Computation of Values
in Primate Prefrontal Cortex

Dominic Barraclough
Hyojung Seo
Daeyeol Lee

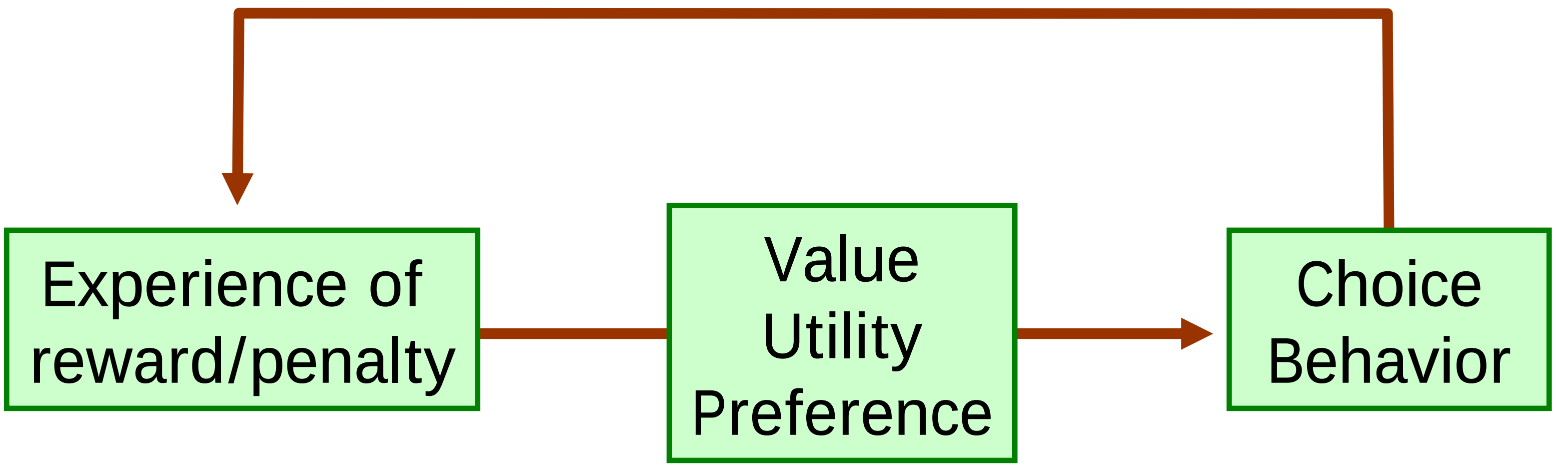
Department of Brain & Cognitive Sciences
Center for Visual Science
University of Rochester

Neurobiology of Decision Making
The Banbury Center
Cold Spring Harbor Laboratory
May 23, 2005

funded by James McDonnell Foundation/NEI/NIMH/NINDS

Hedonic Psychophysics (KT, 1984):

Objective value (money) vs. Subjective value (utility)

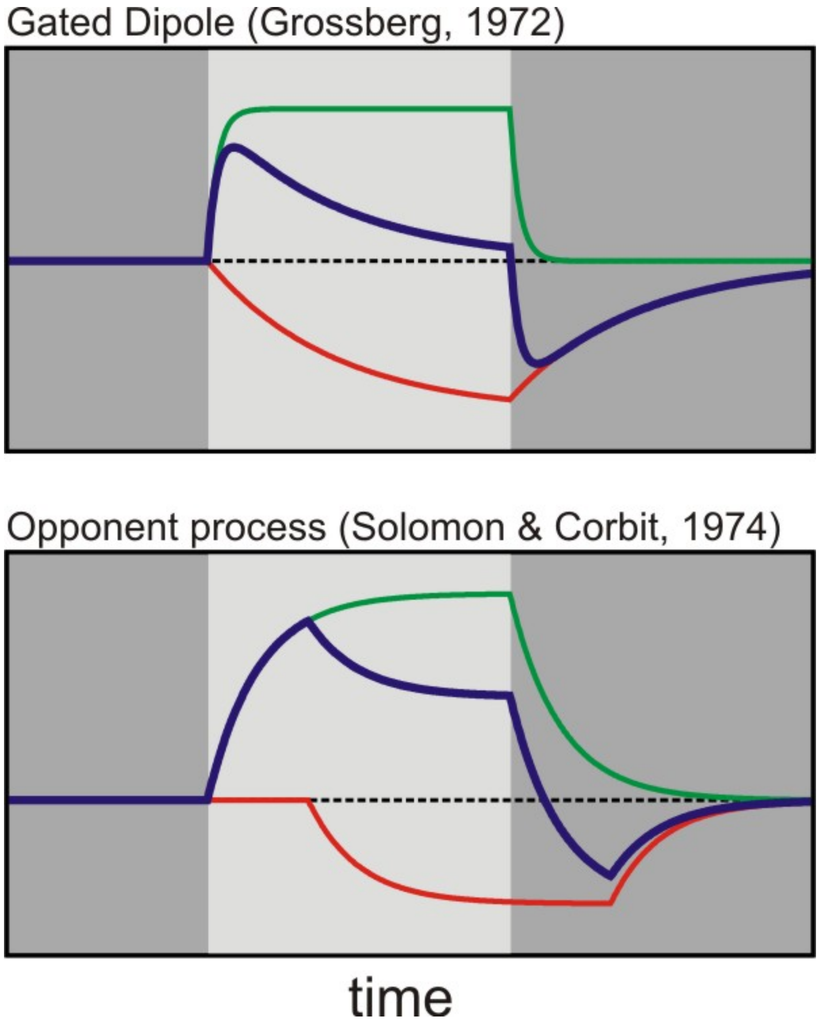


- Hedonic adaptation (Helson, 1947)
- Opponent-process theory (Solomon & Corbit, 1974)
- Prospect theory (Kahneman & Tversky, 1979)
- Incentive contrast (Flaherty, 1982)
- Average reward reinforcement learning (Puterman, 1994)

Kahneman & Tversky (1984):

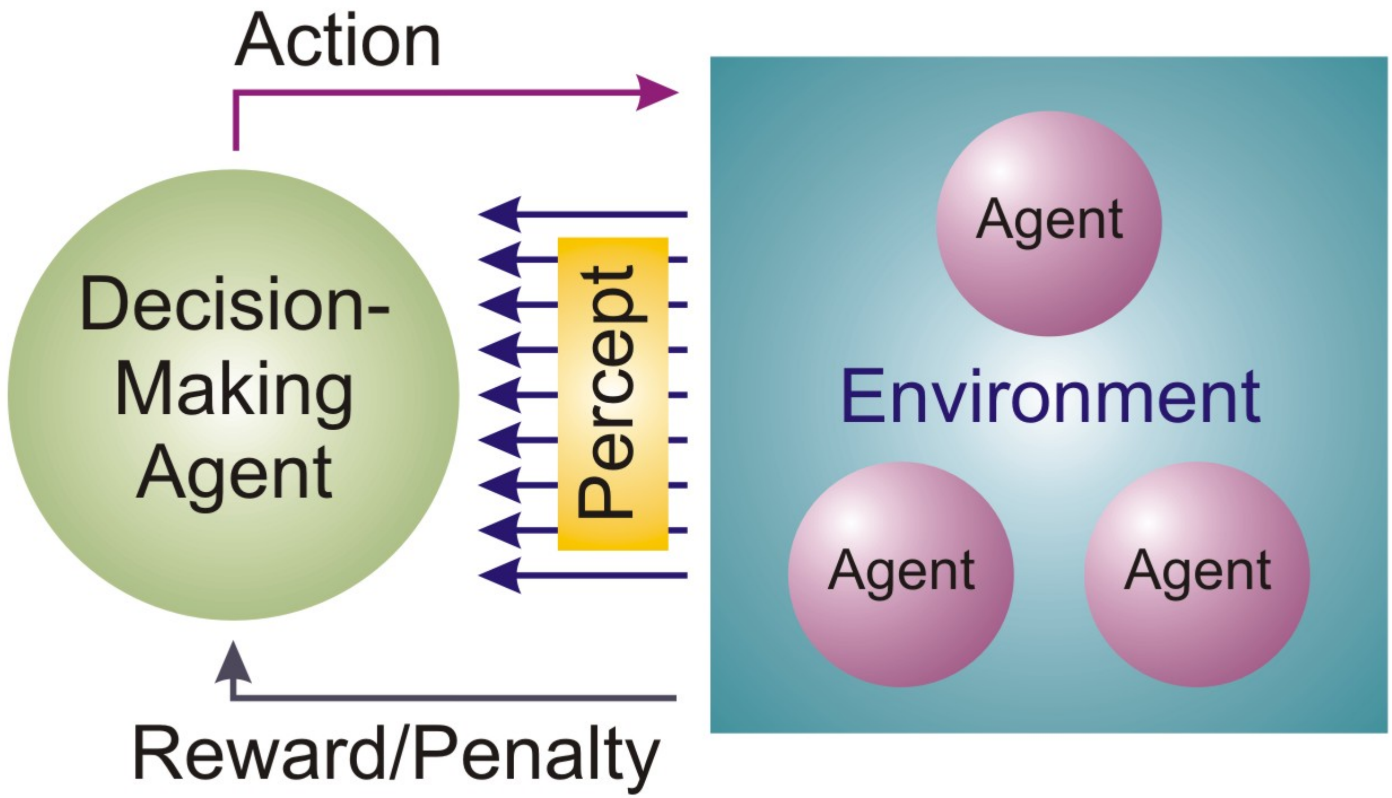
The most basic problem of **hedonic psychophysics** is the determination of the level of adaptation or aspiration that separates positive from negative outcomes. The **hedonic reference point** is largely determined by the objective status quo, but it is also affected by expectations and social comparisons. An objective improvement can be experienced as a loss, for example, when an employee receives a smaller raise than everyone else in the office. The experience of pleasure or pain associated with a change of state is also critically dependent on the **dynamics of hedonic adaptation**. Brickman and Campbell's (1971) concept of the hedonic treadmill suggests the radical hypothesis that rapid adaptation will cause the effects of any objective improvement to be short-lived.

Models of hedonic adaptation

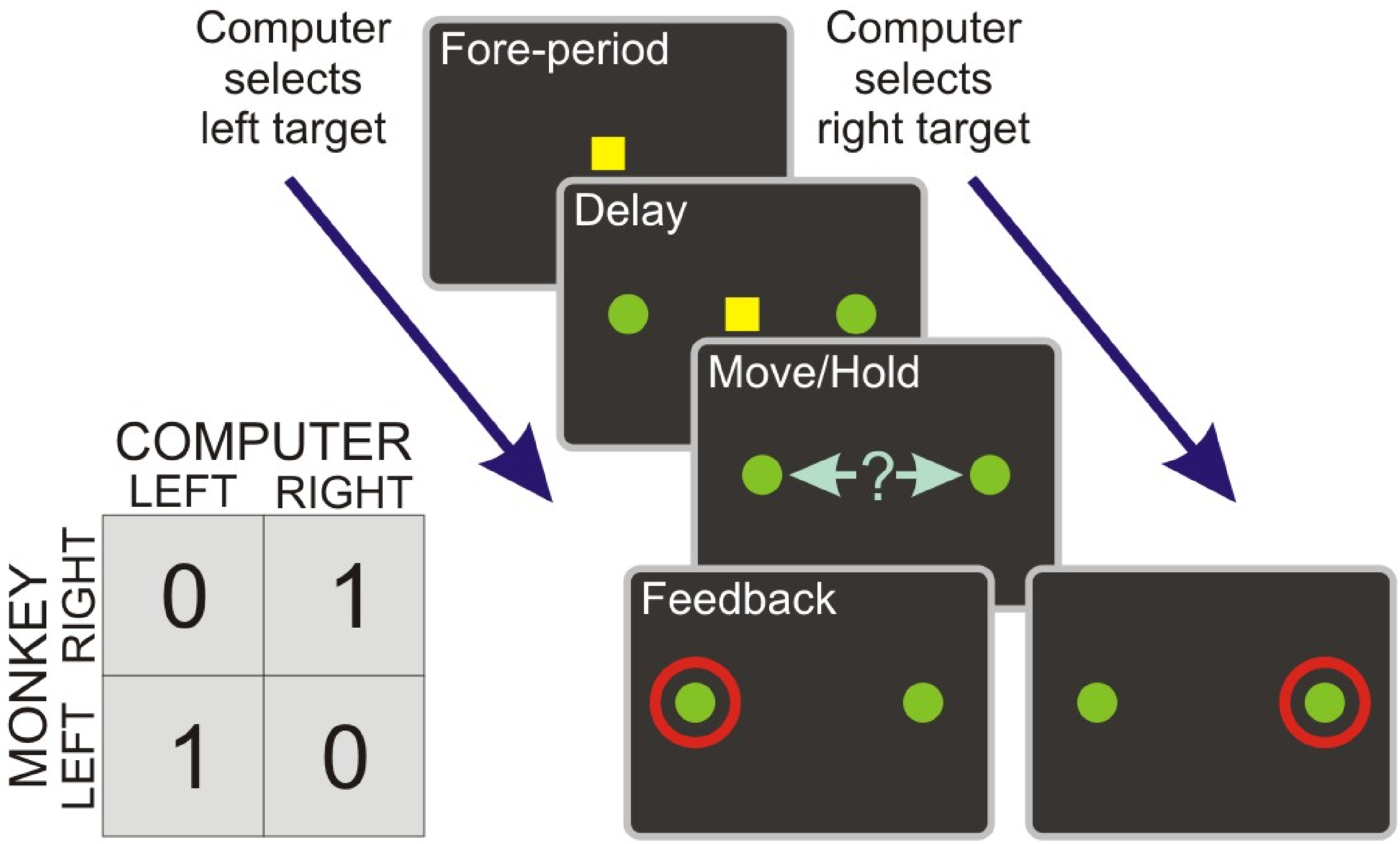


Use of competitive games to study decision making

1. Requires the subject to make stochastic choices (mixed strategy).
2. Decision making in a multi-agent environment is difficult,
   but more interesting.



Matching Pennies: Oculomotor Free-choice Task



Computer's decision-making strategy

Algorithm 0: random

Algorithm 1: biased against the monkey's choice sequence (partially exploitative)

Algorithm 2: biased against the monkey's choice+reward sequence (fully exploitative)

Even against the most exploitative opponent, the animal's choice was not completely random.

Reinforcement learning model

Action Selection:

$$\text{logit } p_t(R) \equiv \log\left(\frac{p_t(R)}{1-p_t(R)}\right) = V_t(R) - V_t(L),$$

Update Rule:

$$V_{t+1}(x) = \alpha\, V_t(x) + \Delta_t(x), \qquad \text{for } x = R \text{ or } L$$

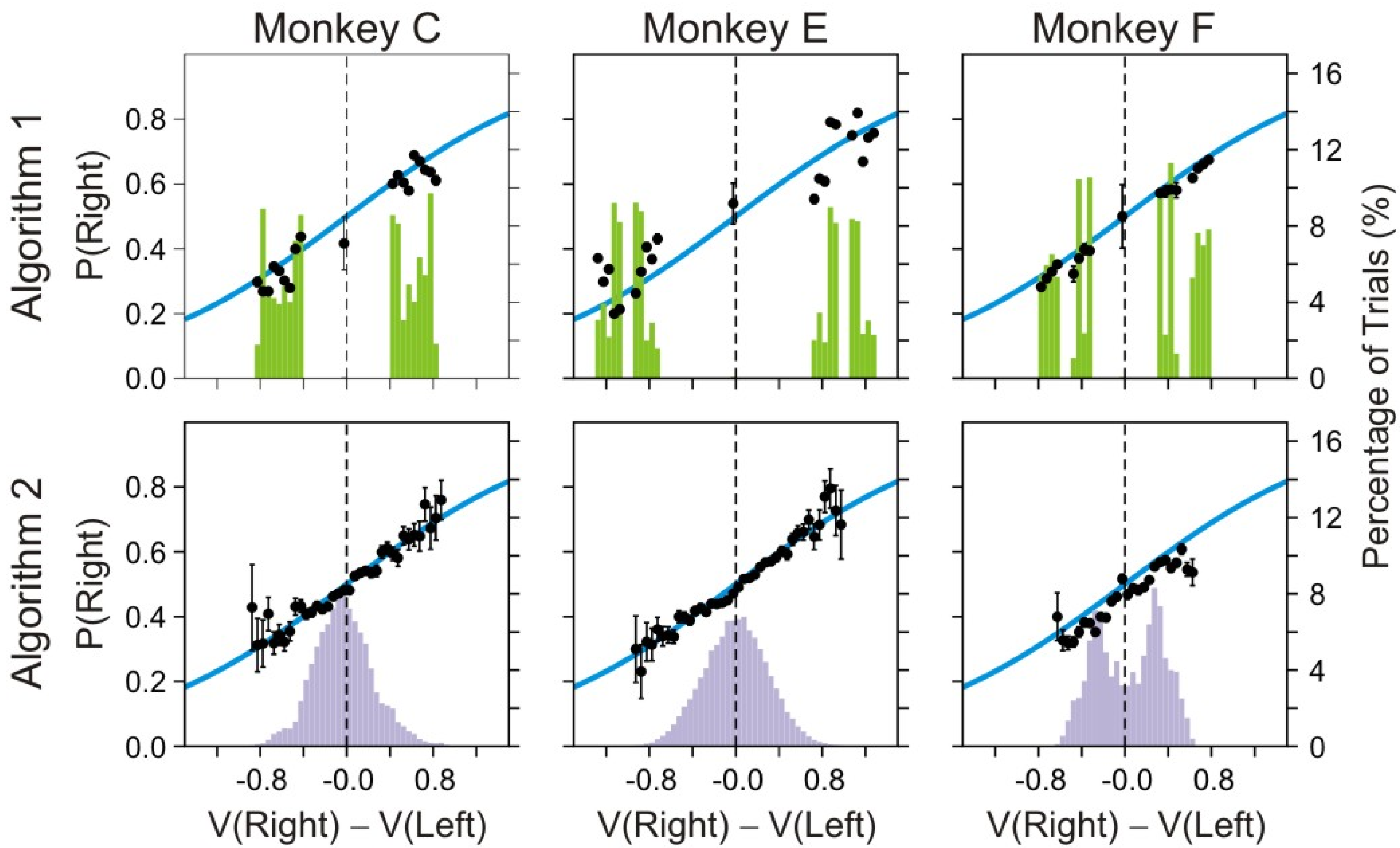
where

$\Delta_t(x) = \Delta_1$ for rewarded target

$\Delta_t(x) = \Delta_2$ for unrewarded target

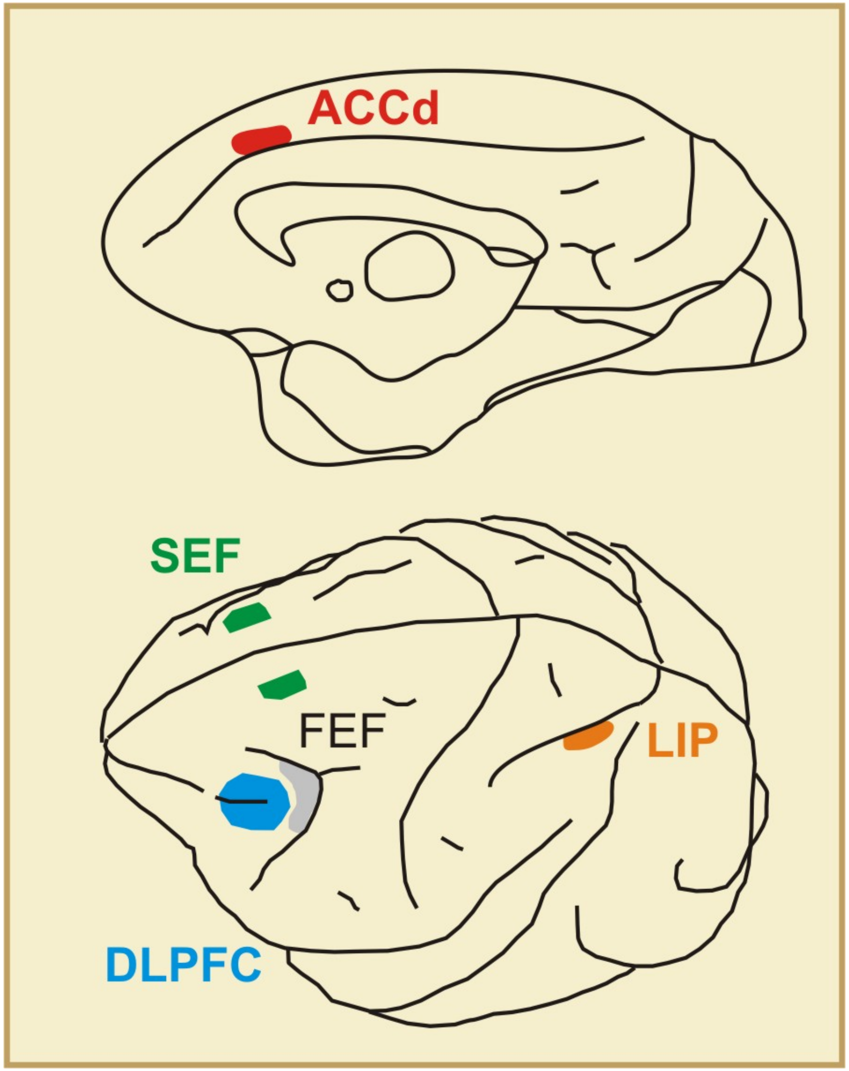
$\Delta_t(x) = 0$ for unselected target.

Reinforcement learning model: result

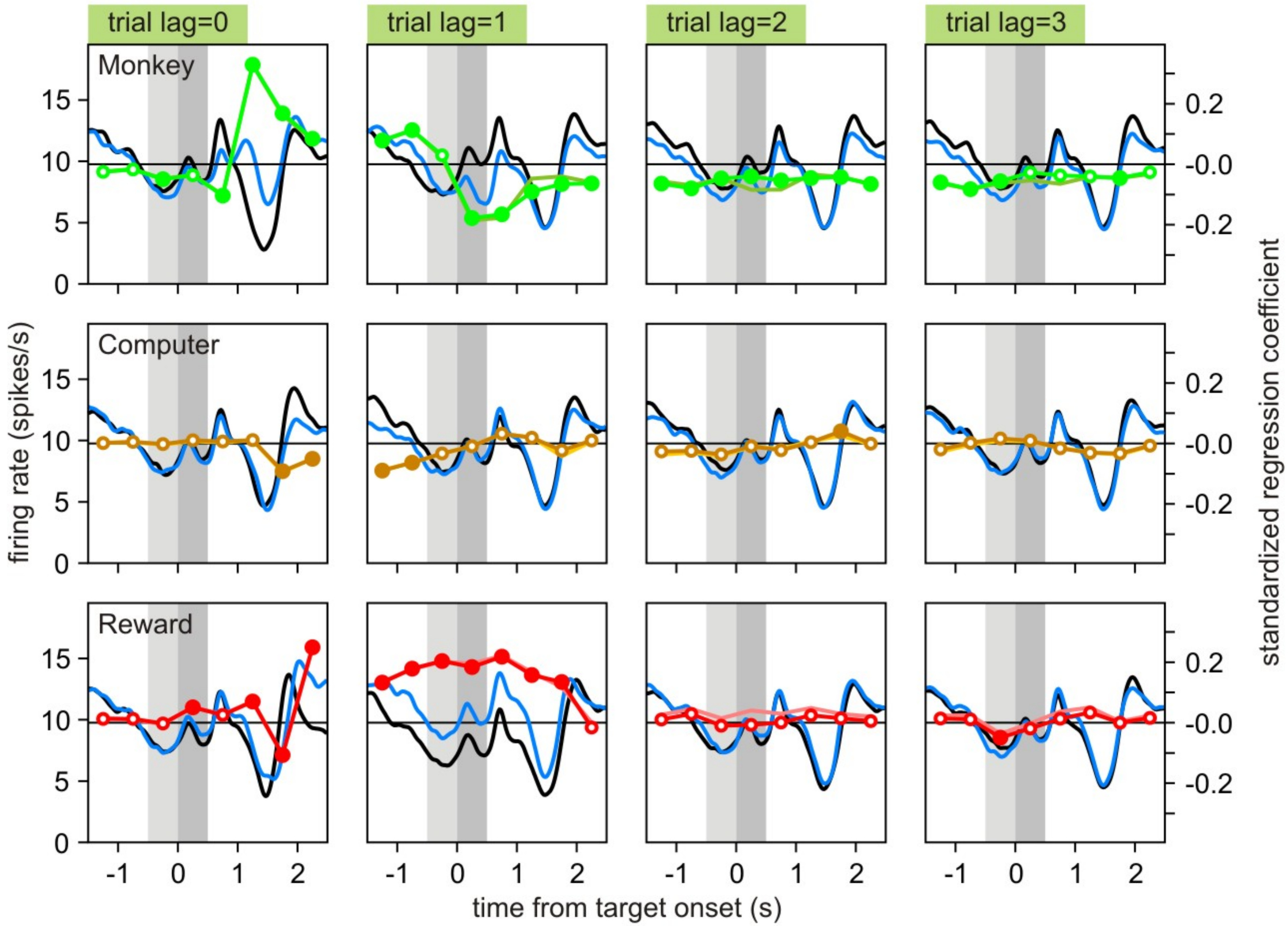


Decision-related signals in the primate cortex

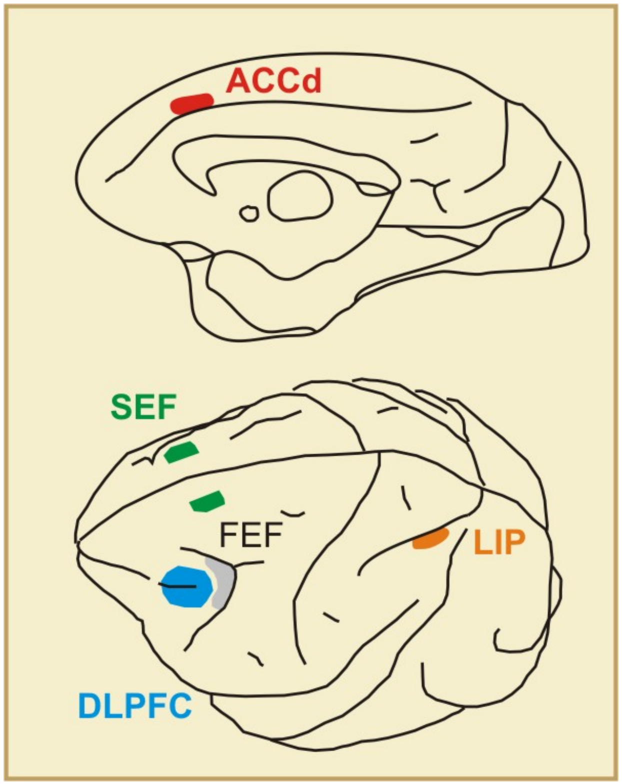
- dorso-lateral prefrontal cortex
  (DLPFC; 227 neurons)
- supplementary eye field
  (SEF; 185 neurons)
- dorsal anterior cingulate cortex
  (ACCd, area 24c;
  154 neurons)
- lateral intraparietal area
  (LIP; 108 neurons)

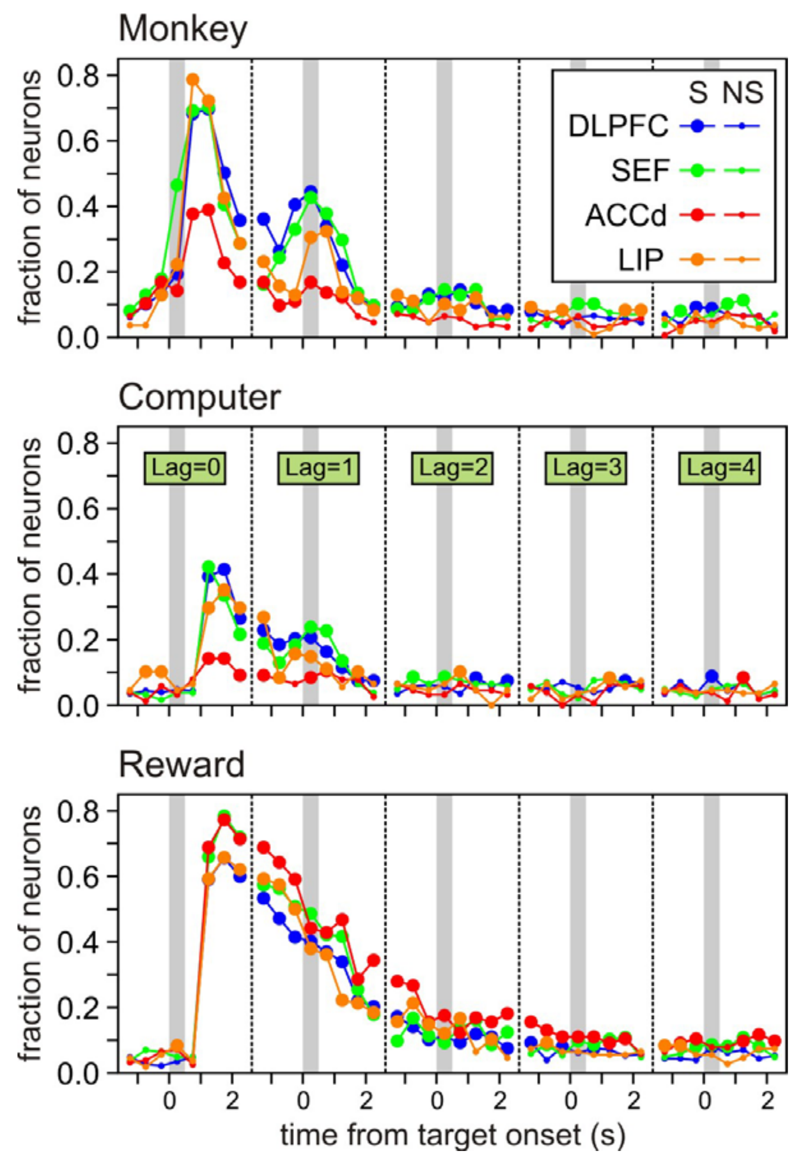


Temporal integration of choice and reinforcement signals in DLPFC

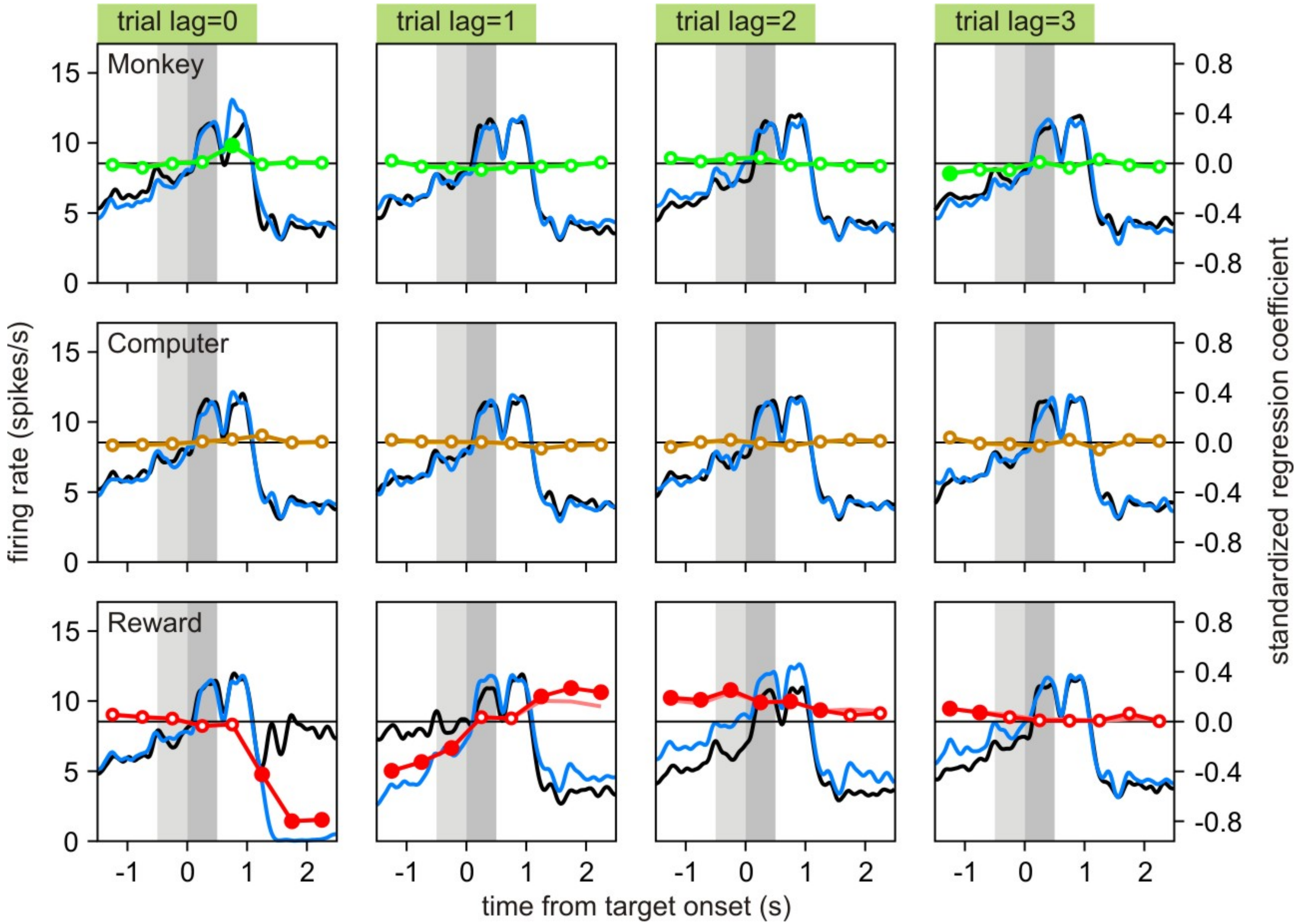


Temporal integration of choice and reinforcement signals in PFC & PPC.





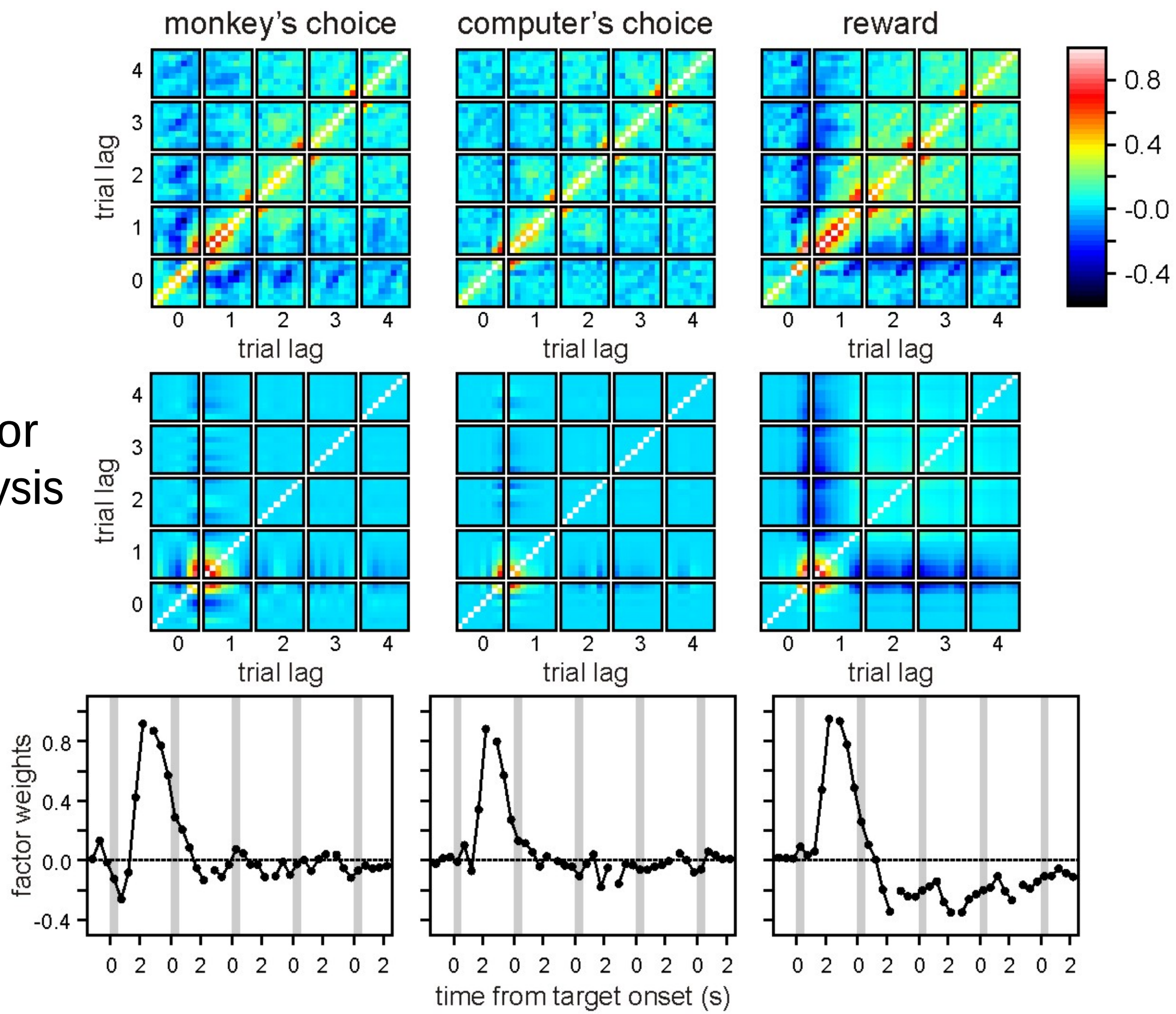
Temporal integration of reinforcement signals in ACCd



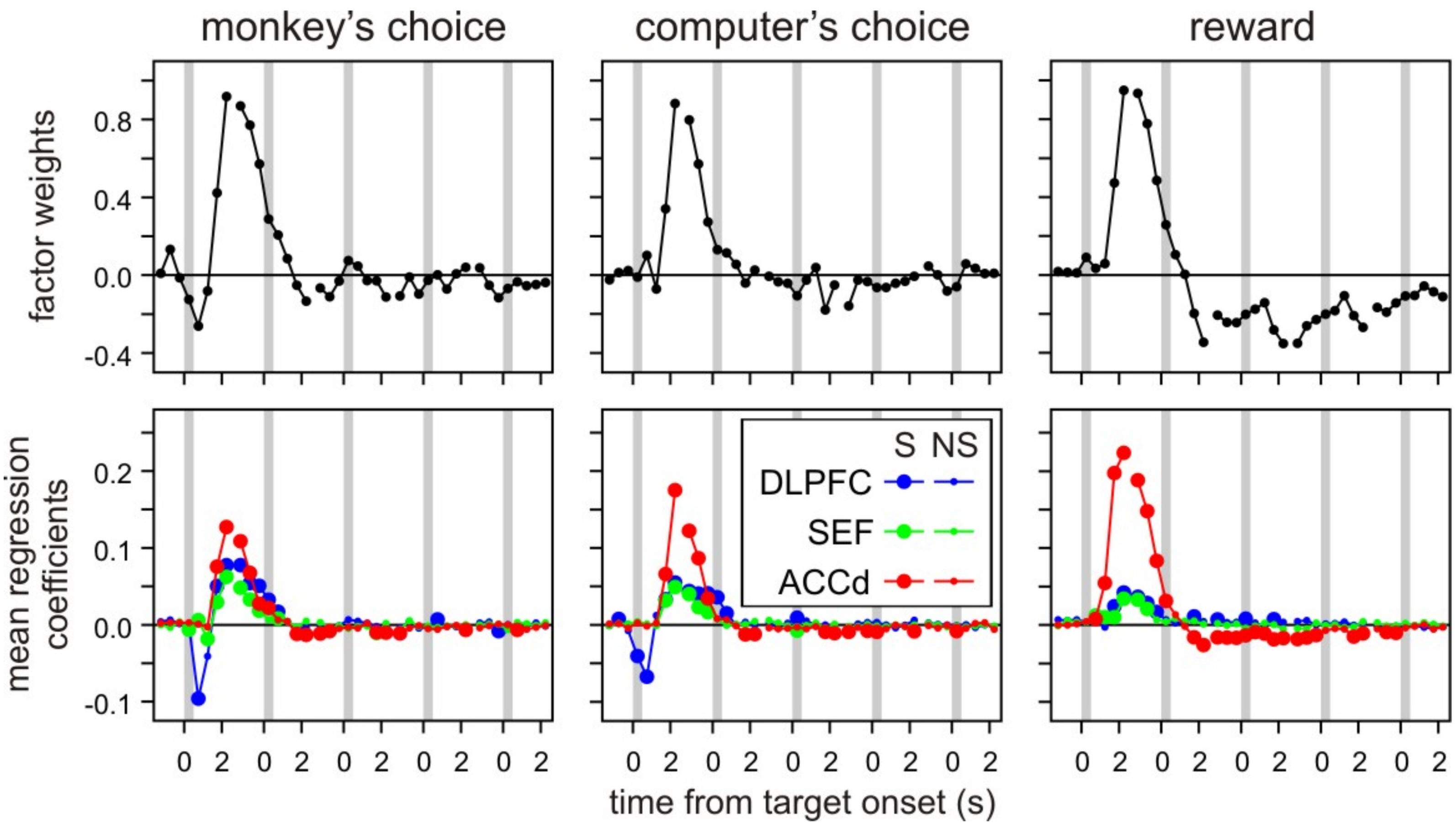
Neural activity is biphasically modulated by reward:
correlation coefficient matrix for regression coefficients (all areas)

Data

Single-factor
Factor Analysis
Model



Reward signals in ACCd display a robust biphasic pattern.



Summary & Conclusions

1. Decisions during a competitive game are biased according to the value functions of alternative choices, as predicted by a reinforcement learning algorithm.

2. Choice and reward signals necessary to update value functions are broadly distributed in the primate prefrontal and parietal cortex.

3. Biphasic reward signals found in the dorsal anterior cingulate cortex might underlie the transformation of physical reward to subjective value, and thereby influence the process of learning optimal decision-making strategies.