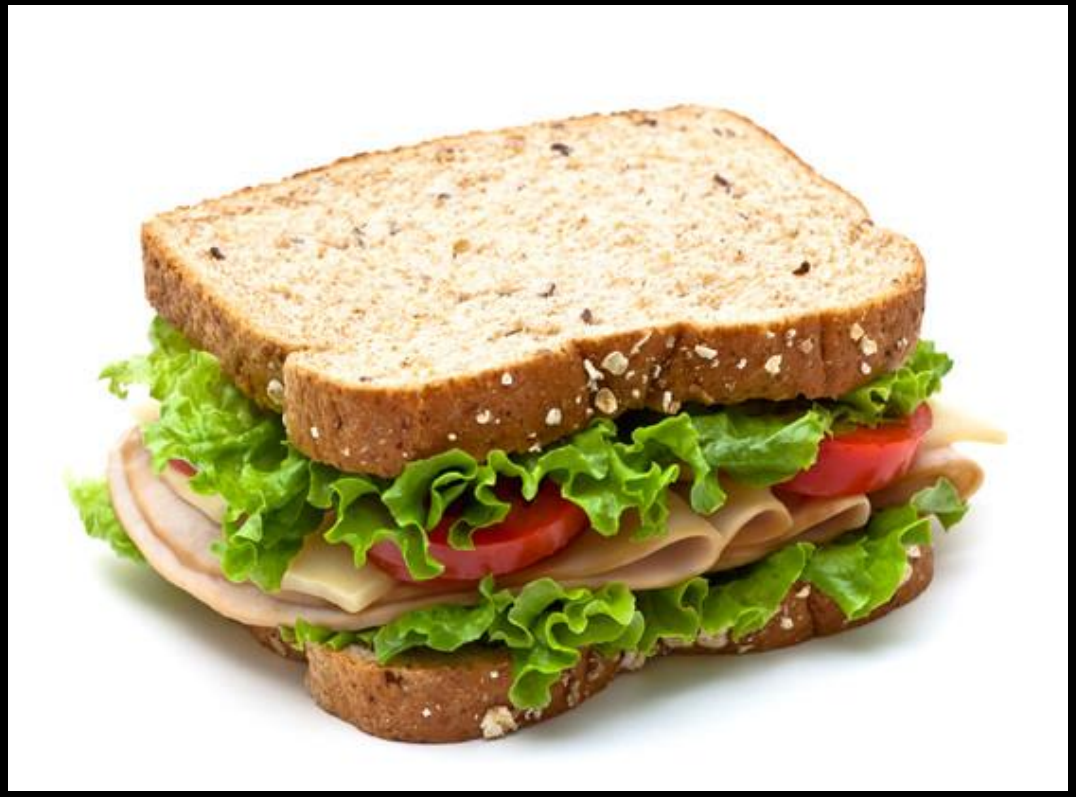
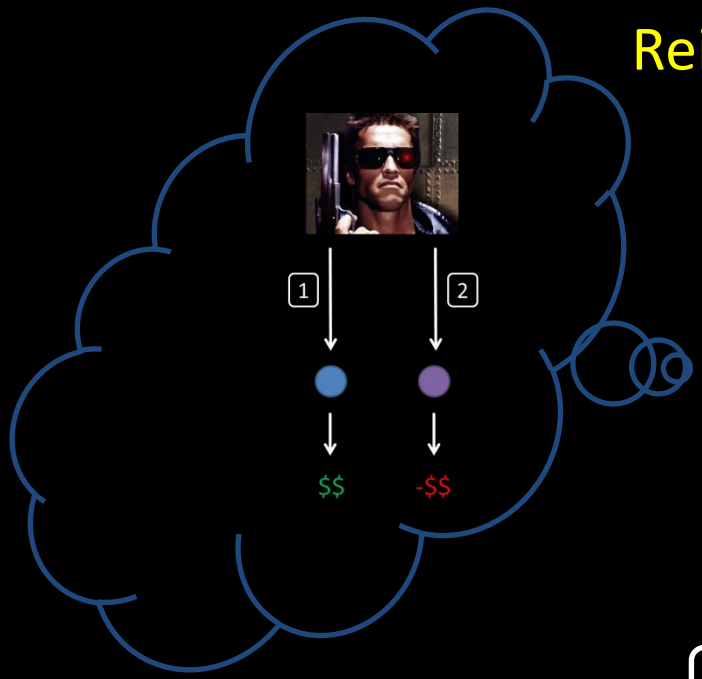
Model-free goal selection

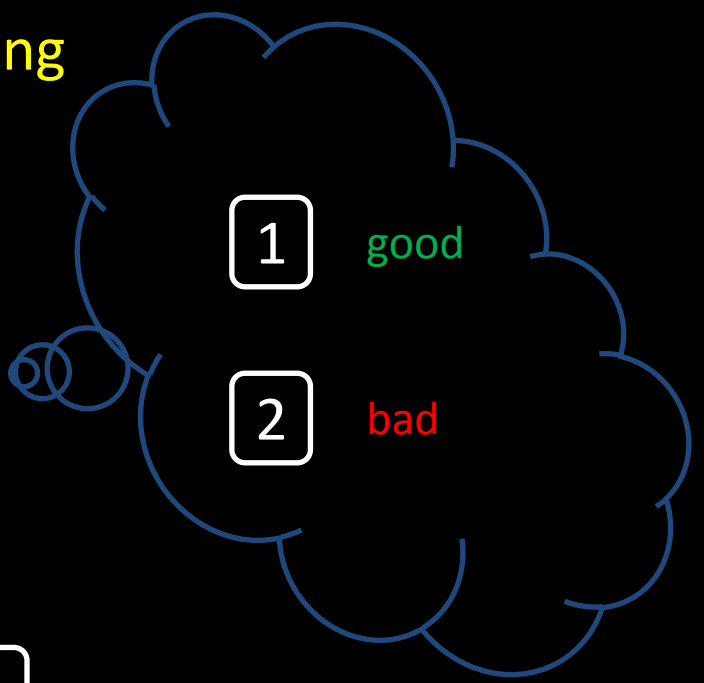
Adam Morris & Fiery Cushman



Abscond with an aardvark, abscond with an alligator…

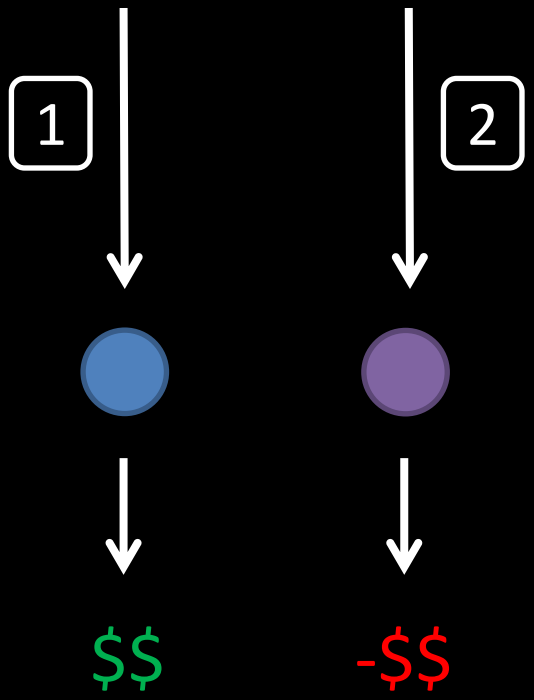
Reinforcement learning





Model-based
reinforcement learning

Goal-directed planning

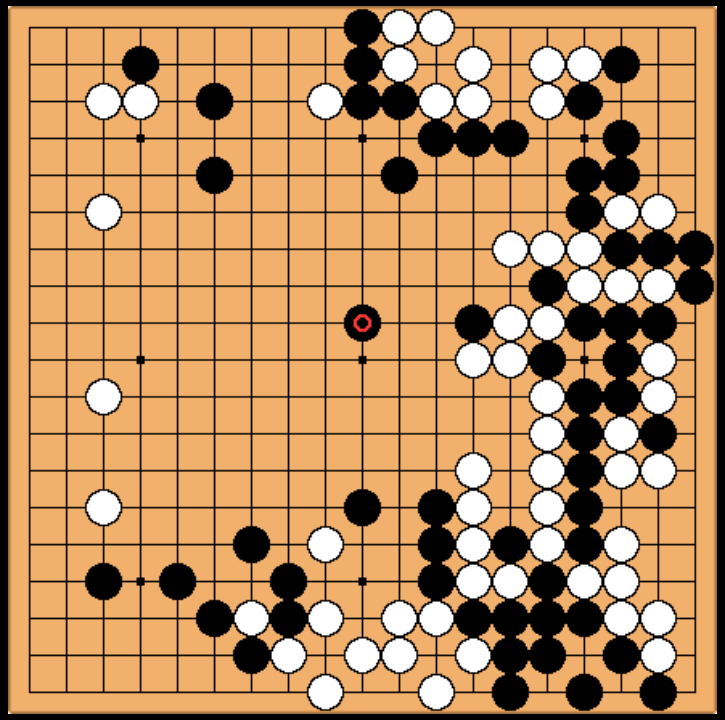


Model-free
reinforcement learning

Habitual actions

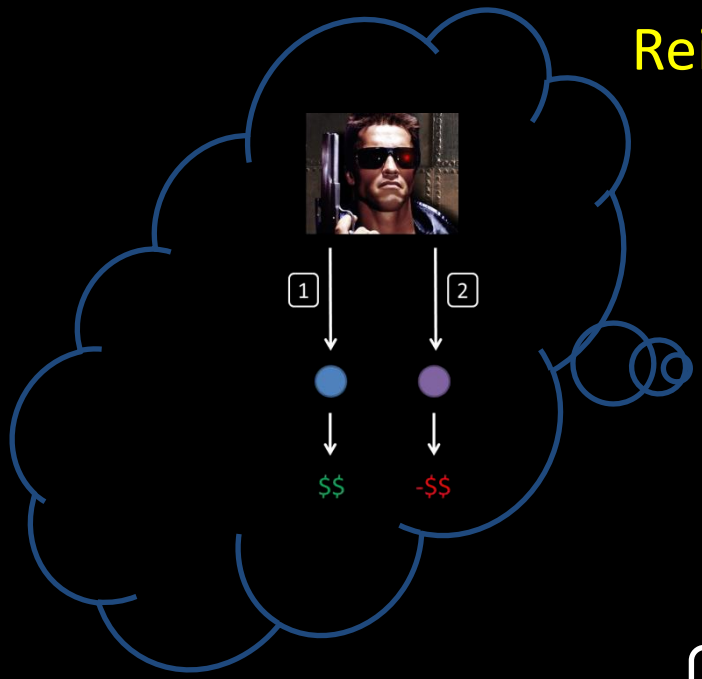
(Sutton & Barto 1998)

Reinforcement learning

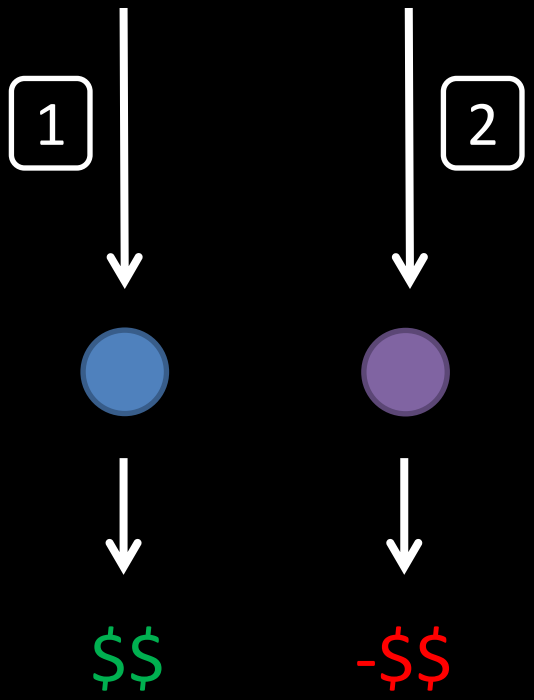


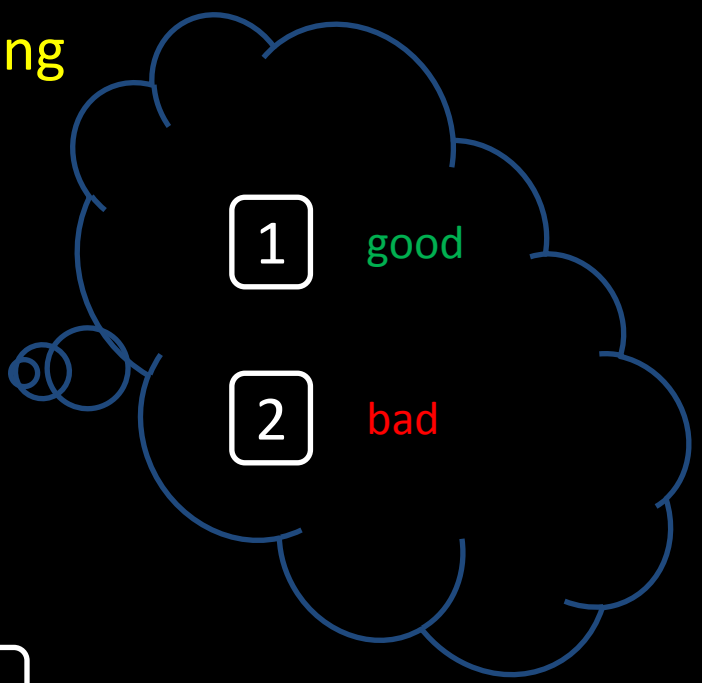
(Silver, Sutton, & Muhler 2012)

Reinforcement learning

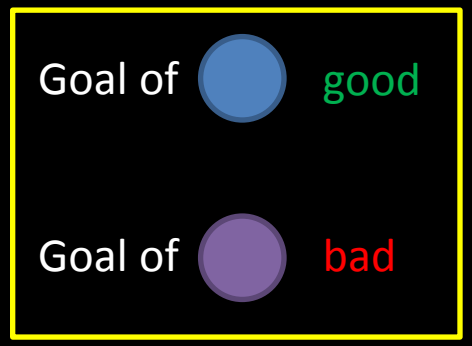


Model-based
reinforcement learning





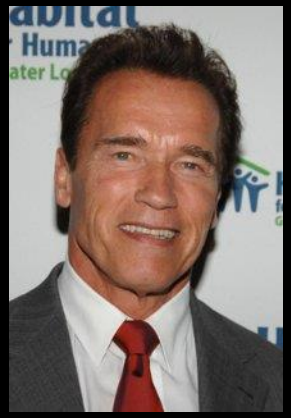
Model-free
reinforcement learning

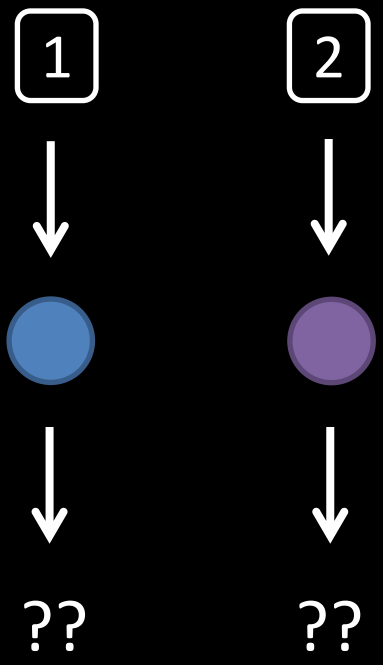


Model-free goal selection?

(Silver, Sutton, & Muhler 2012)

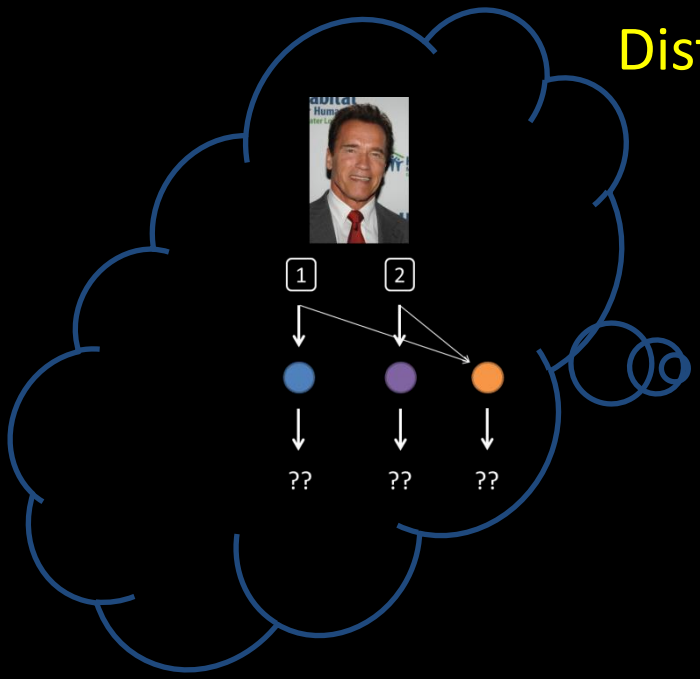
Distinguishing MB & MF

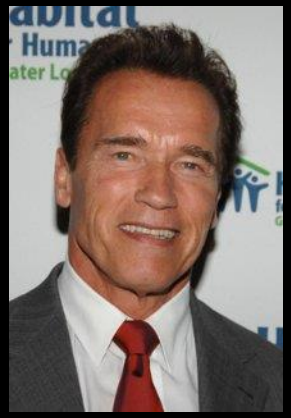


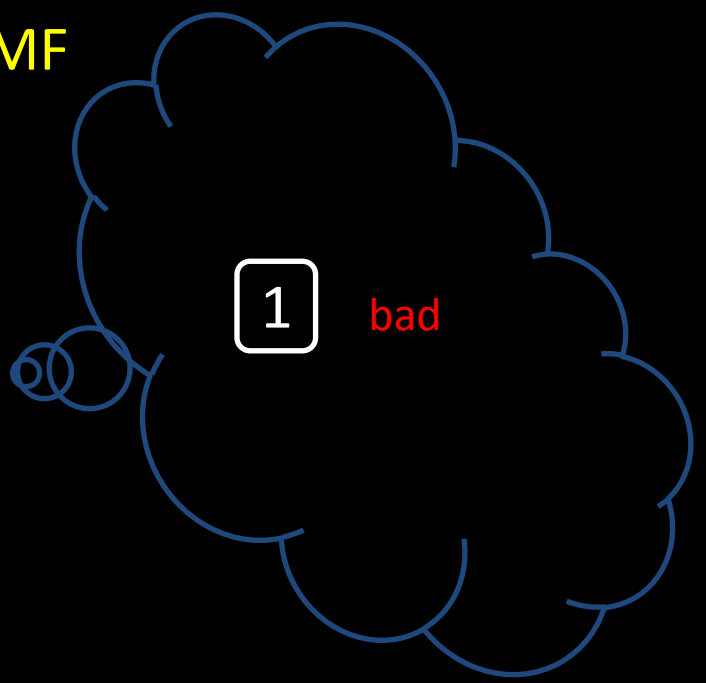


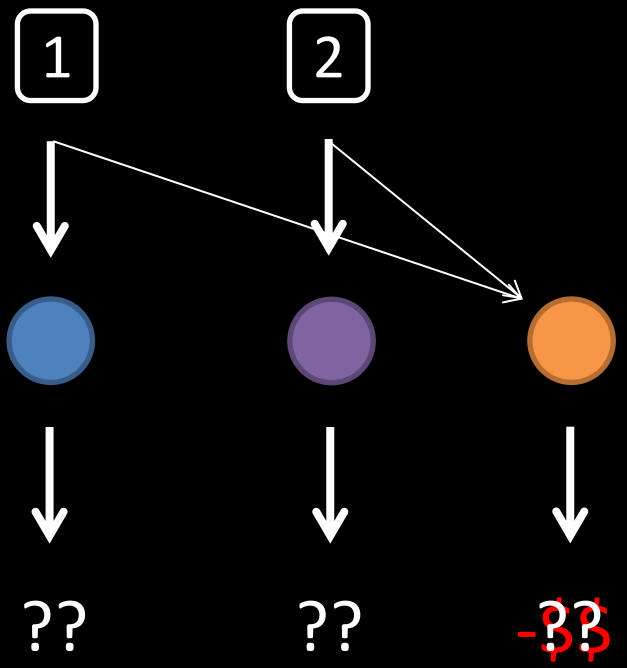
(Gläscher et al. 2010)

Distinguishing MB & MF





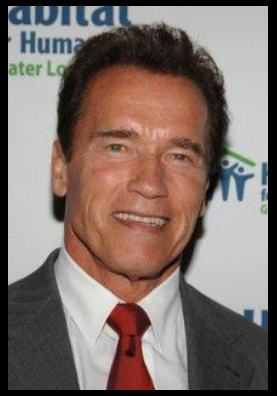


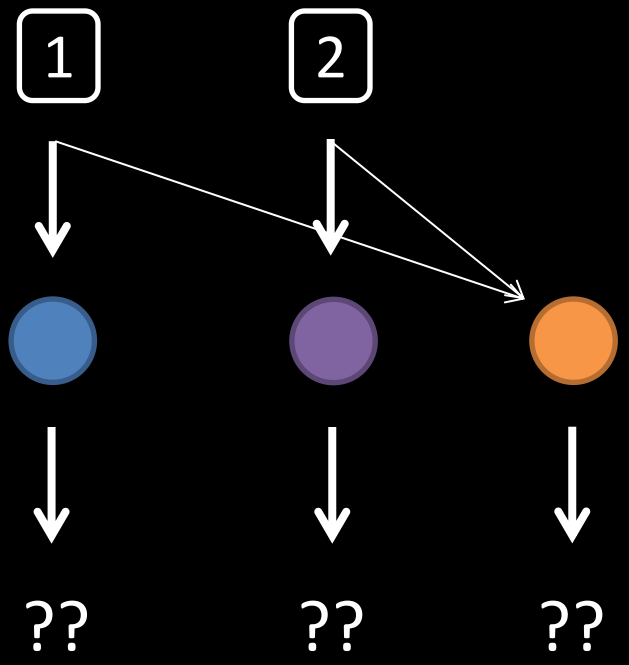


Goal of bad

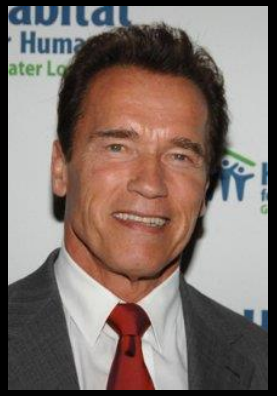
(Gläscher et al. 2010)

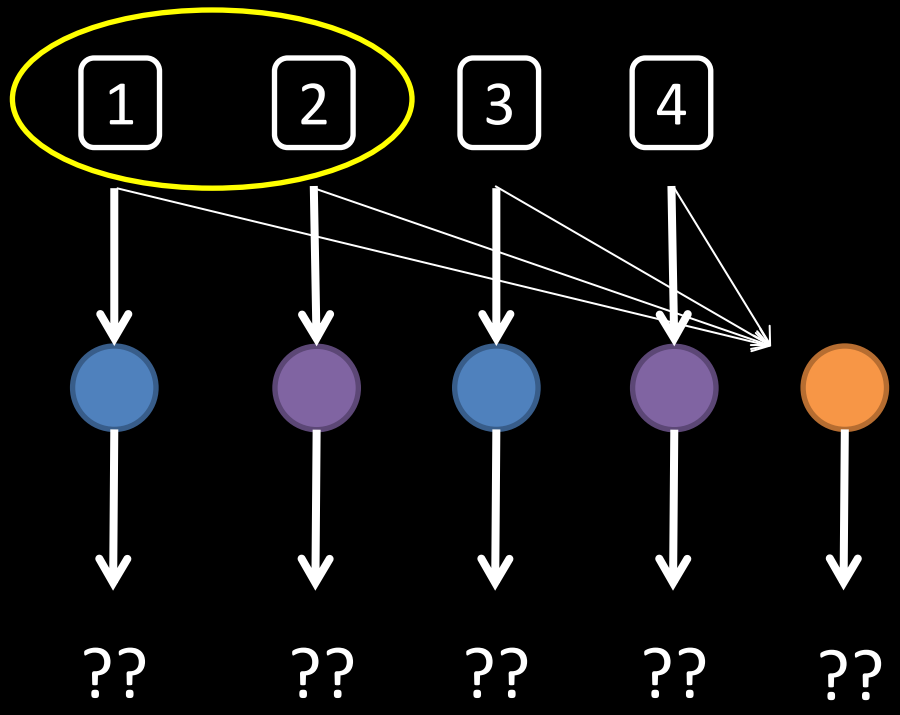
Our experiment



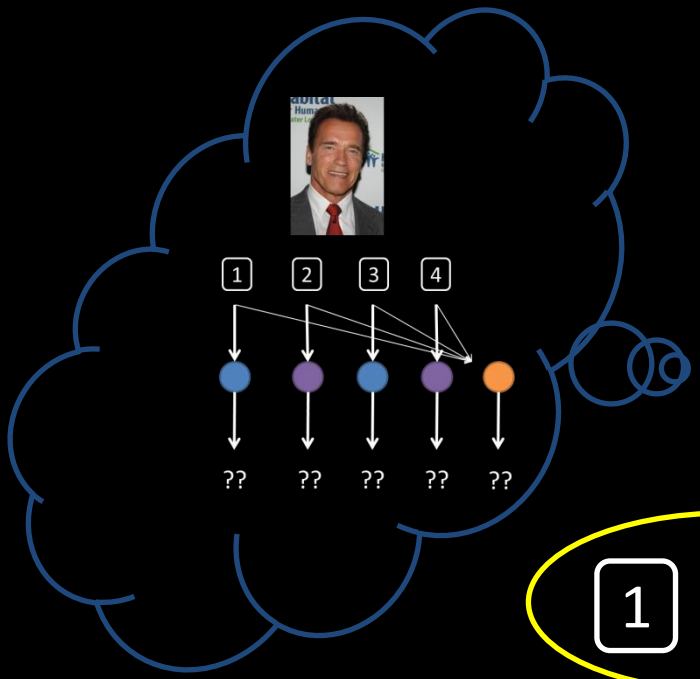


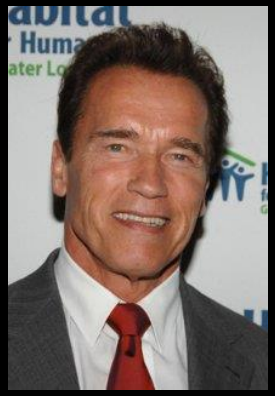
Our experiment

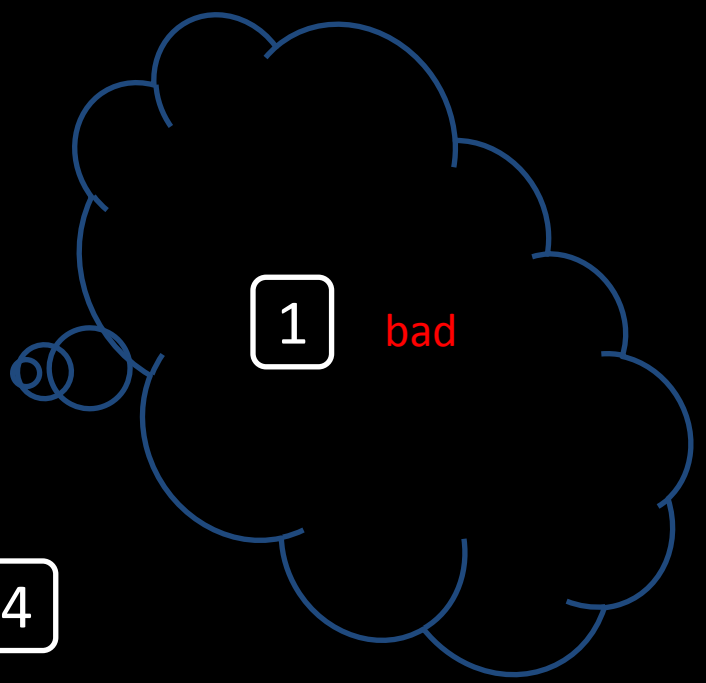


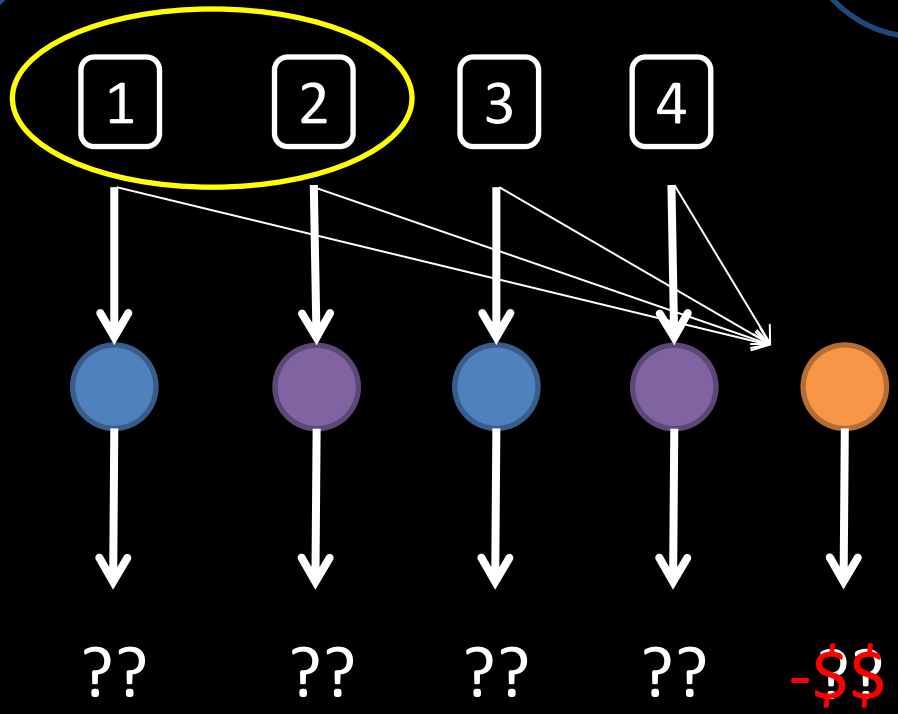


Our experiment



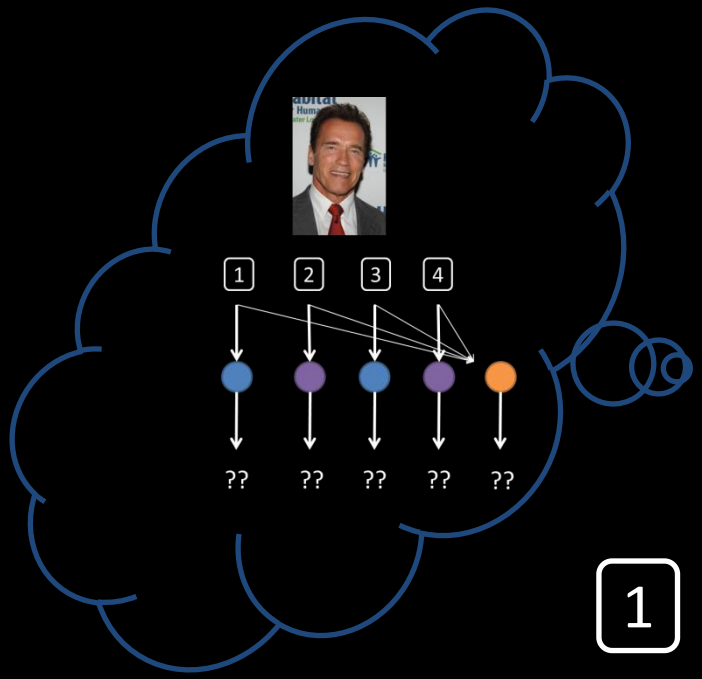


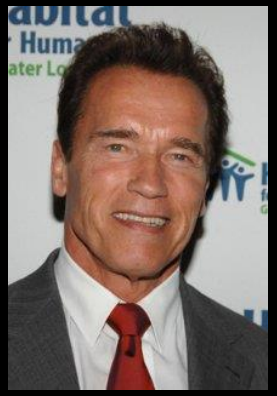


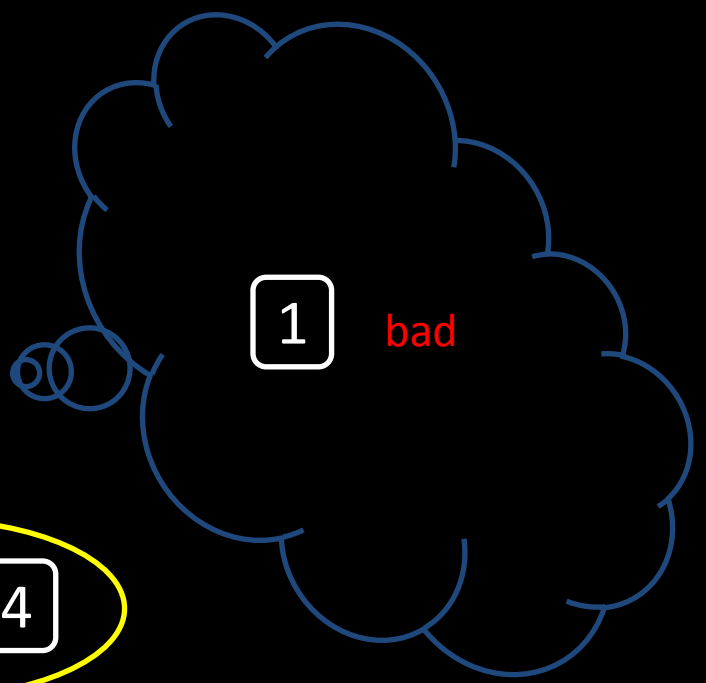


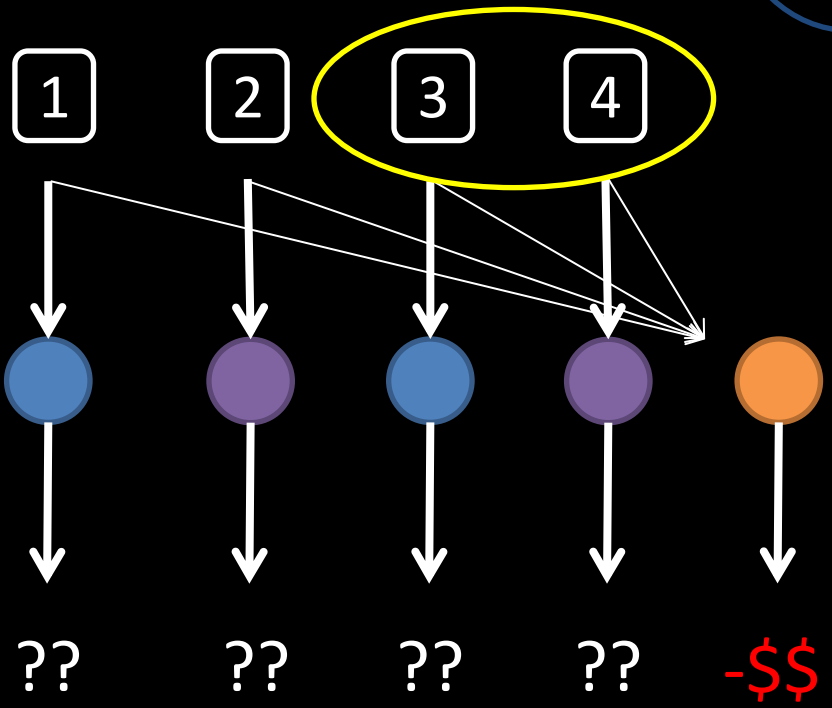
Goal of bad

Our experiment



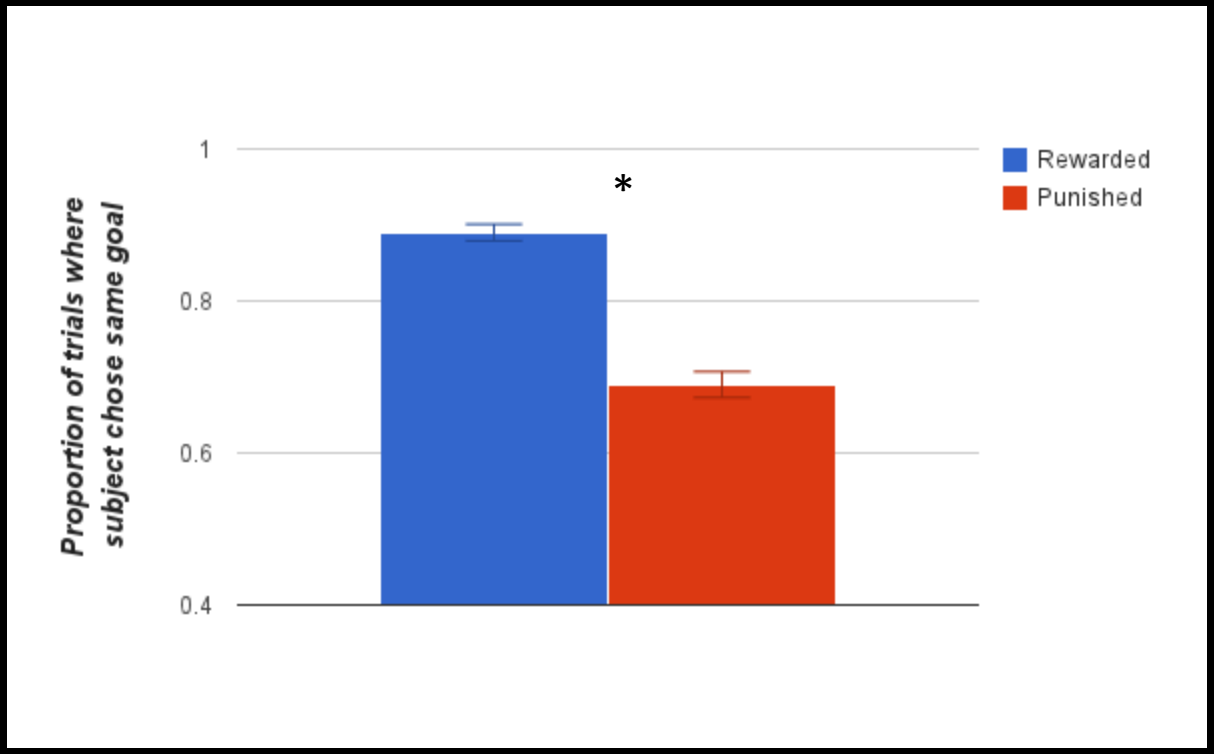






Goal of bad

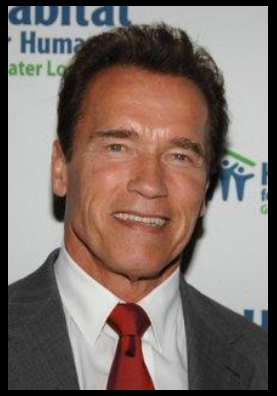
Our experiment

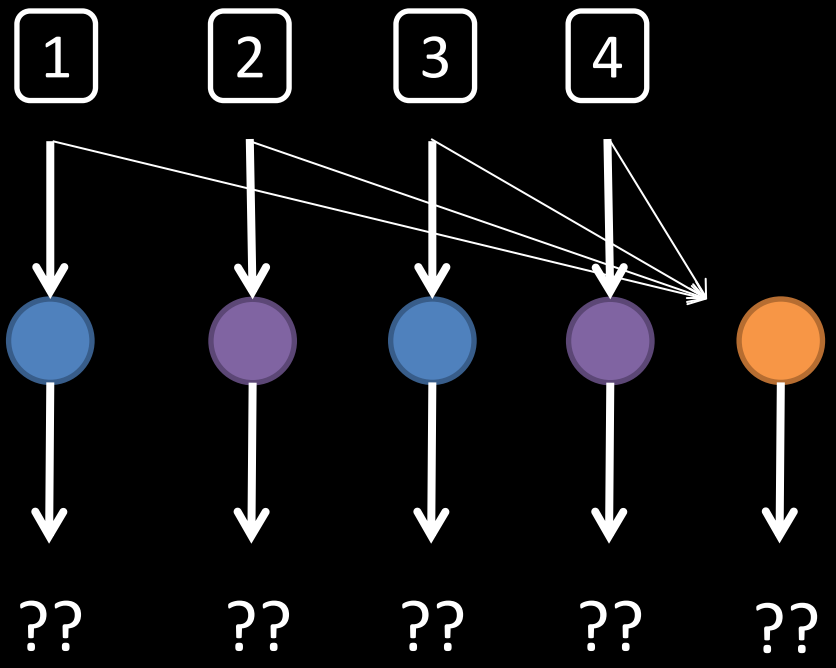


Orange's value → **Choice on next trial**

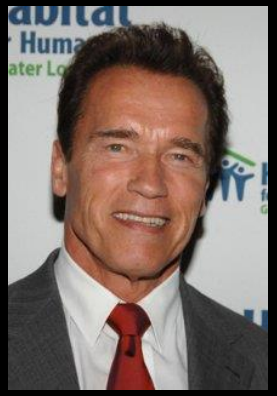
$\beta = .191$, $p < .0001$

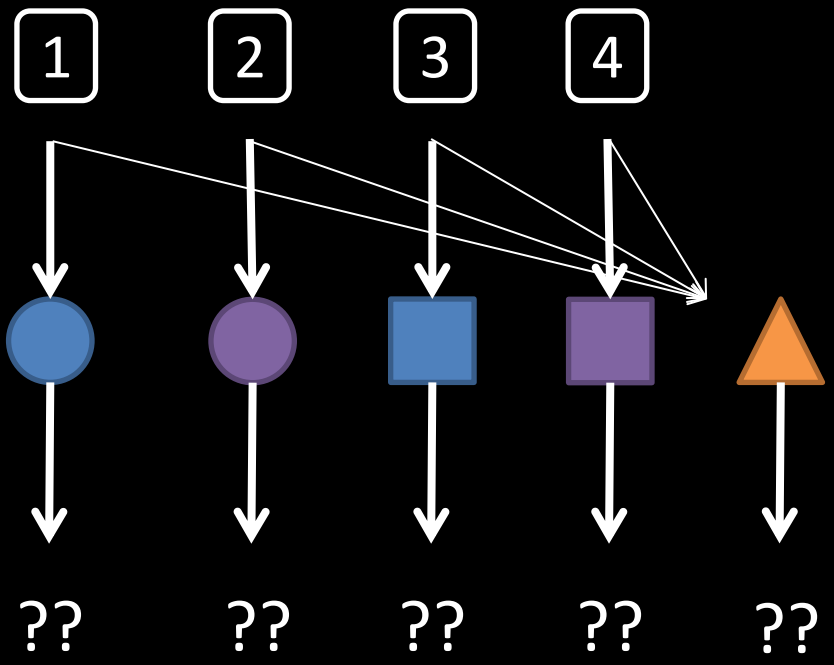
Experiment #2



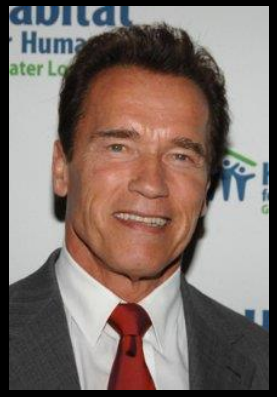


Experiment #2

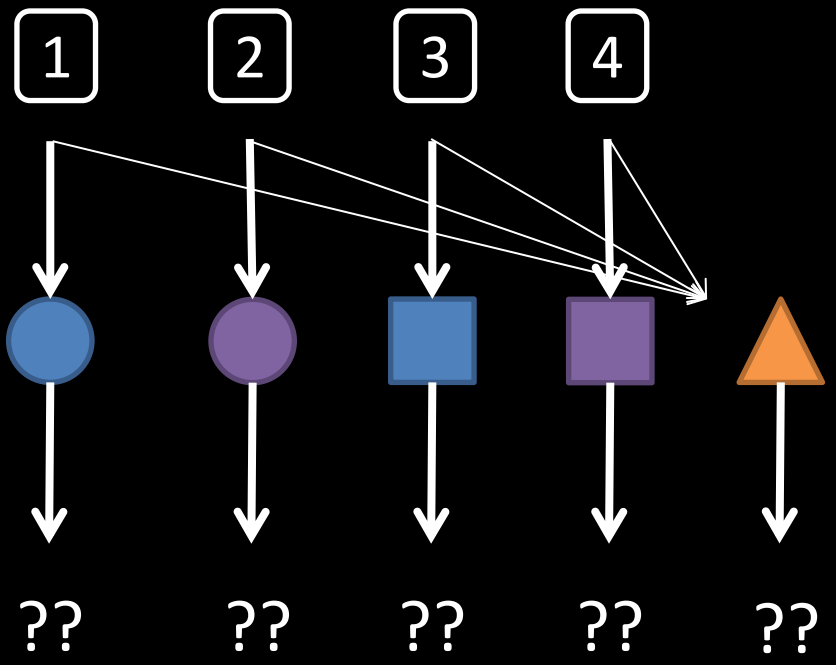




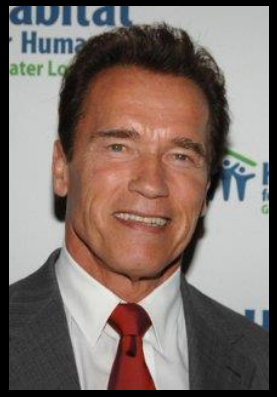
Experiment #2



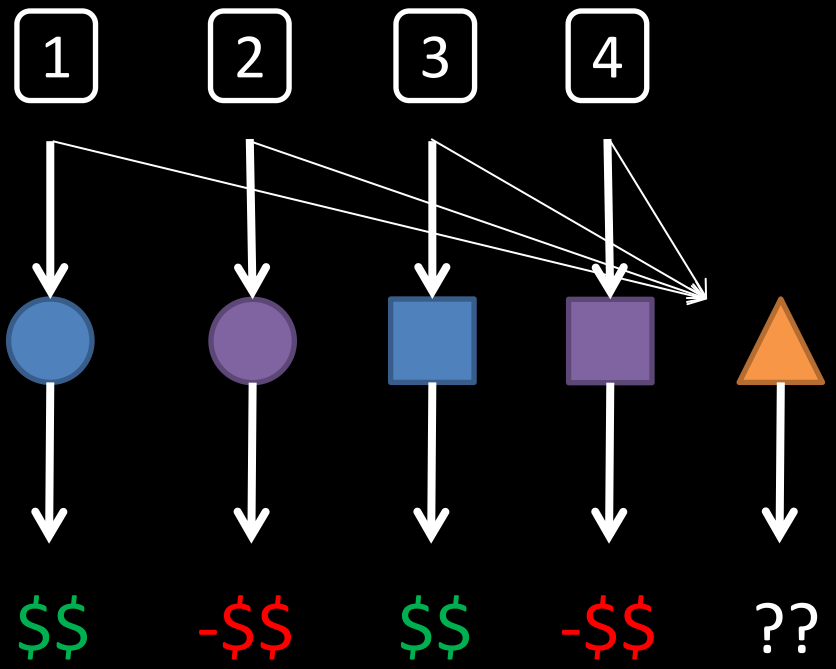
Color trial



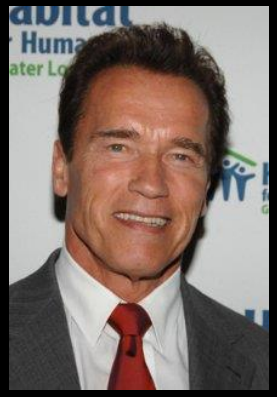
Experiment #2



Color trial

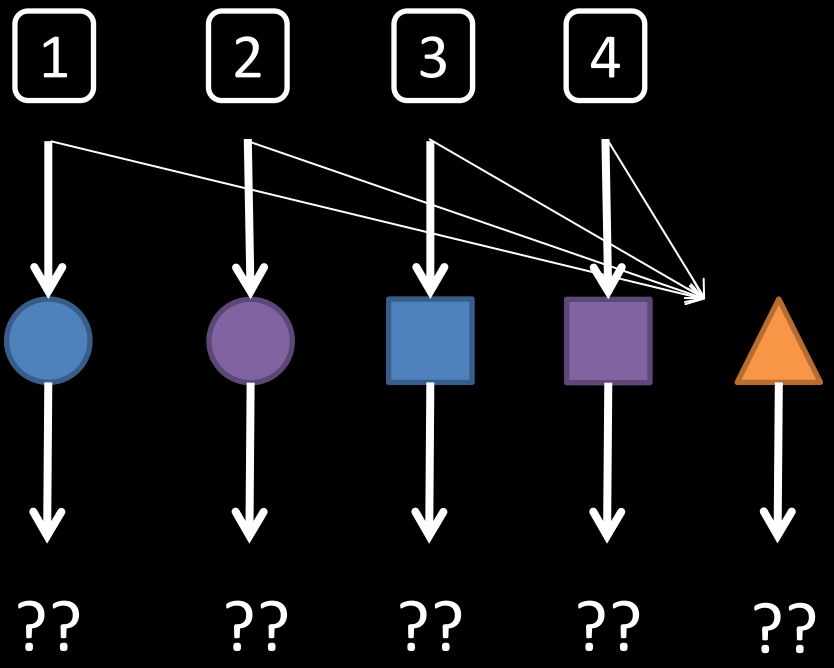


Experiment #2

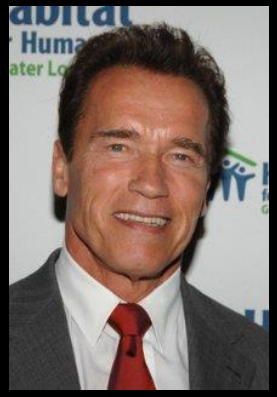


Color trial

Shape trial

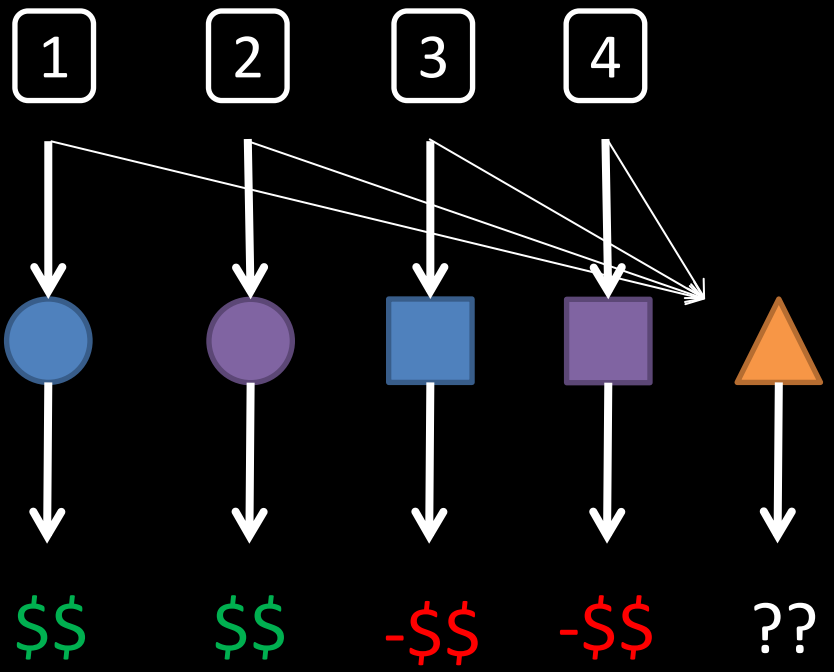


Experiment #2

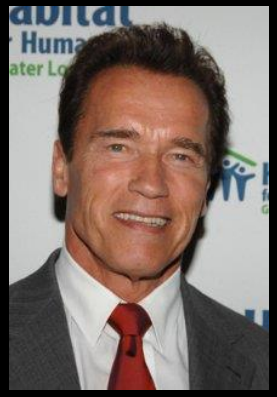


Color trial

Shape trial

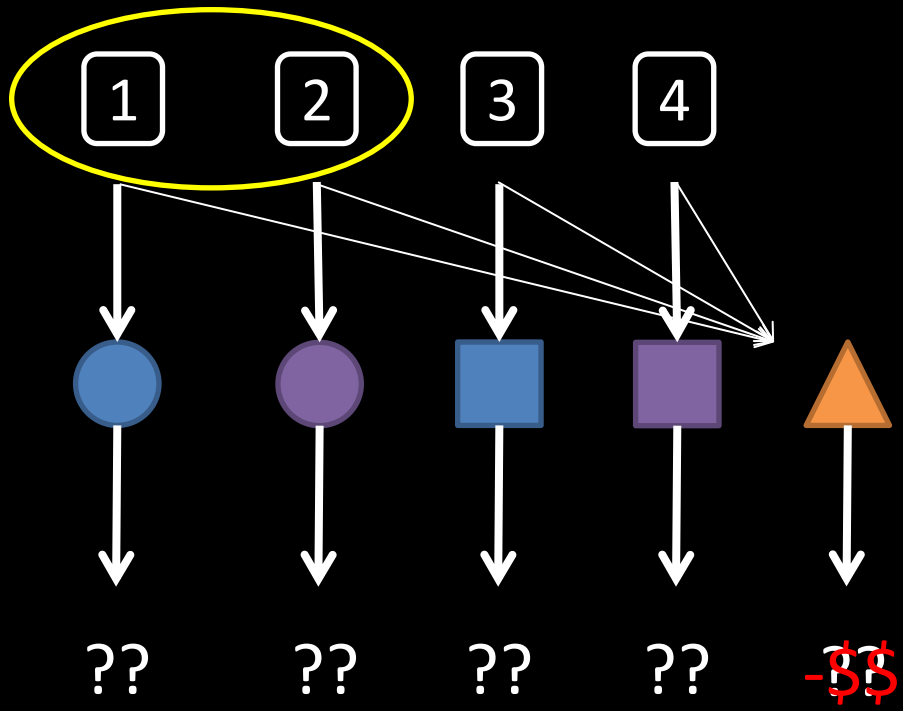


Experiment #2

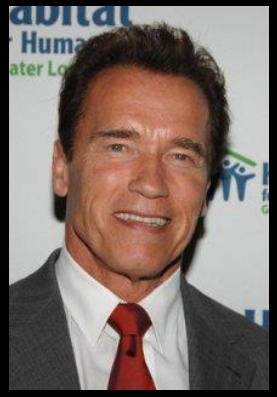


Color trial

Shape trial

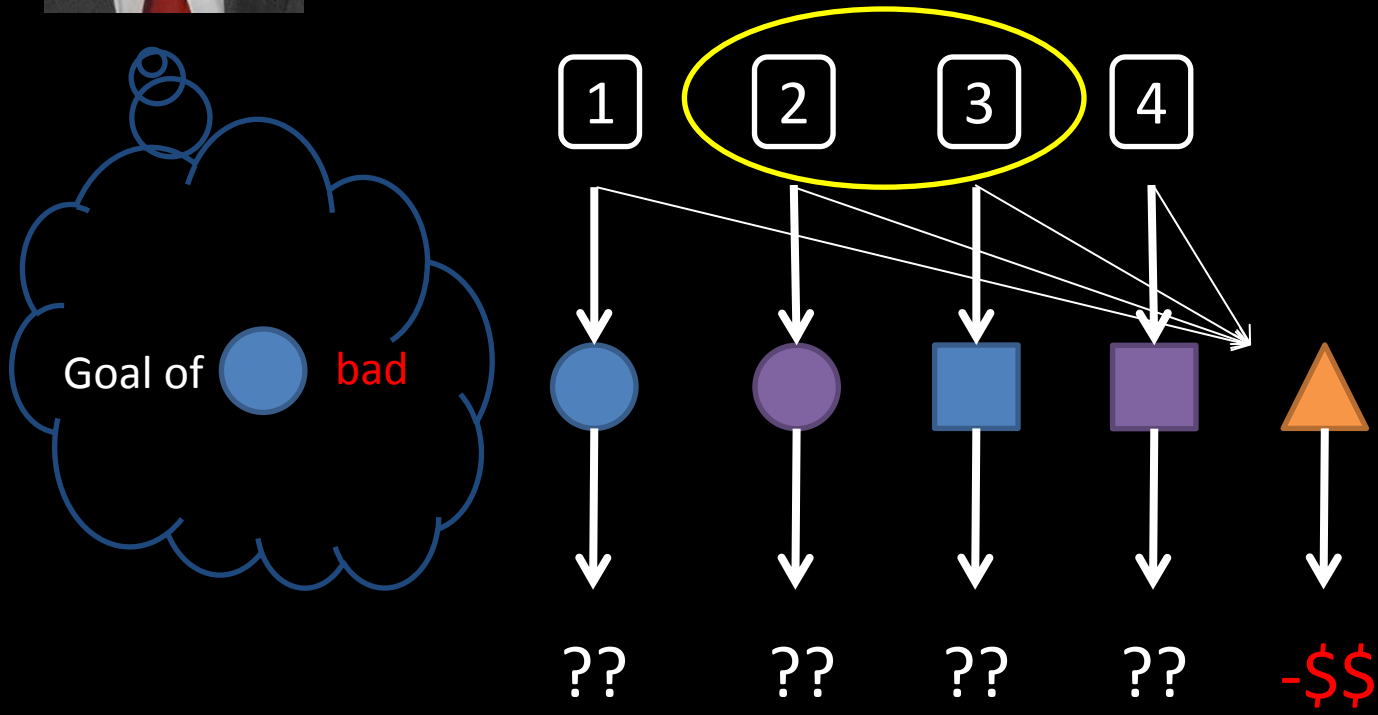


Experiment #2



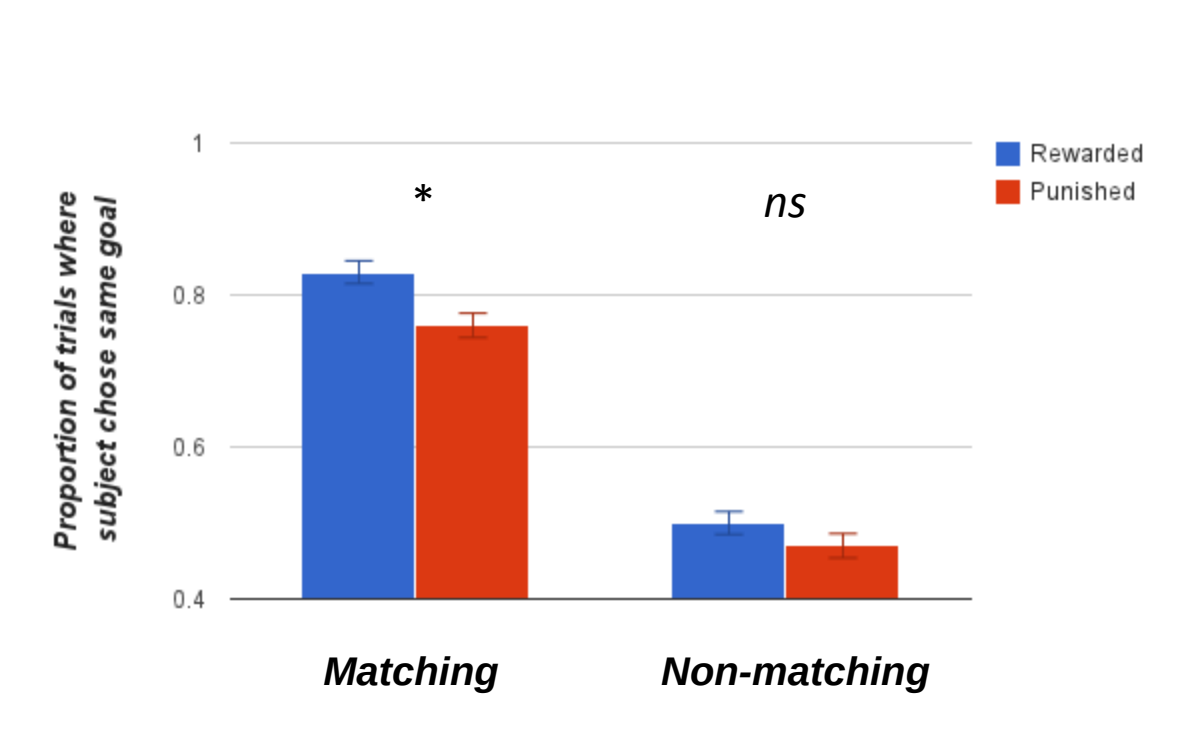
Color trial

Shape trial



Prediction: We'll only see the effect on "matching" critical trials.

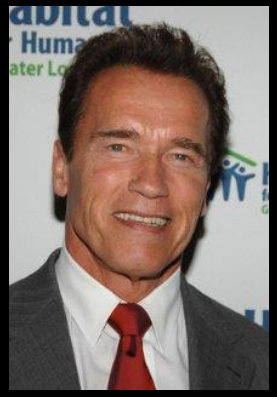
Experiment #2

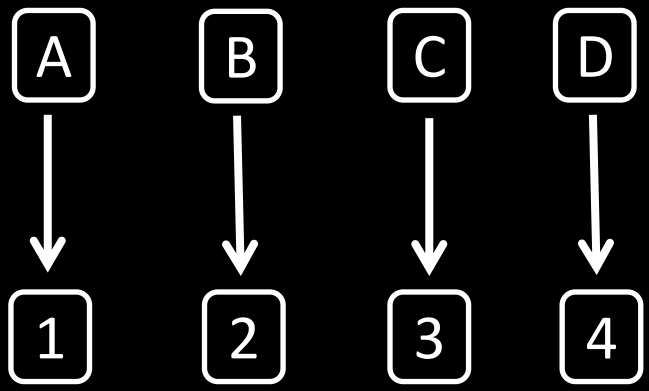


Orange's value * Type of critical trial → Choice on next trial

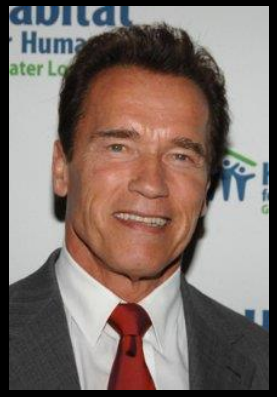
$\beta = .049$, $p < .01$

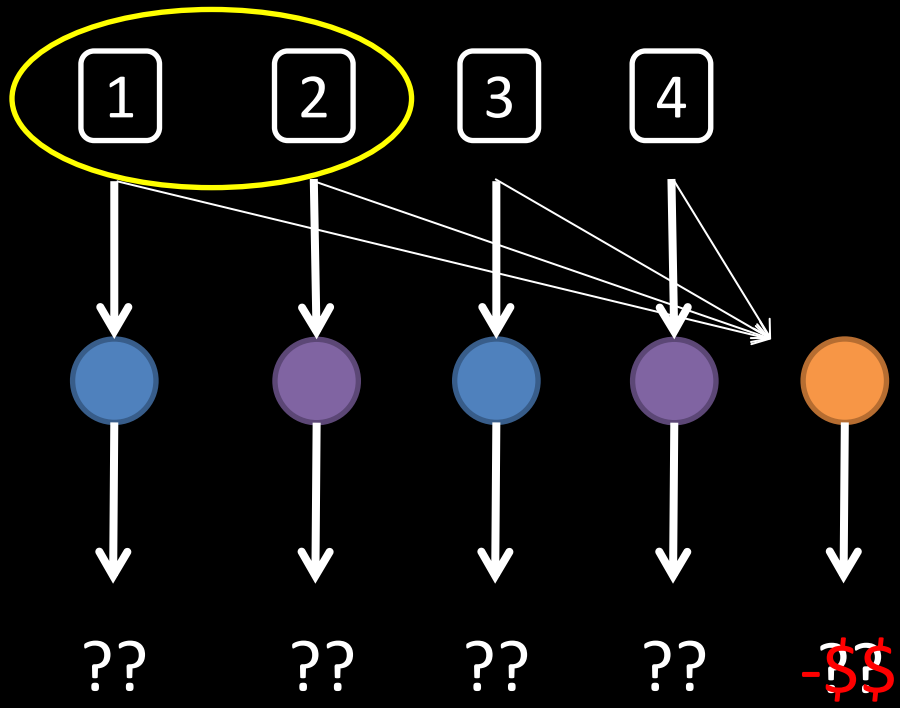
Experiment #3



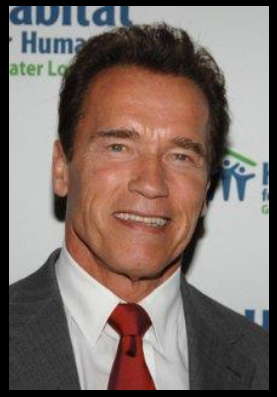


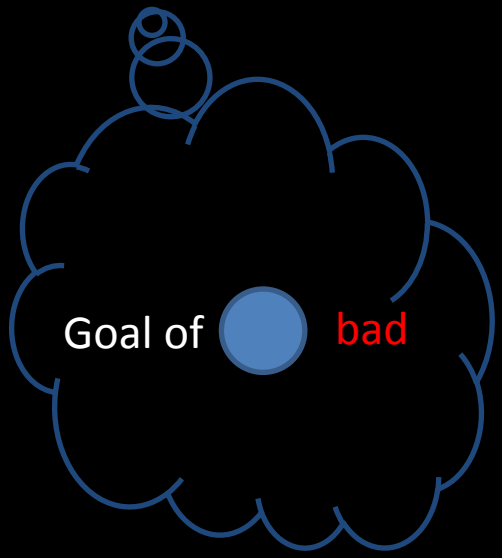
Experiment #3

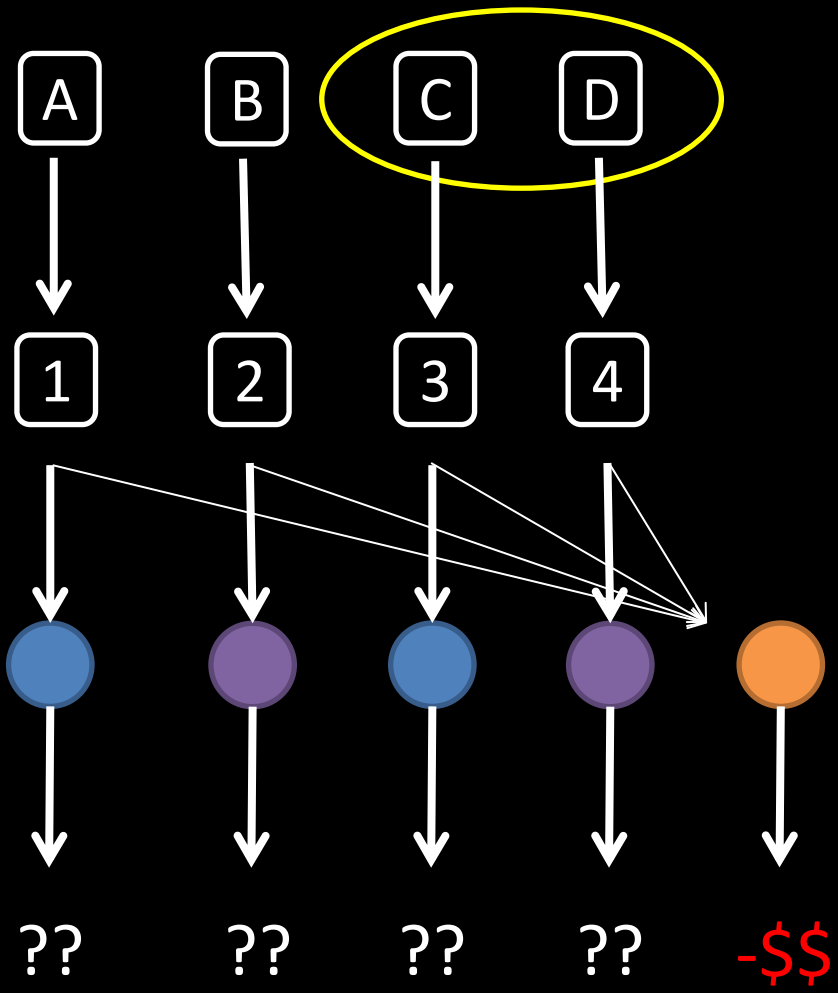




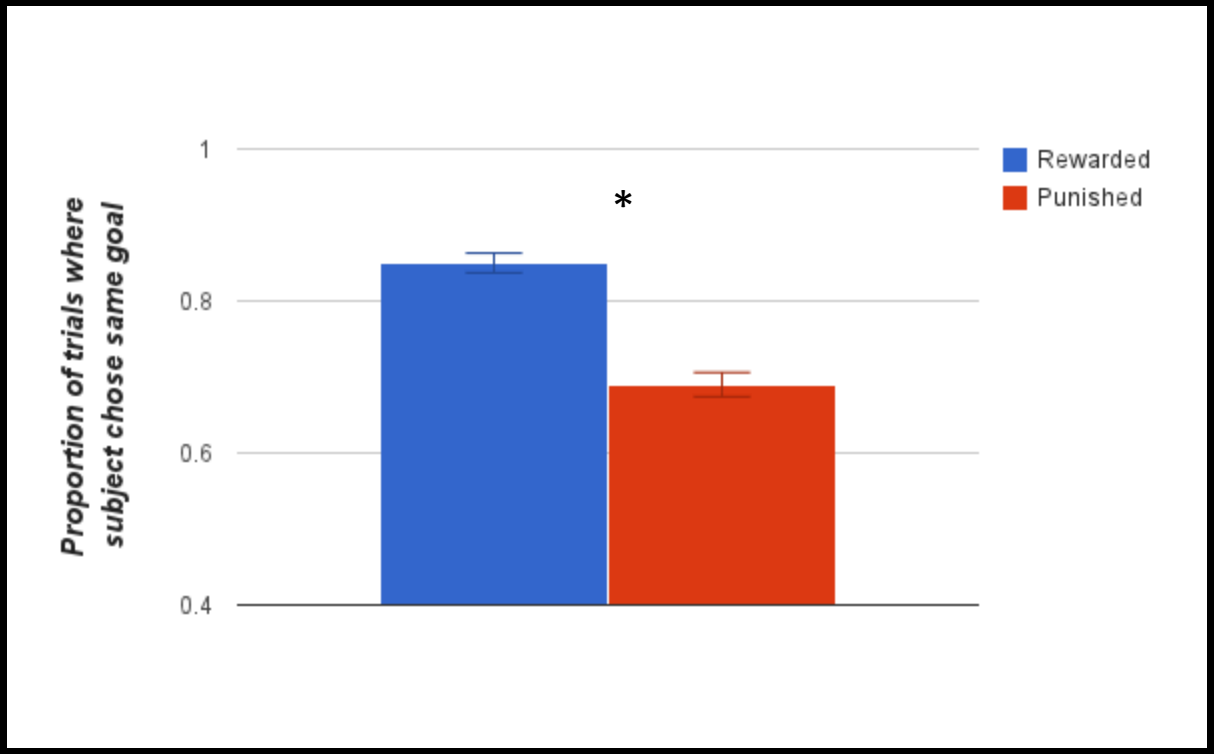
Experiment #3







Experiment #3

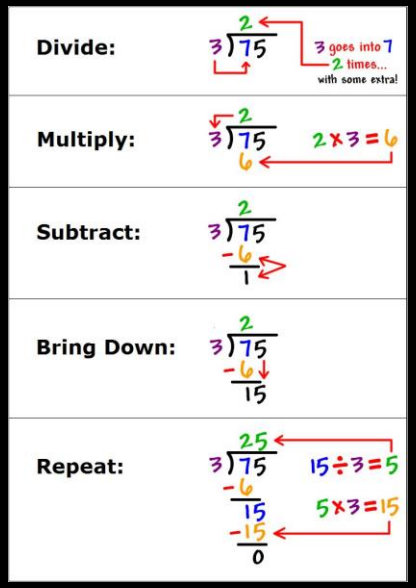


Orange's value → Choice on next trial

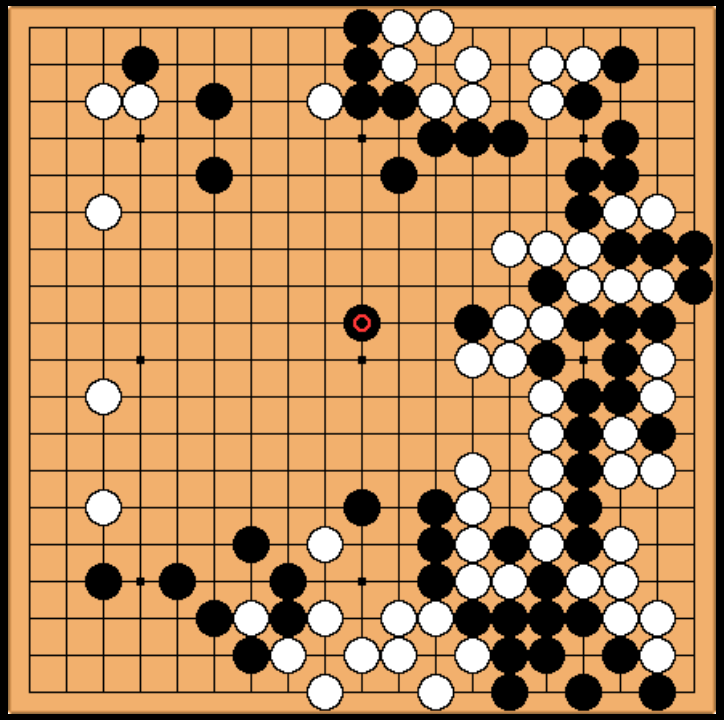
$\beta = .143$, $p < .0001$

Psychological Phenomena

Habit-like features of goals (Huang & Bargh 2014)



AI Implications



Thank you!