures across different contexts
- This effect is stronger in environments where different contexts share the same underlying transition dynamics and weaker in environments where the transition dynamics differ between contexts -> **clear evidence of efficient structure learning**
- Valence – whether an outcome is positive or negative – has a strong impact on the rate with which new information about state transitions is integrated into beliefs -> **evidence of asymmetric belief updating**
- Future work will examine how shared versus separate use of state transition knowledge between contexts corresponds to differences in neural representations