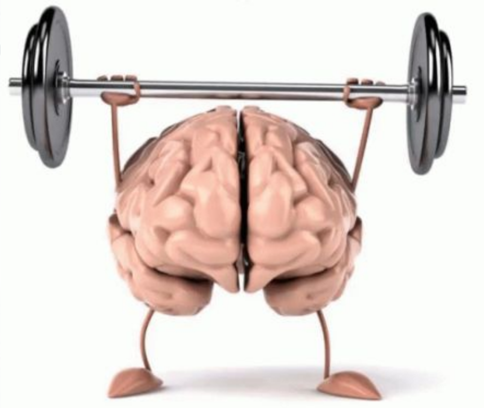
Lesson 3:
Model-based learning in a two-step task

Model-based planning

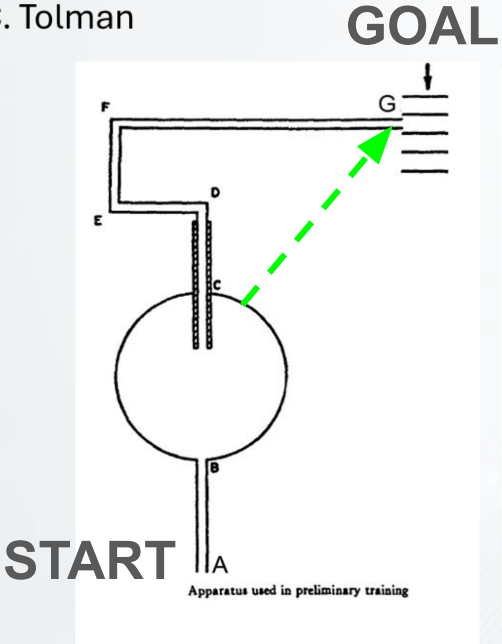
Model-based planning

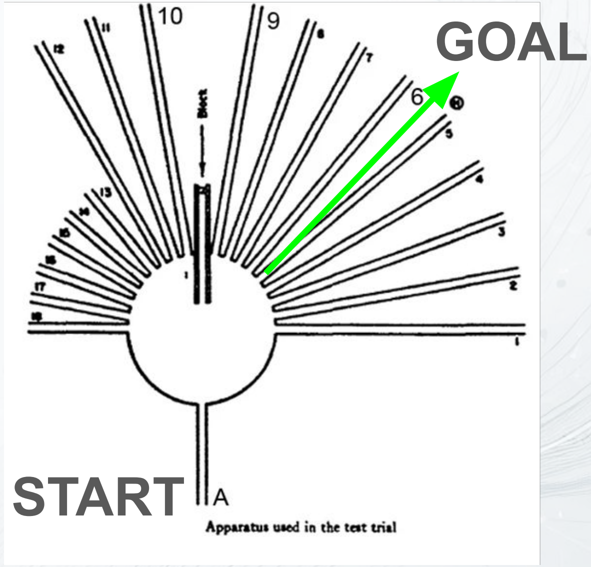
Model-based planning

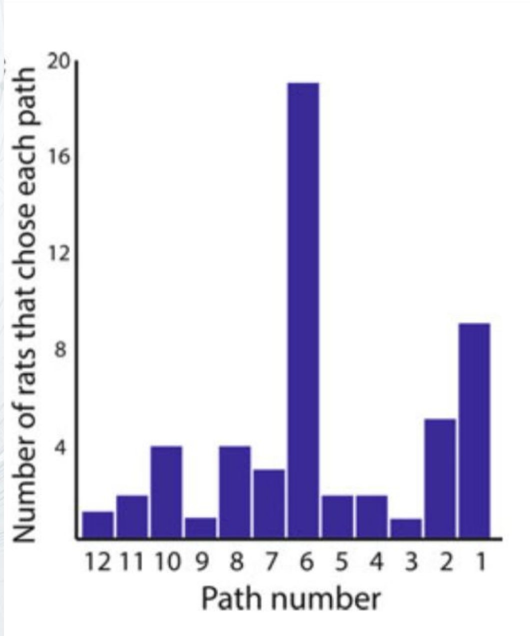


Latent learning

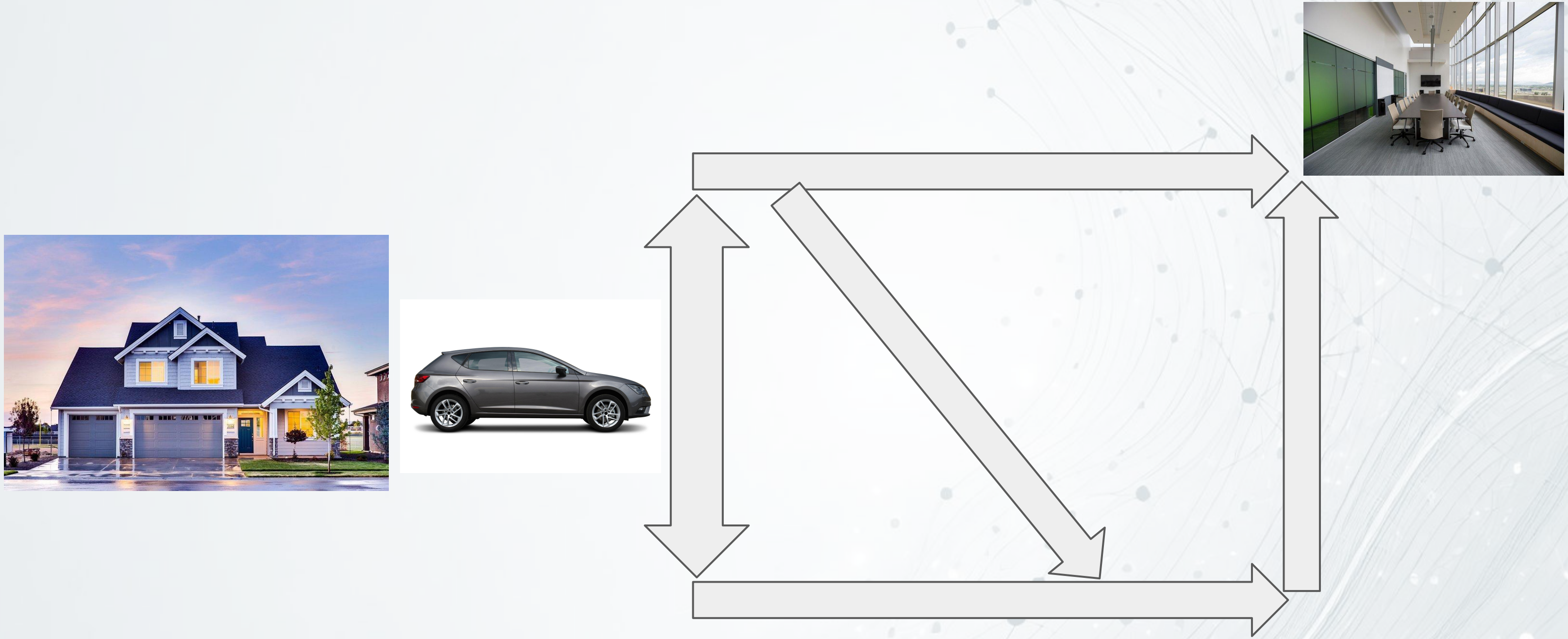
Edward C. Tolman



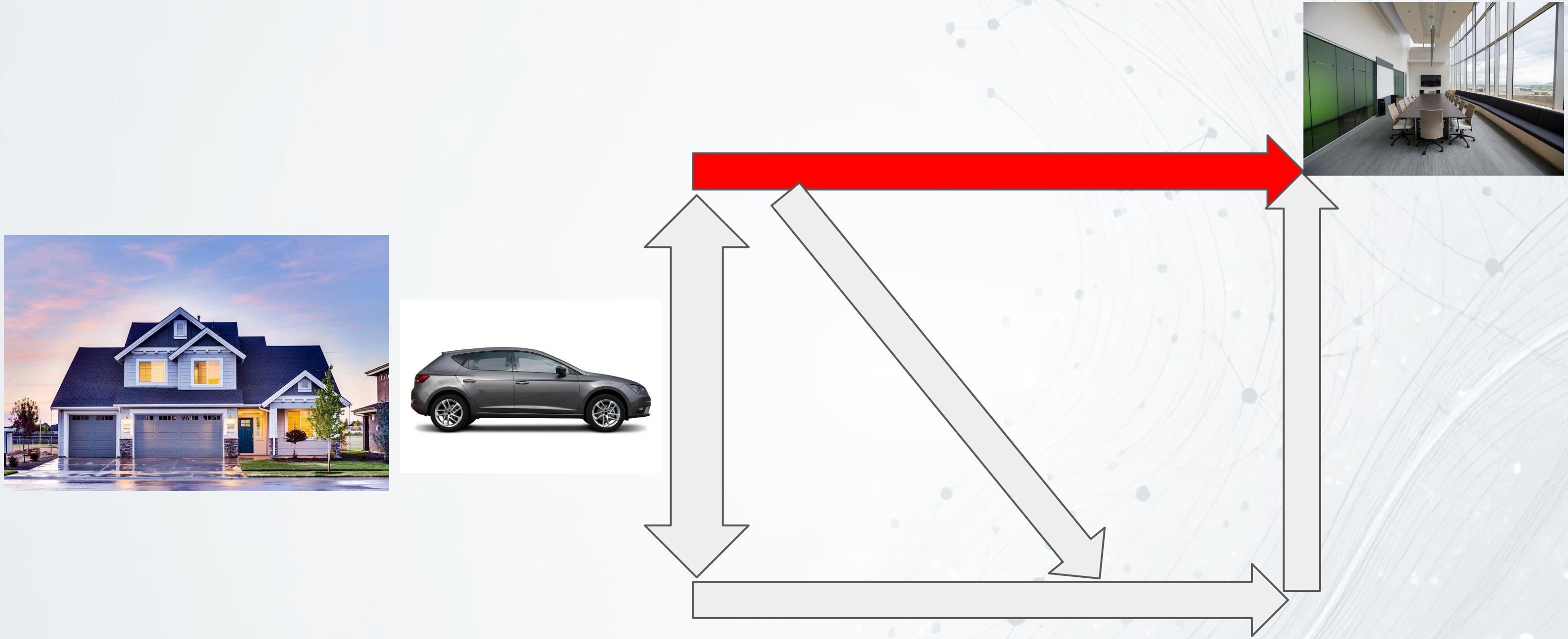




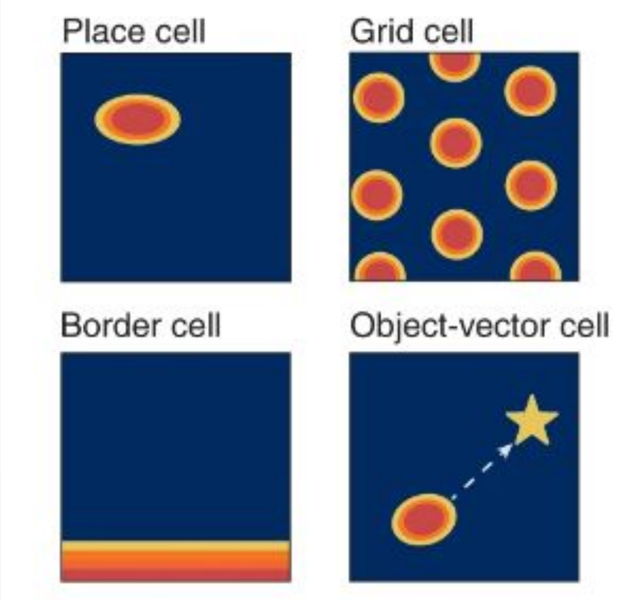
Tolman et al., 1946 I



When the environment changes, model-based planning allows quick adaptation

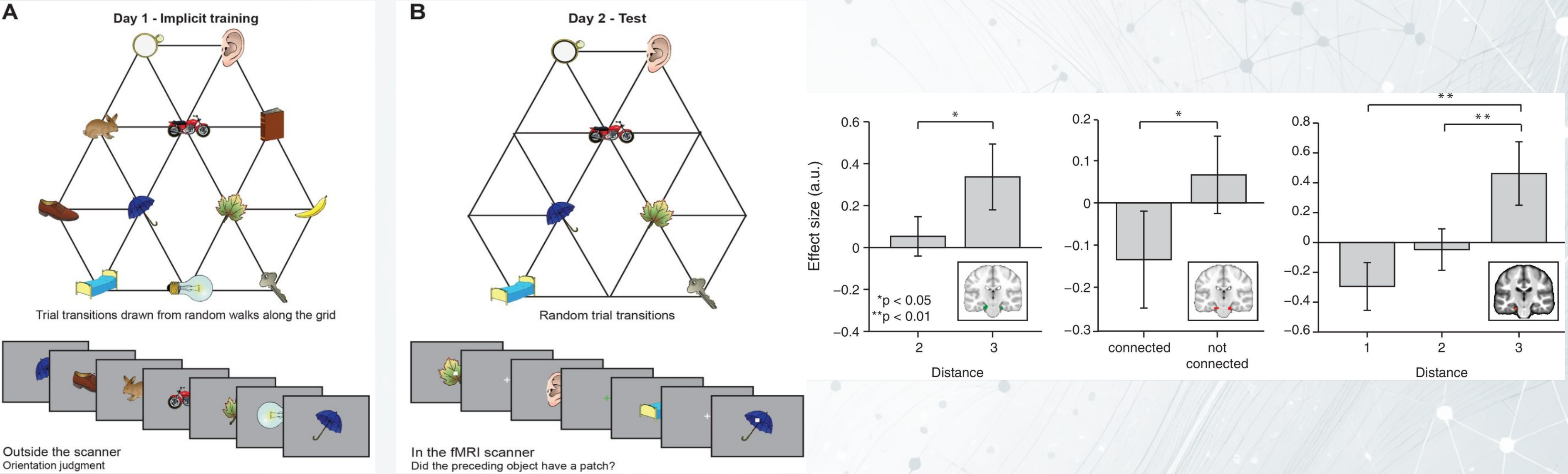


Cognitive maps and the brain



O'Keefe, 1970; Hafting et al 2005; Hoydal et al., 2019; Whittington et al., 2022

Abstract cognitive maps

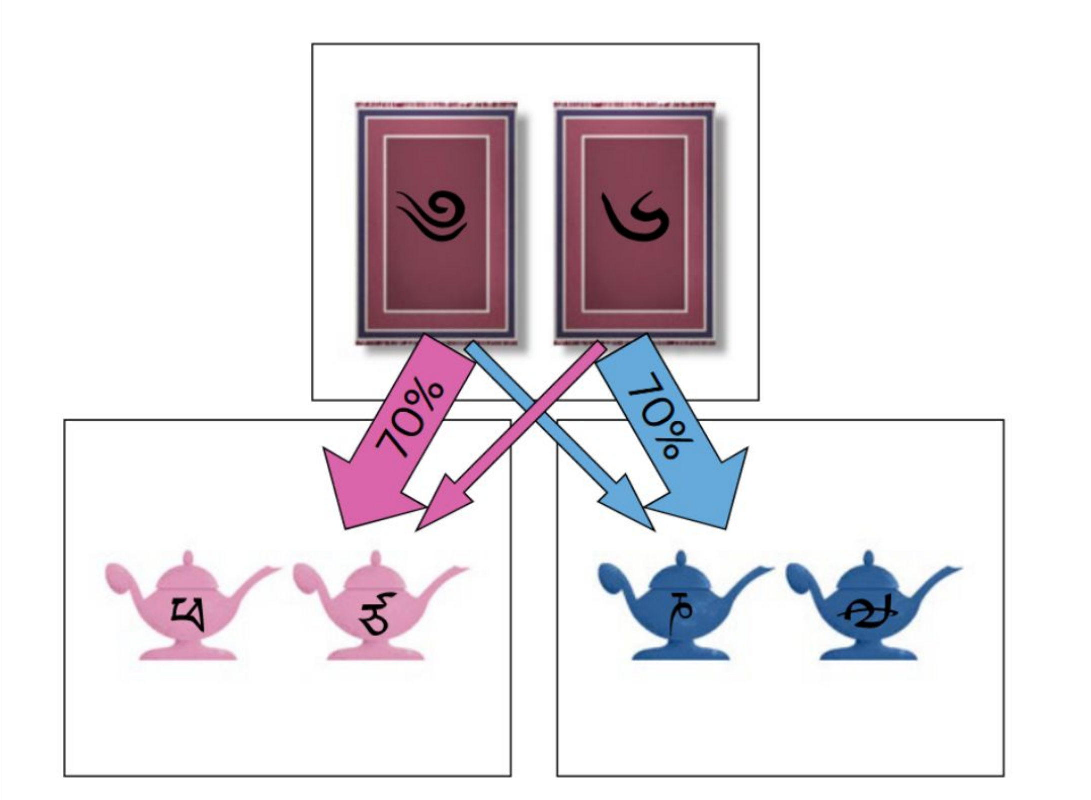


Garvert et al., 2017

Model-based learning

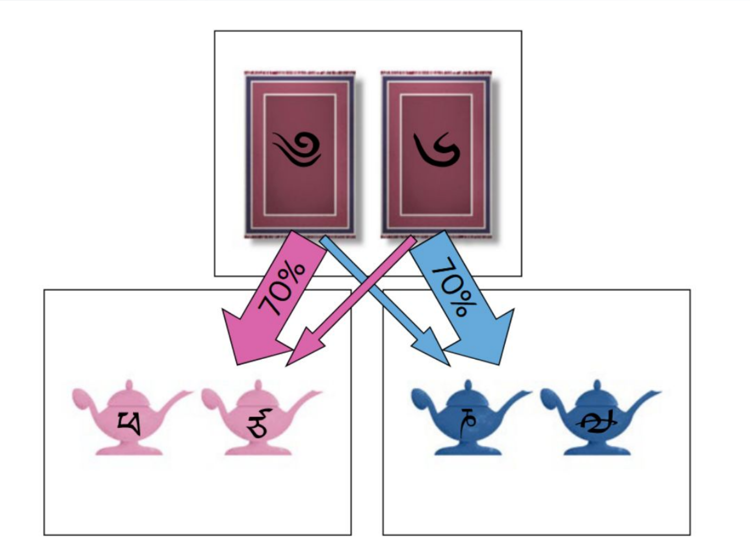
(Two-step task with probabilistic transitions)

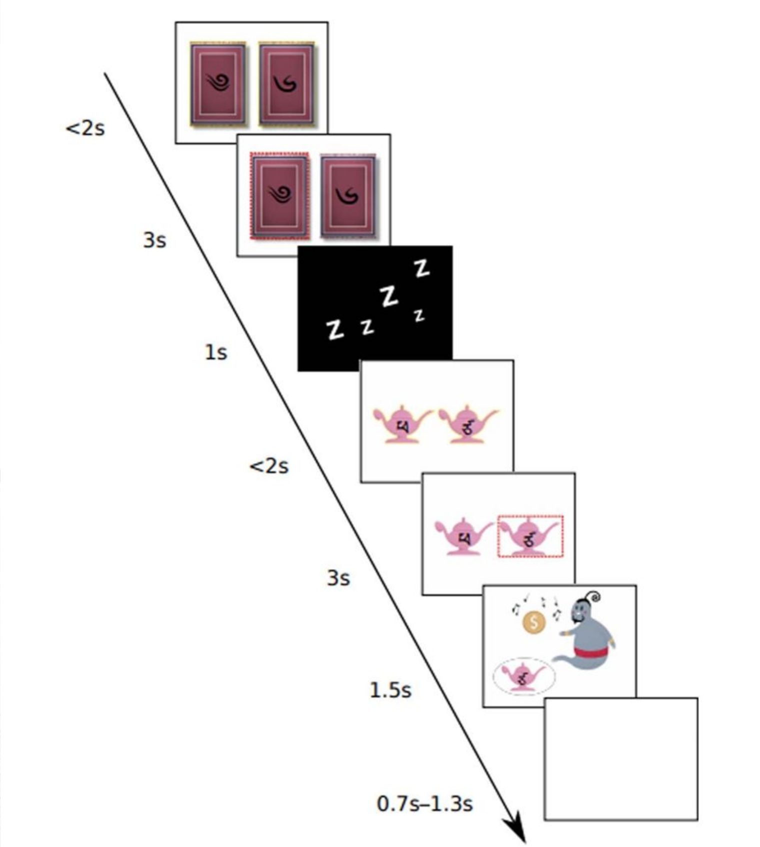
Two-step task



Daw et al., 2005; 2011
Feher da Silva et al. 2020; 2023

Two-step task





Daw et al., 2005; 2011
Feher da Silva et al. 2020; 2023

Hybrid model

Parameters

Learning rate - α

Inverse temperature - □

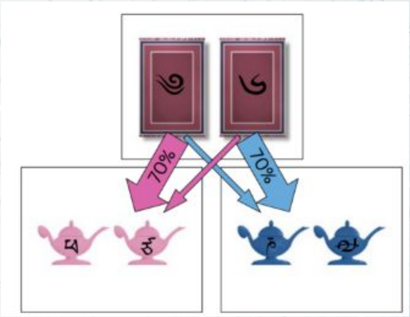
Eligibility factor - λ

Model based weighting - ω

Hybrid model

A

B/C



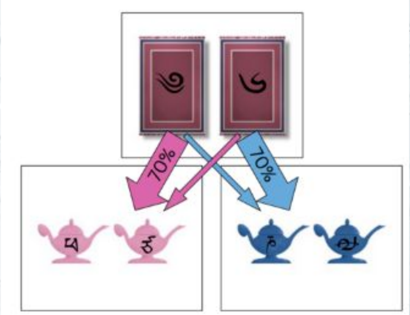
Model

$$State = \{A, B, C\}$$

$$Action = \{1, 2\}$$

Hybrid model

A



B/C

Model

$$State = \{A, B, C\}$$

$$Action = \{1, 2\}$$

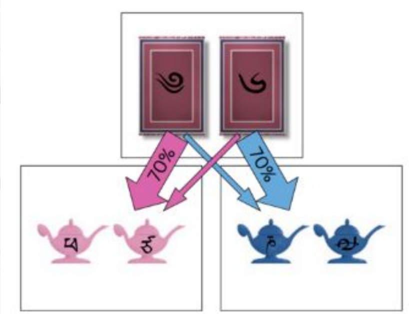
$$Q^{MF}(s, a)$$

$$Q^{MB}(s, a)$$

Hybrid model

A

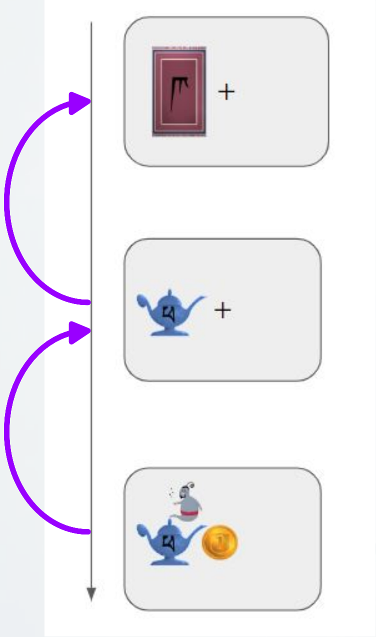
B/C



Model-free action value updating

PE$_1$

PE$_2$



$$PE_2 = r_t - Q(s_2, a_1)_t$$

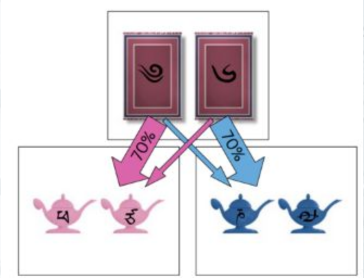
$$PE_1 = Q^{MF}(s_2, a_2)_t - Q^{MF}(s_1, a_1)_t$$

$$Q^{MF}(s_1, a_1)_{t+1} = Q^{MF}(s_1, a_1)_t + \alpha \cdot PE_1 + \alpha \cdot \lambda \cdot PE_2$$

$$Q^{MF}(s_2, a_2)_{t+1} = Q^{MF}(s_2, a_2)_t + \alpha \cdot PE_2$$

Hybrid model

A



B/C

MB action value updating

$$Q^{MB}(s_A, a_j) = P(s_B \mid s_A, a_j) \cdot \max_{a \in \{a_A, a_B\}} Q^{MF}(s_B, a_i) + P(s_C \mid s_A, a_j) \cdot \max_{a \in \{a_A, a_B\}} Q^{MF}(s_C, a_i)$$

$$P(s_B \mid s_A, a_1) = 0.7 \qquad P(s_C \mid s_A, a_1) = 0.3$$

$$P(s_B \mid s_A, a_2) = 0.3 \qquad P(s_C \mid s_A, a_2) = 0.7$$

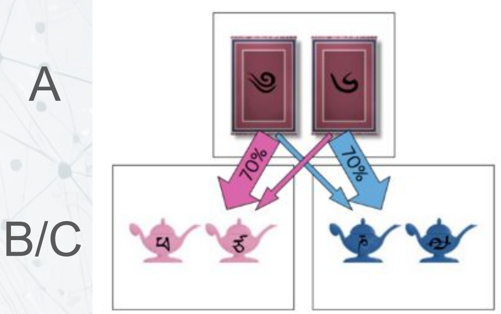
$Q^{MF}(s=B) = [0.5,0.8]$

$Q^{MF}(s=C) = [0.2,0.4]$

$Q^{MB}(s=A,a=1) = 0.7 \cdot \max[0.5,0.8] + 0.3 \cdot \max[0.2,0.4]$

$Q^{MB}(s=A,a=2) = 0.7 \cdot \max[0.2,0.4] + 0.3 \cdot \max[0.5,0.8]$

Hybrid model



Action selection

$$Q^{\text{net}}(a) = \omega \cdot Q^{MB}(s = A, a) + (1 - \omega) \cdot Q^{MF}(s = A, a)$$

$$P(a \mid s = A) = \frac{\exp[\beta \cdot Q_{\text{net}}(a)]}{\sum_{a'} \exp[\beta \cdot Q_{\text{net}}(a')]}$$

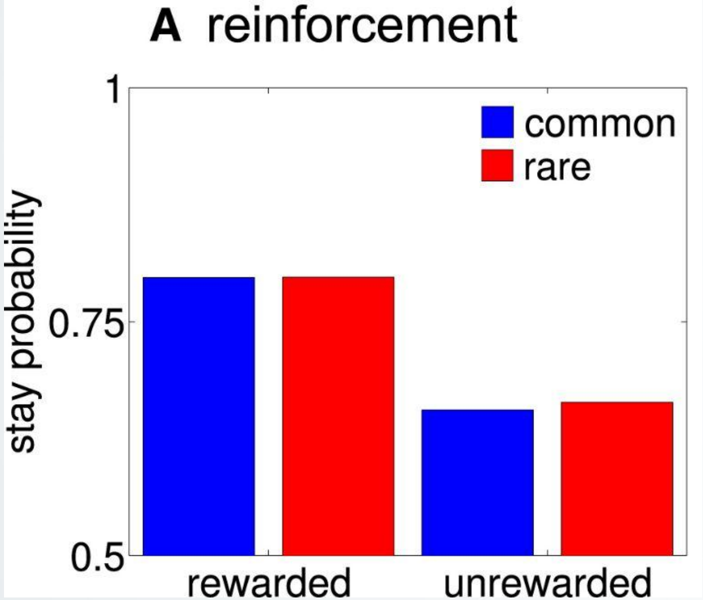
$$P(a \mid s_2) = \frac{\exp[\beta \cdot Q^{\text{MF}}(s_2, a)]}{\sum_{a'} \exp[\beta \cdot Q^{\text{MF}}(s_2, a')]}$$

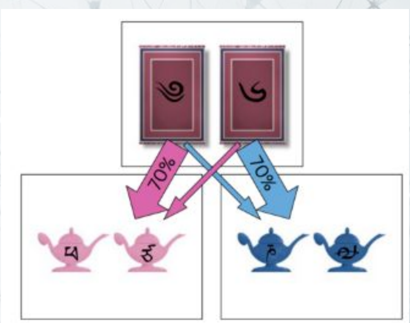
Step-by-step simulation

[google sheet]

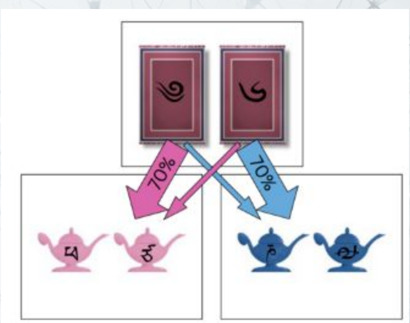
Simulation code in github

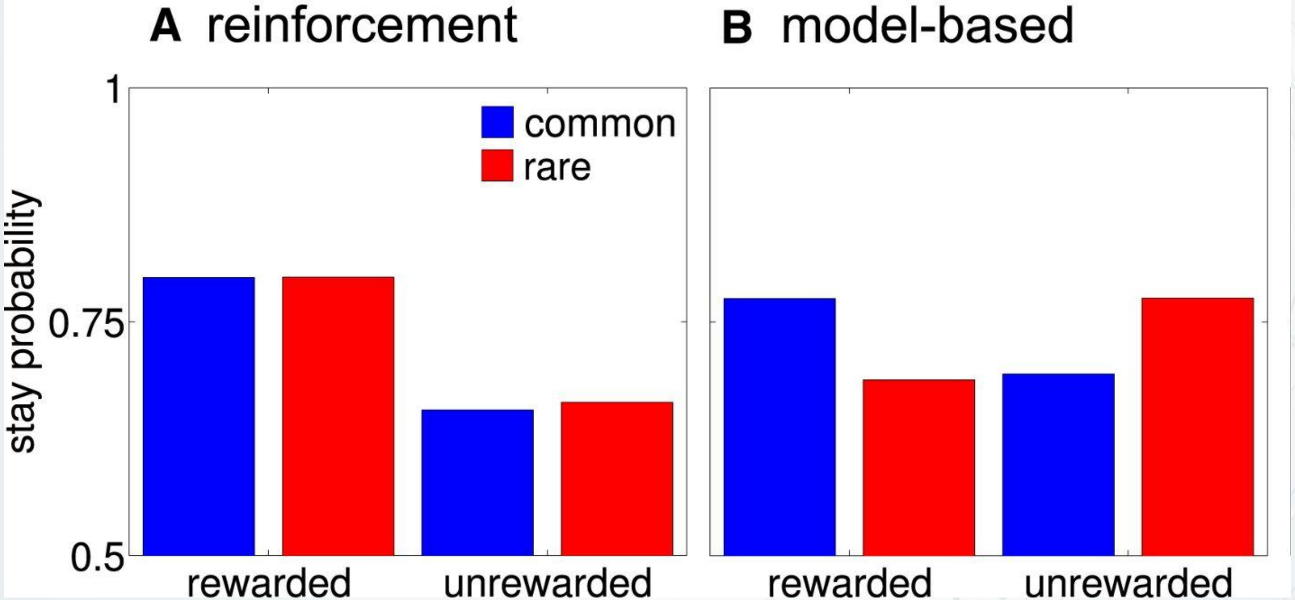
Hard Medium Easy



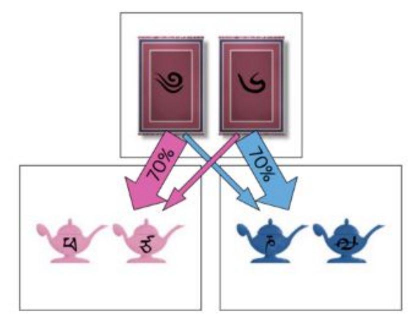


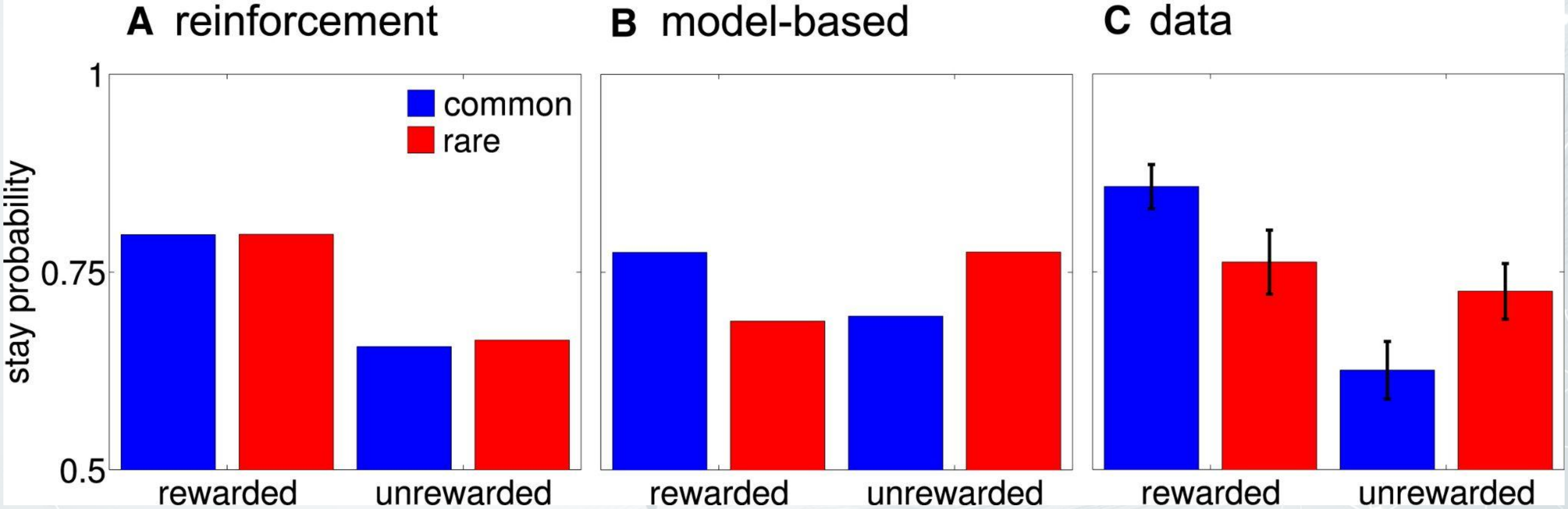
Daw et al., 2005; 2011
Feher da Silva et al. 2020; 2023





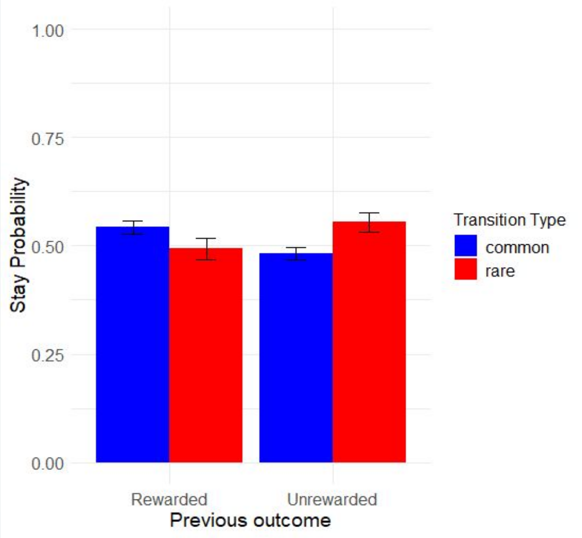
Daw et al., 2005; 2011
Feher da Silva et al. 2020; 2023

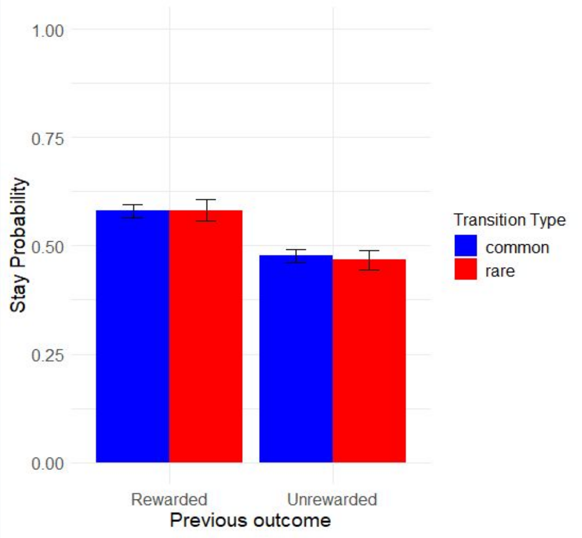


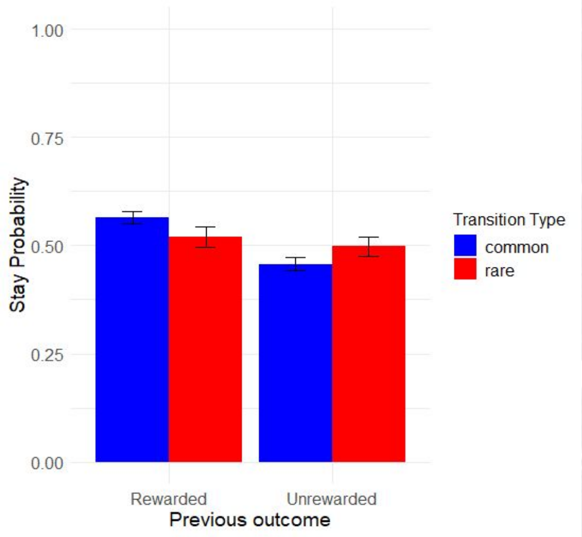


Daw et al., 2005; 2011
Feher da Silva et al. 2020; 2023

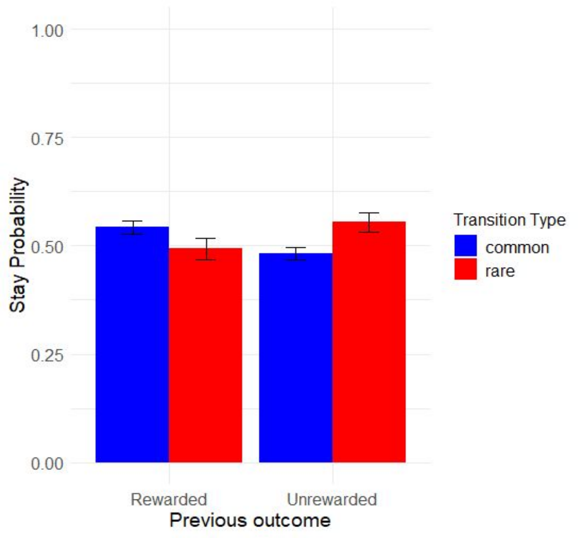
Where’s omega=0?

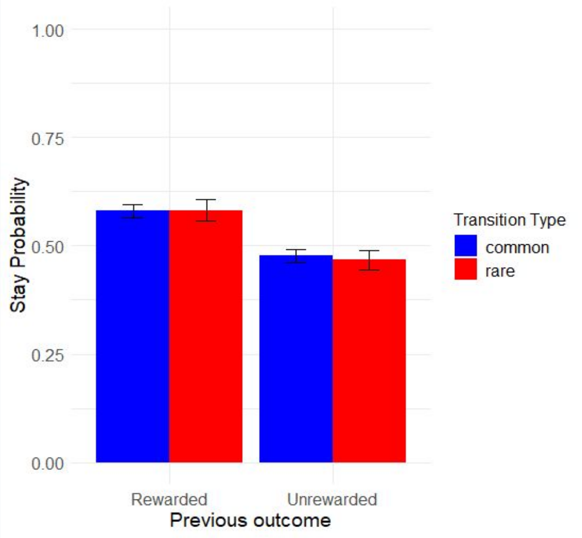


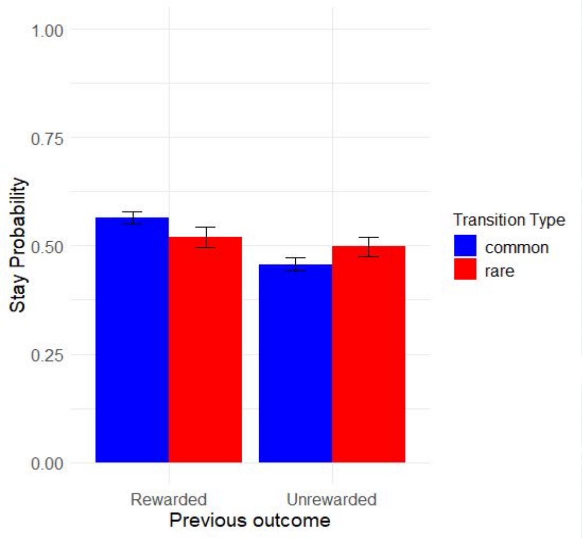




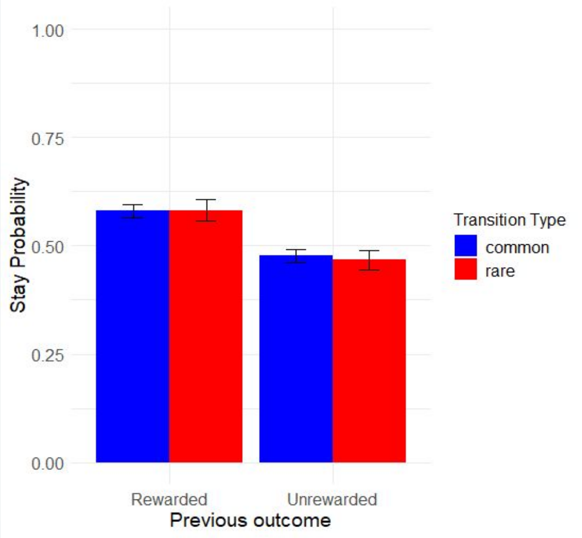
Where's omega=1?



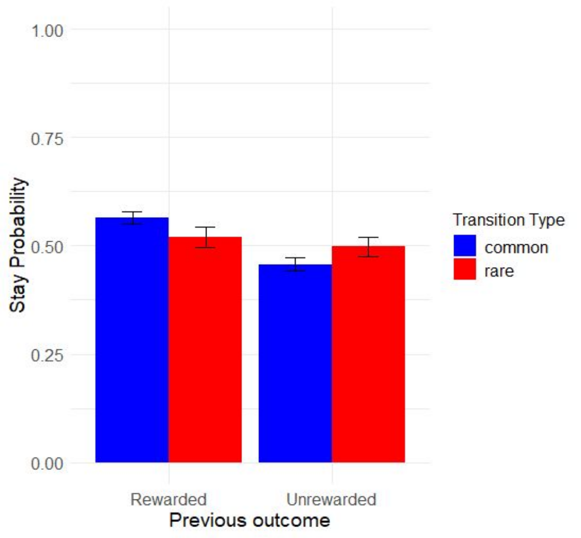




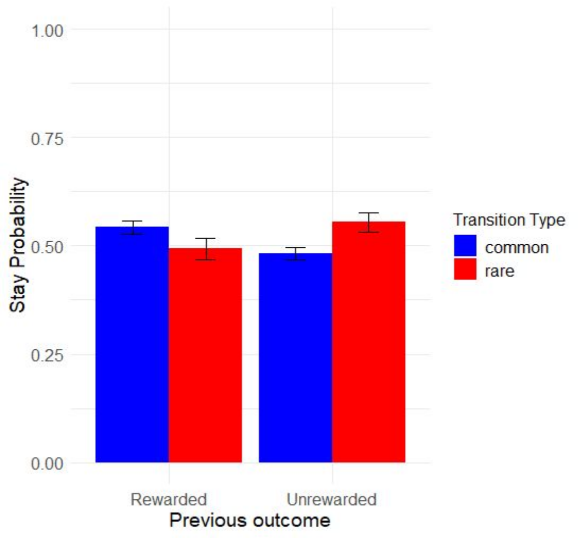
$\omega=0$



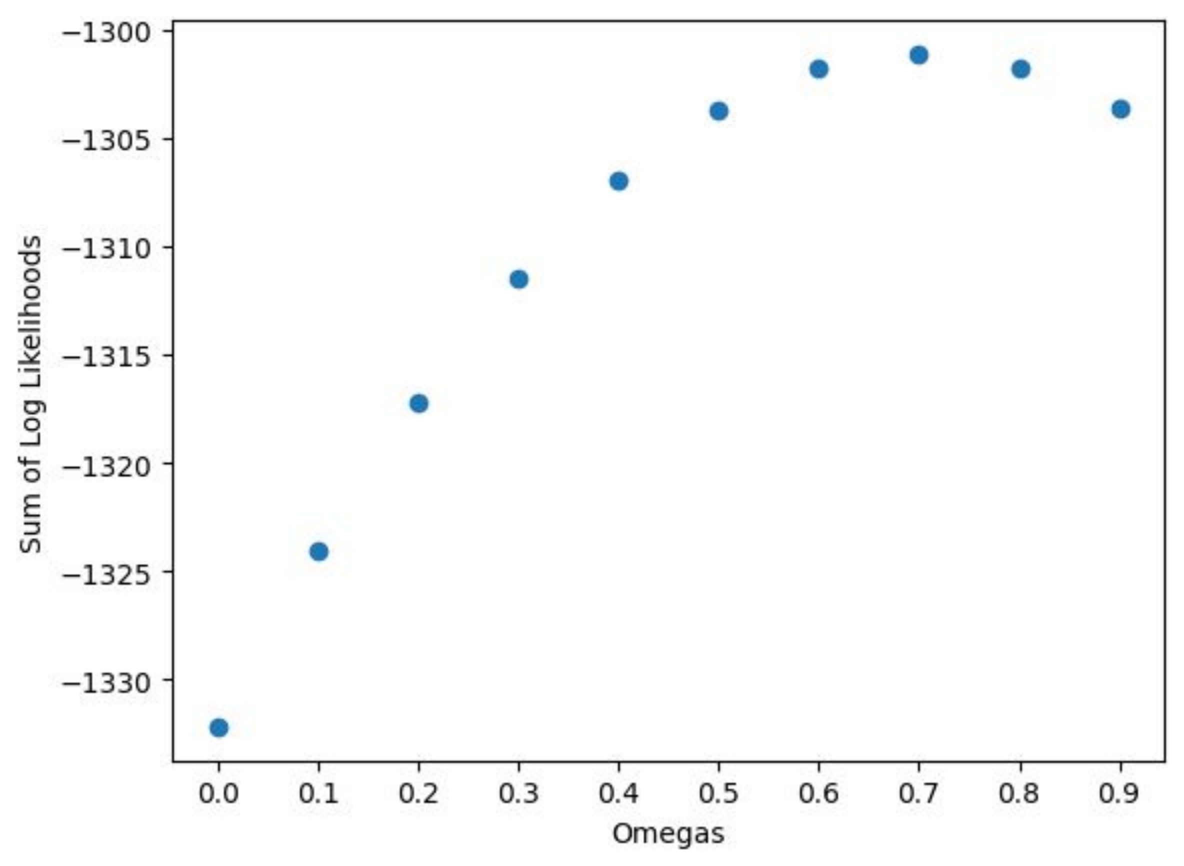
$\omega=0.5$



$\omega=1$



Grid search for omega (github)



Step-by-step estimation

[google sheet]

Thank you for listening :)

Extra reading:

1. See Collins&Cockburn 2020 for a recent review on the topic.

2. An agent could also represent a model of the environment in terms of how outcomes are associated with one another (Moran et al., 2019, 2021a, 2021b).

3. Recent results propose that a story-like framing of the instructions could lead participants to be much more model-based (Feher da Silva & Hare, 2020, 2023).