Increasing model-based decision strategies in humans

Maria K Eckstein[1], Klaus Wunderlich[2], and Anne GE Collins[1]

[1]Department of Psychology, University of California, Berkeley, USA;
[2]Graduate School of Systemic Neurosciences, Ludwig Maximilian University, Munich, Germany

Introduction

Humans are equipped with various strategies for solving problems. Two such strategies are **model-based** and **model-free** decision making [1].

Model-based strategies (*goal-directed* in psychology): Problem solvers employ an internal model, specifying which actions lead to which results in which states, to infer the sequence of actions that leads to the desired outcome.

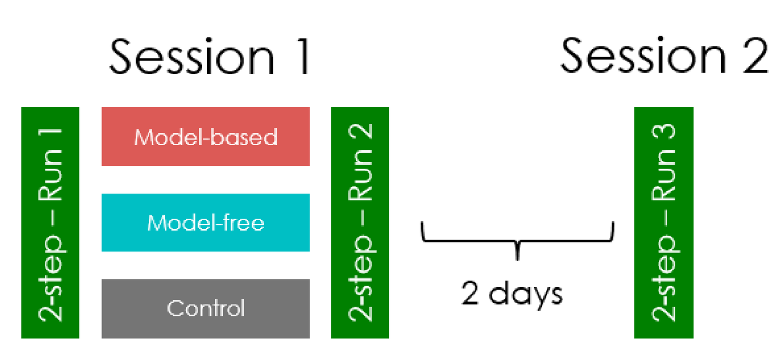
Model-free strategies (*habitual* in psychology): Decisions are made by repeating action that previously led to reward. This leads to the formation of habits, stable stimulus-response associations.

Model-free strategies are fast and computationally cheap, but inflexible. Humans fall back to these strategies under stress [2]. **Model-based** strategies are flexible, but slow and demanding. They are better suited to attain long-term goals.

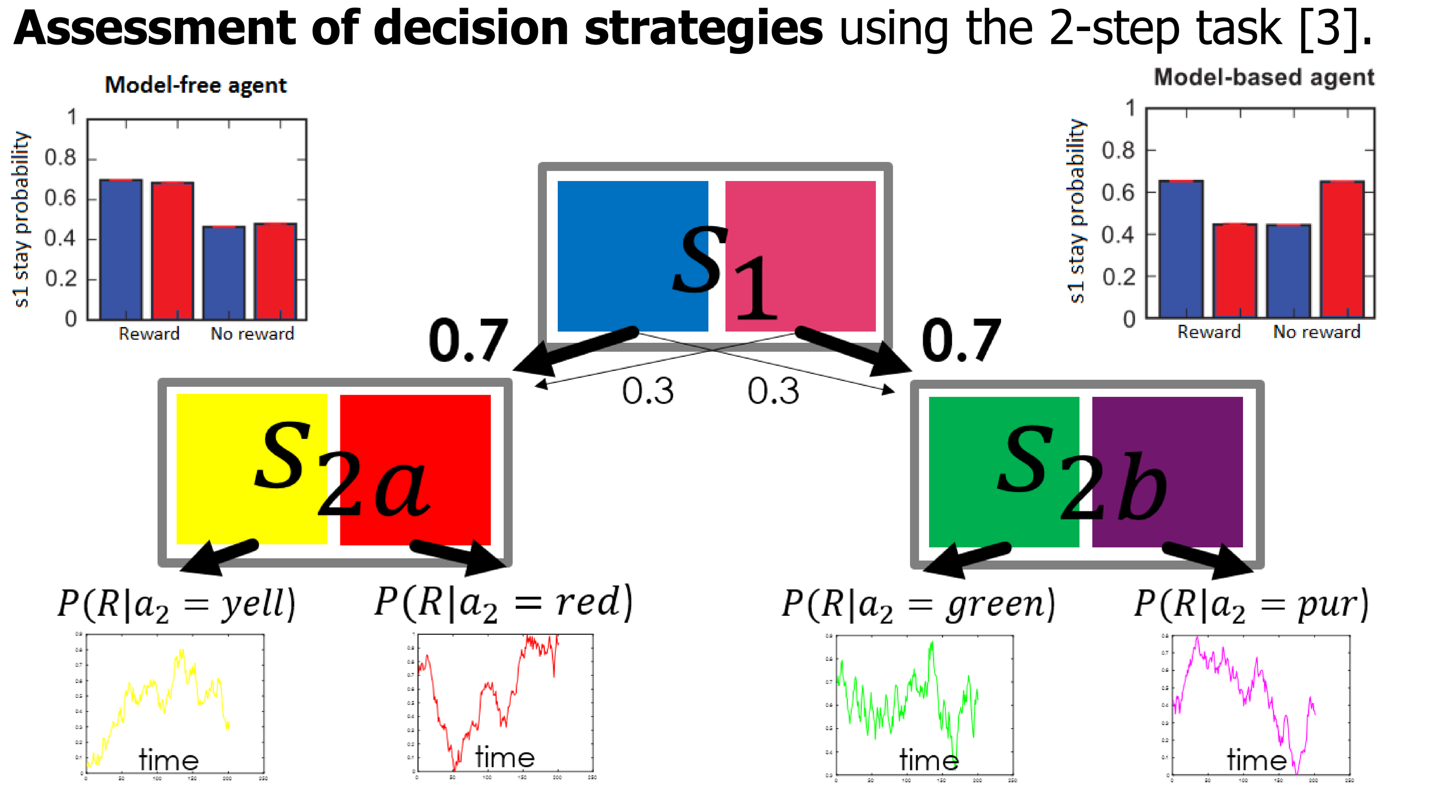
Research is lacking as to how model-based strategies can be improved.

Procedure

Overview.

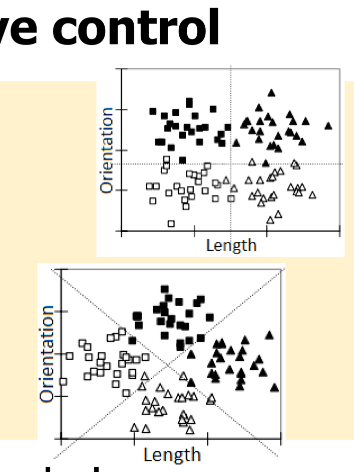


Assessment of decision strategies using the 2-step task [3].



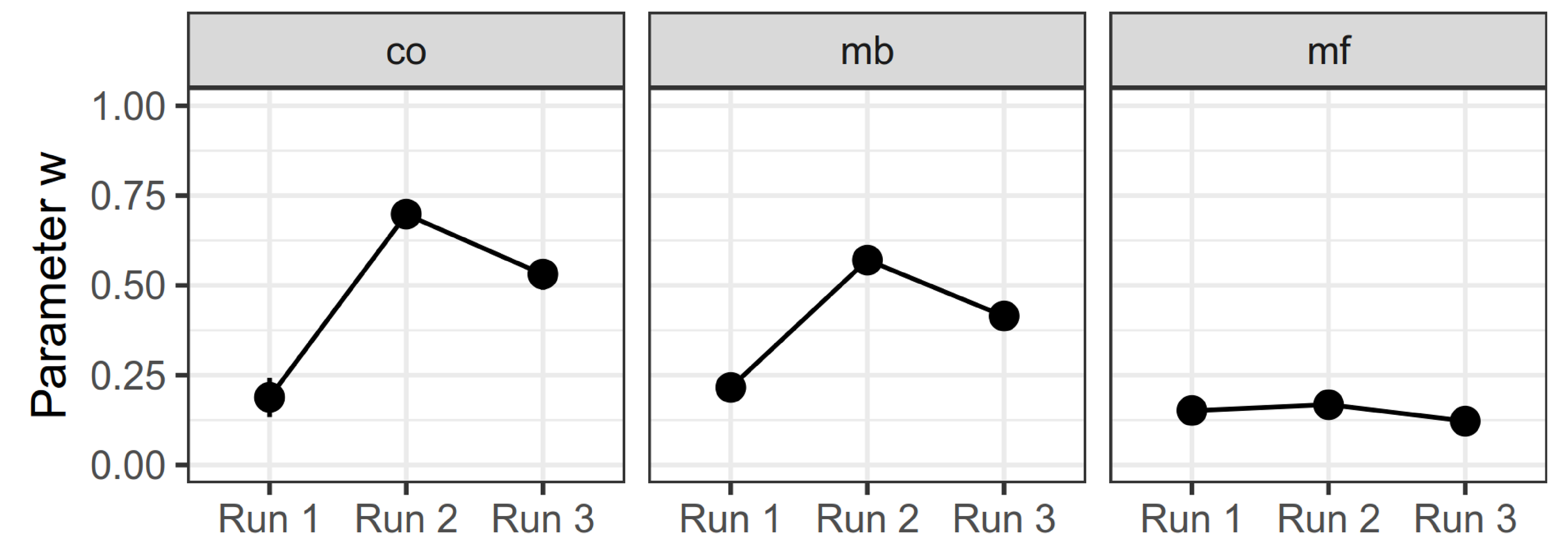
Model-based tasks, model-free tasks, or active control

Model-based: Tower of London;
rule-based category learning [4]
Model-free: Simple operant condition task [5];
information-integration category learning [4].
Control: Number comparison [6];
orientation discrimination [7].



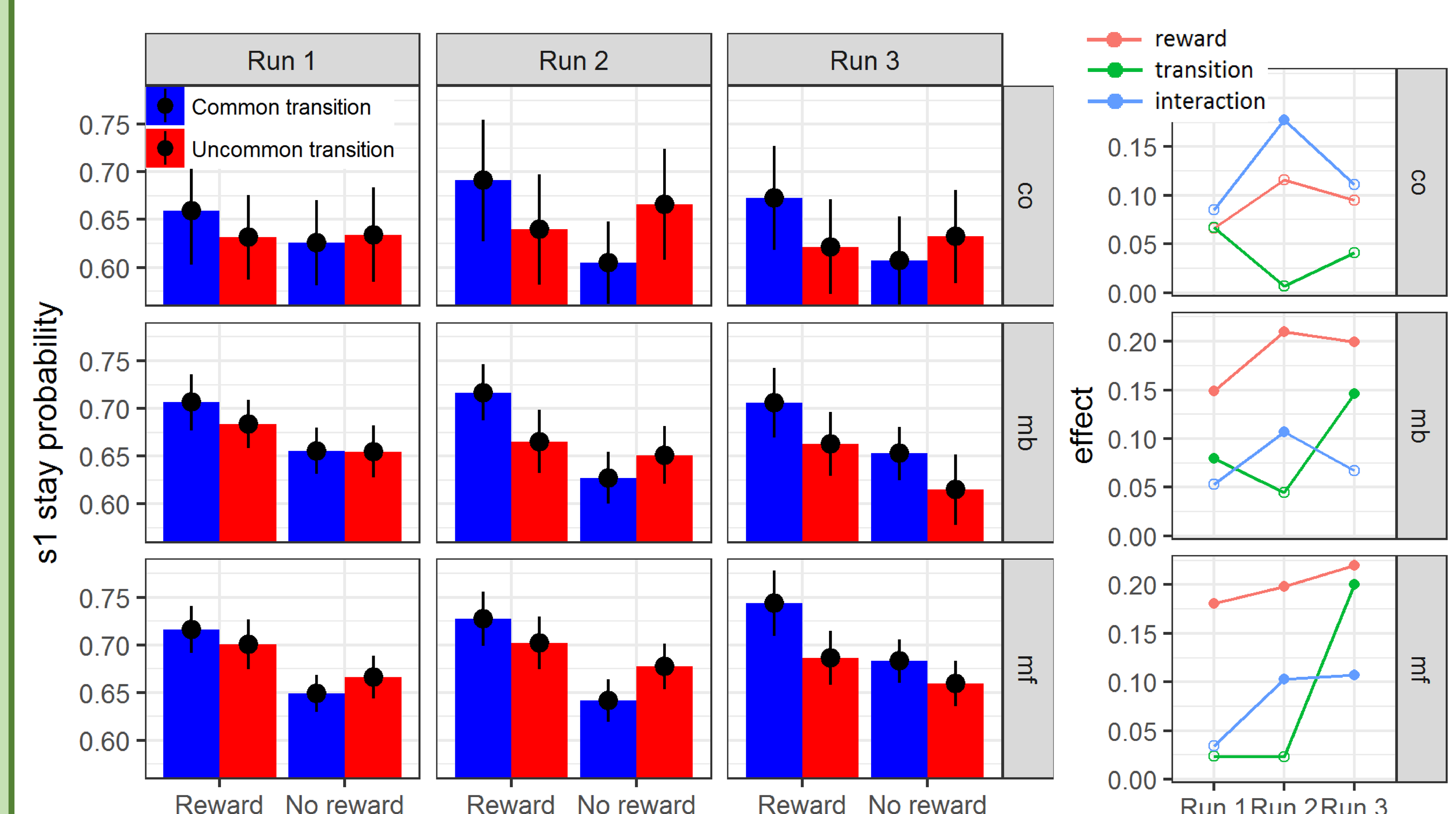
Participants. 116 healthy young adults from UC Berkeley.

Training Effects: Computational Analysis



Parameter *w* increased in participants with model-based training (mb: t(38)=8.7, p < 0.001) and control (co: t(18)=9.4, p , 0.001), suggesting an **increase in model-based decision making**. Both run ($\chi^2(2)$=170.6, p < 0.001) and group ($\chi^2(2)$=88.1, p < 0.001) had significant effects on *w*. The interaction between run and group was also significant ($\chi^2(4)$=94.8, p < 0.001).

Training Effects: Behavioral Analysis



Model-based patterns were evident after all trainings: Run 2, interaction between reward and transition in control group (co: beta=0.18, z=2.05, p=0.040); after model-based training (mb: beta=0.11, z=2.54, p=0.011); and after model-free training (mf: beta=0.10; z=2.92, p=0.004).

Model-free patterns (main effect of reward) were evident in model-based (mb) and model-free (mf) training groups in all runs (all beta>0.15, all z>3.41, all p < 0.001).

Computational Model

We used a reinforcement learning model reported previously [e.g., 8]. Agents select actions based on estimated values, using a softmax rule.

State-2 action values are updated using basic Rescorla-Wagner:
$Q(s_2, a_2) = Q(s_2, a_2) + \alpha(r - Q(s_2, a_2))$

State-1 action values are updated in two ways, **model-based**:
$Q_{mb}(s_1, a_1) = \sum_{s_2} p(s_2|s_1, a_1) \cdot max(Q_{mb}(s_2, a_2))$

and **model-free:**
$Q_{mf}(s_1, a_1) = Q_{mf}(s_1, a_1) + \alpha(\lambda Q_{mf}(s_2, a_2) - Q_{mf}(s_1, a_1))$

Final state-1 action values are a combination of model-based and model-free values: $Q(s_1, a_1) = w \cdot Q_{mb}(s_1, a_1) + (1 - w) \cdot Q_{mf}(s_1, a_1)$

This model was fit to participants' behavior using maximum likelihood in a hierarchical model.

Conclusion

Model-based decision making was evident after all three trainings, and *increased* after model-based training and control.

Behavioral analyses showed that participants took the model of the task into account after training, but not before (*model-based* component). Participants were sensitive to reward at all times (*model-free* component).

Computational modeling showed that *model-based decision making increased significantly* after training compared to before in the model-based and control groups.

References. [1] Sutton, R. & Barto, A. (1998). *Reinforcement Learning: An Introduction.* Cambridge, MA: MIT Press.
[2] Schwabe, L. & Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: From goal-directed to habitual control of action. *Behavioural brain research*, 219(2), 321–328.
[3] Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, *69*(6), 1204–1215.
[4] Maddox, W. T. & Ashby, F. G. (2004). Dissociating explicit and procedural-learning based systems of perceptual category learning. *Behavioural Processes*, 66(3), 309–332
[5] Tricomi, E. & Lempert, K. M. (2015). Value and probability coding in a feedback-based learning task utilizing food rewards. Journal of Neurophysiology, 113(1), 4–13. doi:10.1152/jn.00086.2014
[6] Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. Trends in Cognitive Sciences, 14(12), 542–551. doi:10.1016/j.tics.2010.09.008
[7] Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. Nature reviews. Neuroscience, 11(1), 53–60. doi:10.1038/nrn2737
[8] Wunderlich, K., Smittenaar, P., & Dolan, R. J. (2012). Dopamine Enhances Model-Based over Model-Free Choice Behavior. Neuron, 75(3), 418–424. doi:10.1016/j.neuron.2012.03.042

Acknowledgements. German Academic Exchange Service (DAAD) for funding; Silvia Bunge for helpful inputs on the experimental design.