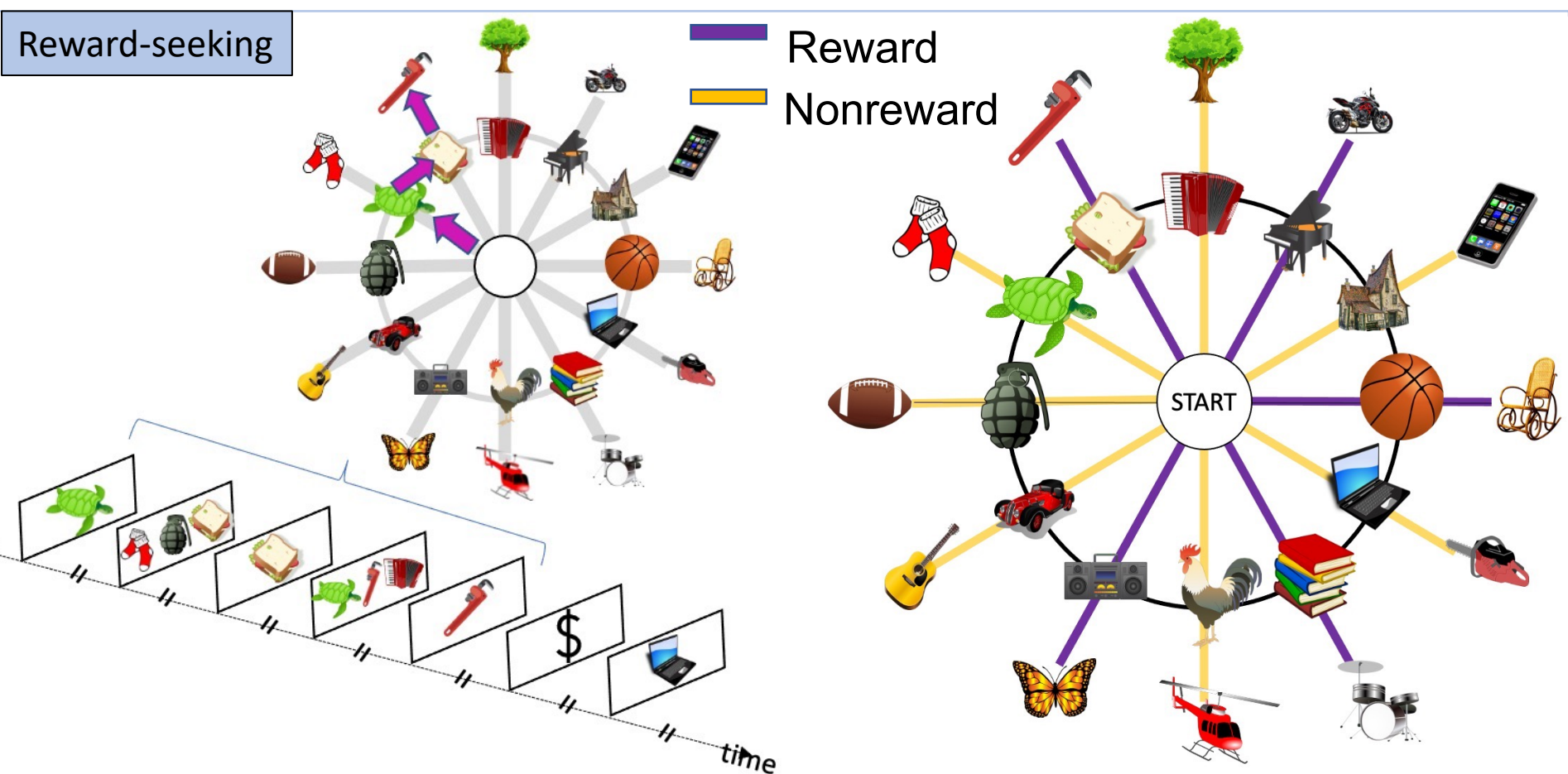


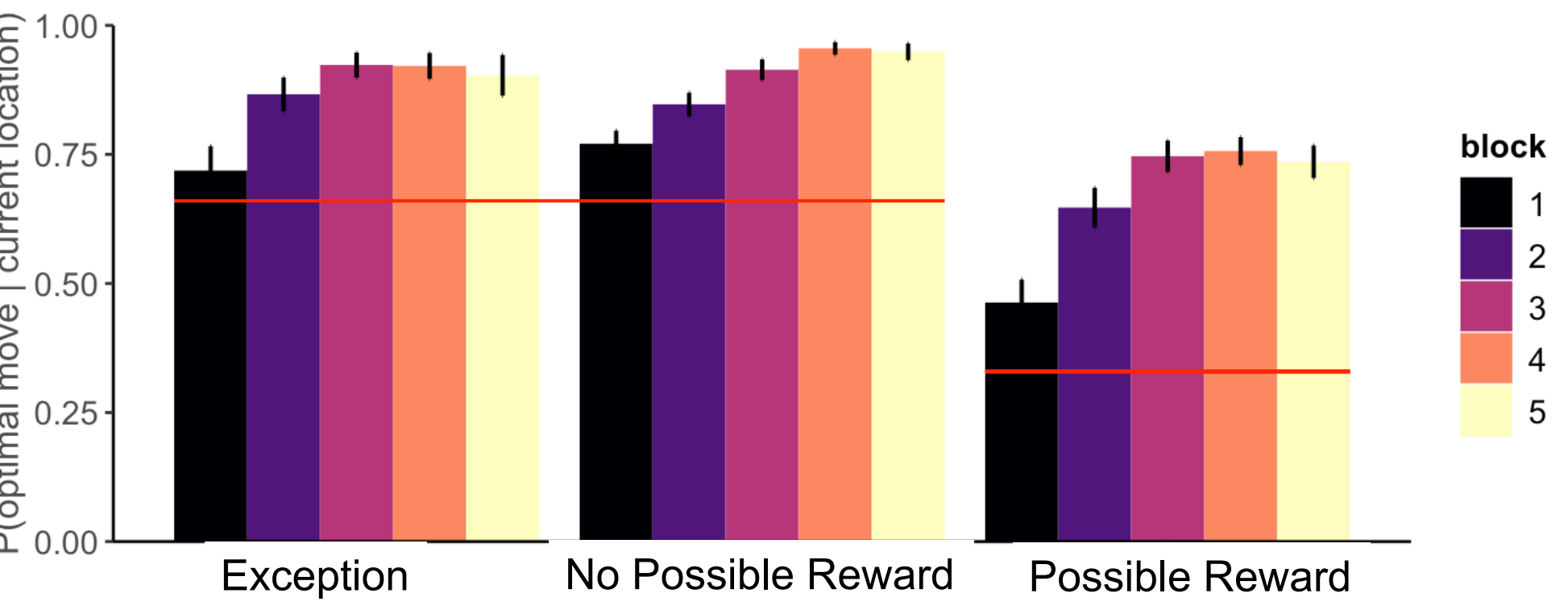
Background:

- Balancing exploration and exploitation is an optimal decision-making heuristic (Gittins, 1979)
- Recently, signs of exploration have been found when people search for reward in both spatial and abstract environments (Wu et al., 2018, 2020)
- Will people seeking reward in an abstract environment display signs of exploration-exploitation tradeoffs?
- Will reward backpropagate to previous abstract locations as it does in animal foraging?

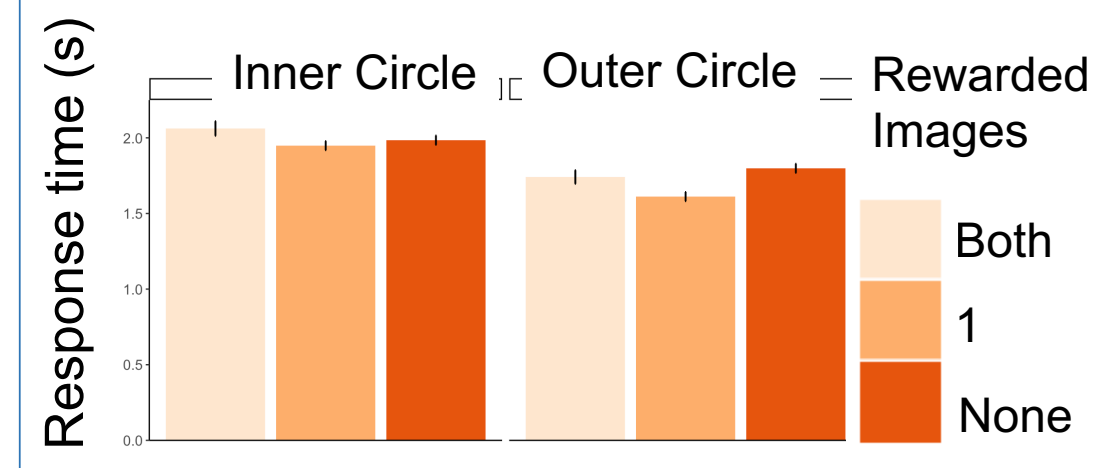
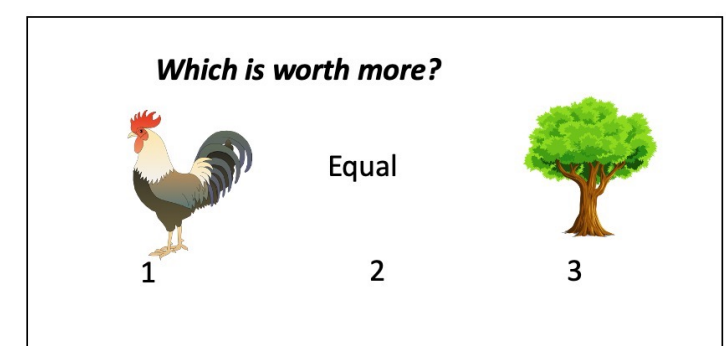
Methods (within-subjects)



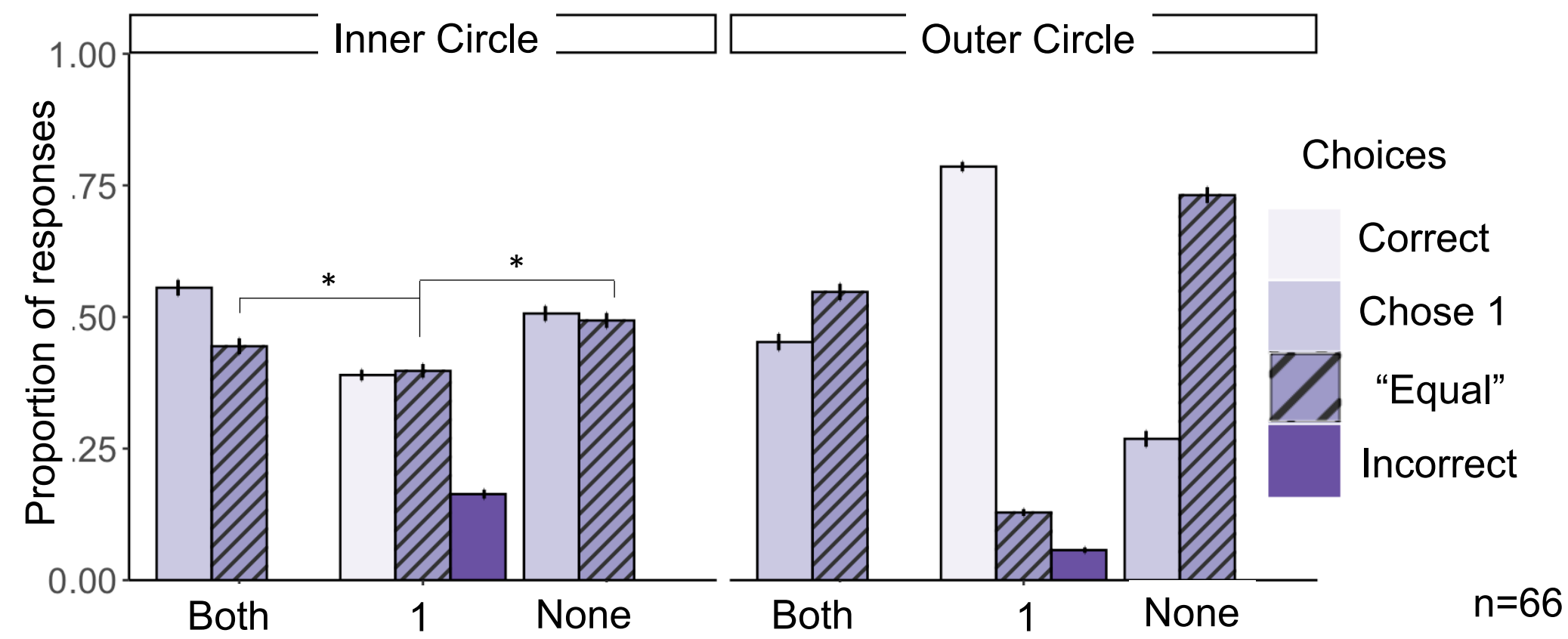
Reward seeking efficiency



Reward memory test



Generalization of reward from outer to inner circle



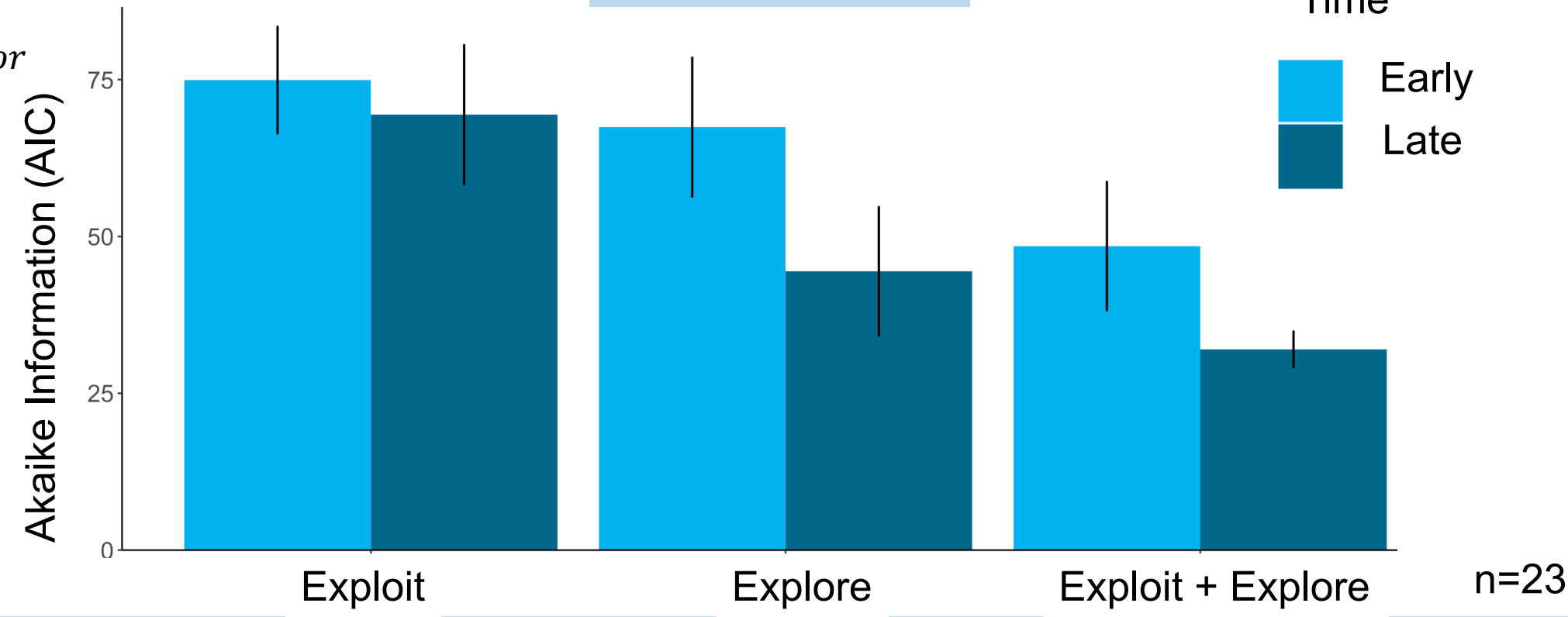
Exploitation-Exploration Models of Behavior

Exploitation = Subjective Value + Prediction Error
 Prediction Error = $\alpha(\gamma R - \text{Subjective Value})$
 α = learning rate
 γ = discount factor
 R = actual reward

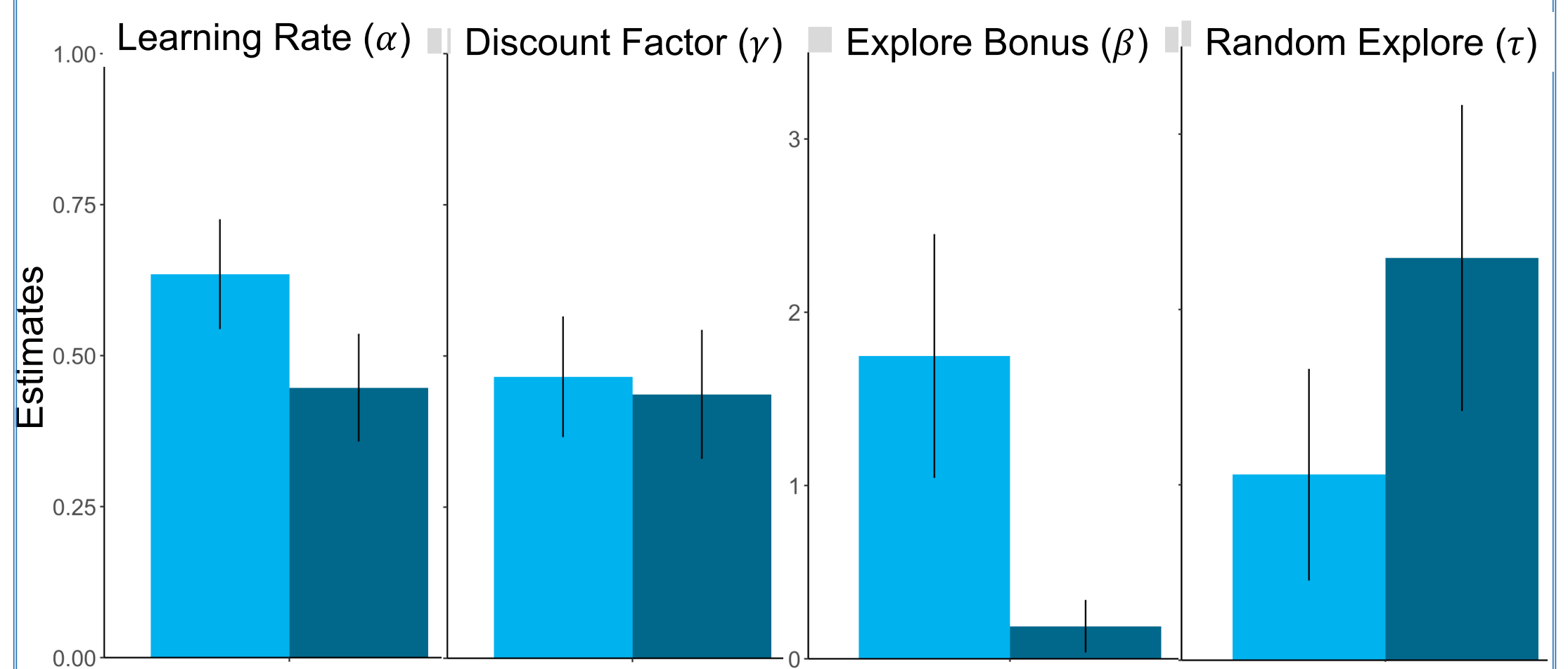
Exploration = $\beta\sqrt{\log(\text{trial\#})/N_i}$
 N = # times current image i was seen

$P(\text{Choice} = k) = \frac{\exp(k/\tau)}{\sum \exp(k_i/\tau)}$
 τ = random exploration

Model Comparison



Parameter Estimates from Explore + Exploit Model



Discussion:

- No evidence for tradeoffs between exploration and exploitation as people navigated an abstract environment
- A combination of Exploitation and Exploration captured observed data better than either factor alone
- Reward did generalize from the rewarded to previous locations in memory

Citations: Gittins, J.C. (1979), J Royal Statistical Soc. B, 41(2), 148-177; Wu, CM et al. (2018) Nature Hum Beh 2, 915-924; Wu, CM et al. (2020) PLOS Comp. Bio., 16(10): e1008384