Predictive Hebbian Learning

Computational Models of Neural Systems

Lecture 5.2

David S. Touretzky

Based on slides by Mirella Lapata

November, 2017

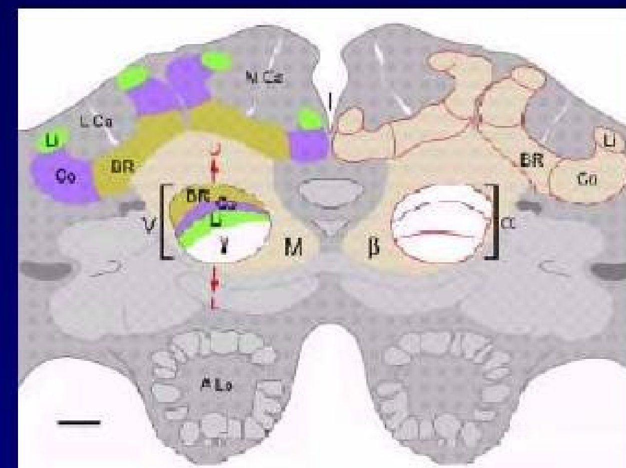
Outline

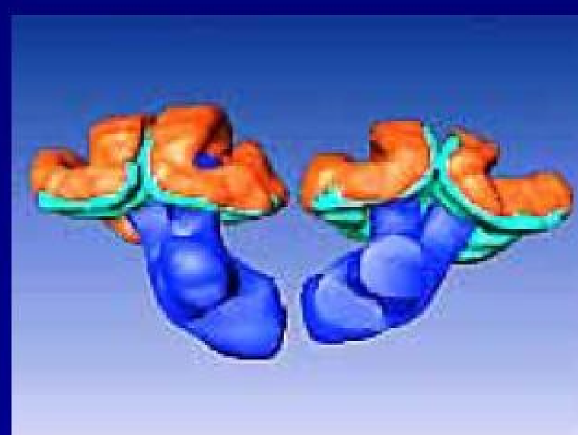
- The bee brain
- Clasical conditioning in honeybees
  - identification of VUMmx1
  - properties of VUMmx1
- Bee foraging in uncertain environments
  - model of bee foraging
  - theory of predictive Hebbian learning
- Dopamine neurons in the macaque monkey
  - activity of dopamine neurons
  - generalized theory of predictive Hebbian learning
  - modeling predictions

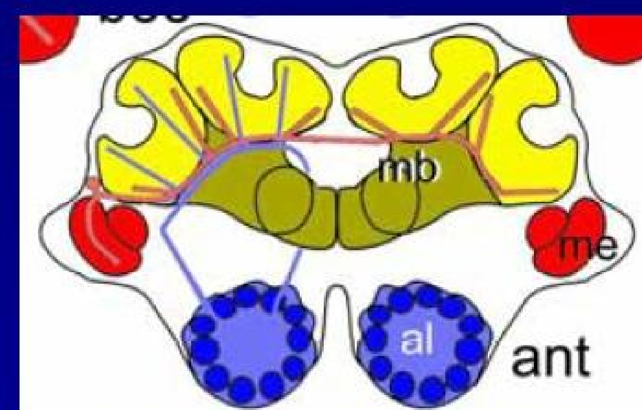
The Bee Brain

- Honeybees have about one million neurons in about 1 mm^3.
  - Fruit flies have only about 100,000 neurons
  - Ants have about 250,000 neurons.
- The mushroom bodies are thought to be involved in learning and memory.

Where is memory located in the honey bee brain?



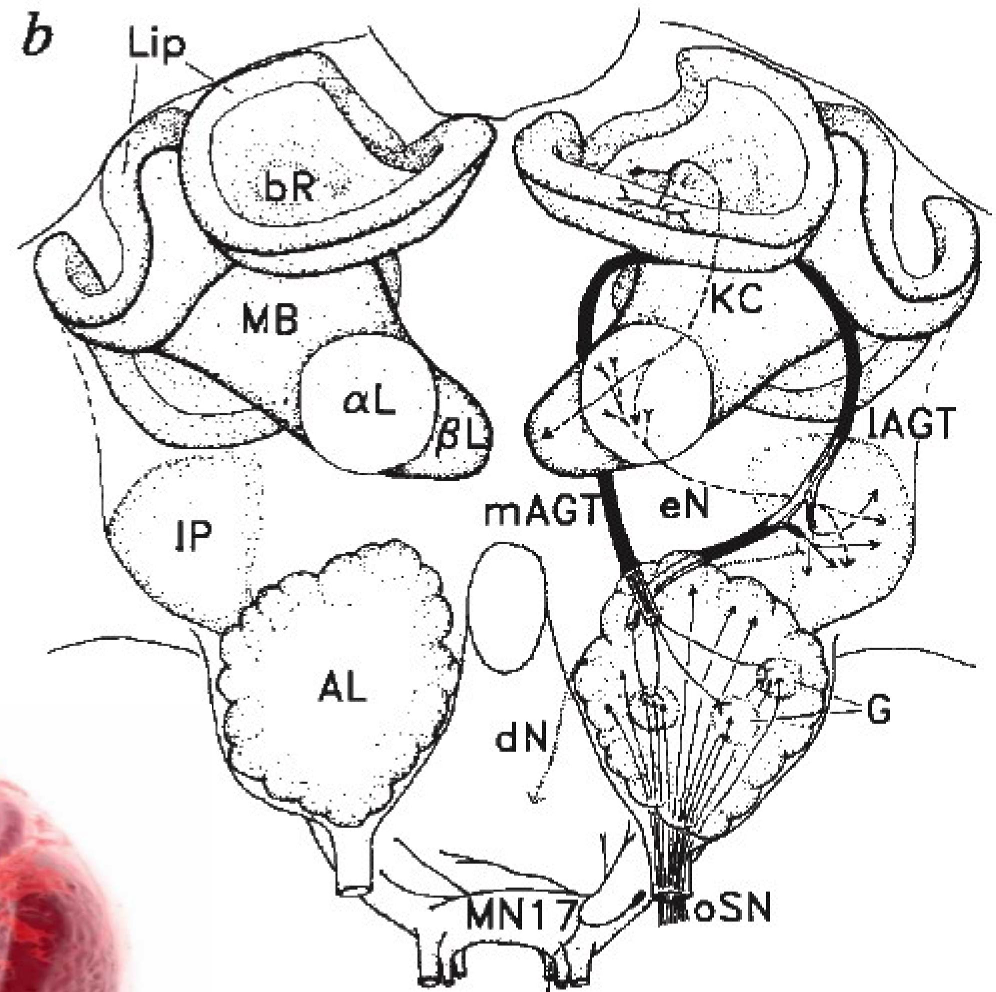


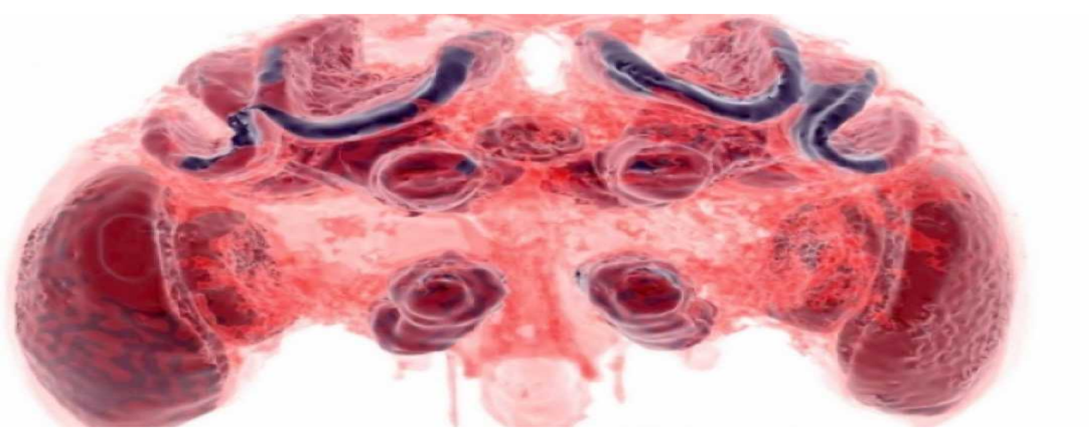


http://web.neurobio.arizona.edu/gronenberg/nrsc581

Anatomy of the Bee Brain

- MB: Mushroom body
- AL: Antenna lobe
- KC: Kenyon cells
- oSN: Olfactory sensory neurons
- MN17: motor neuron involved in PER





Questions

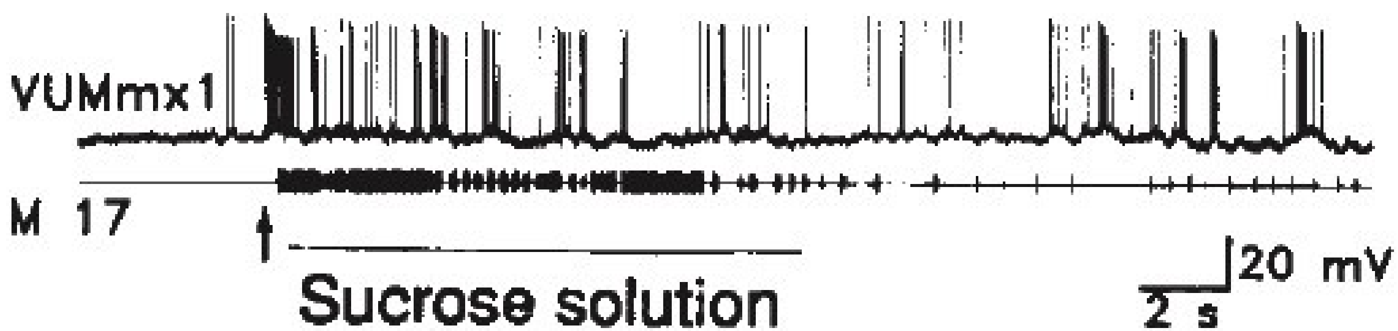
- What are the cellular mechanisms responsible for classical conditioning?
- How is information about the unconditioned stimulus (US) represented at the neuronal level?
- What are the properties of neurons mediating the US?
  - Response to US
  - Convergence with the conditioned stimulus (CS) pathway
  - Reinforcement in conditioning
- How to identify such neurons?

Experiments on Honeybees

- Bees fixed by waxing dorsal thorax to small metal table.
- Odors were presented in a gentle air stream.
- Sucrose solution applied briefly to antenna and proboscis.
- Proboscis extension was seen after a *single pairing* of the odor (CS) with sucrose (US).

Measuring Responses

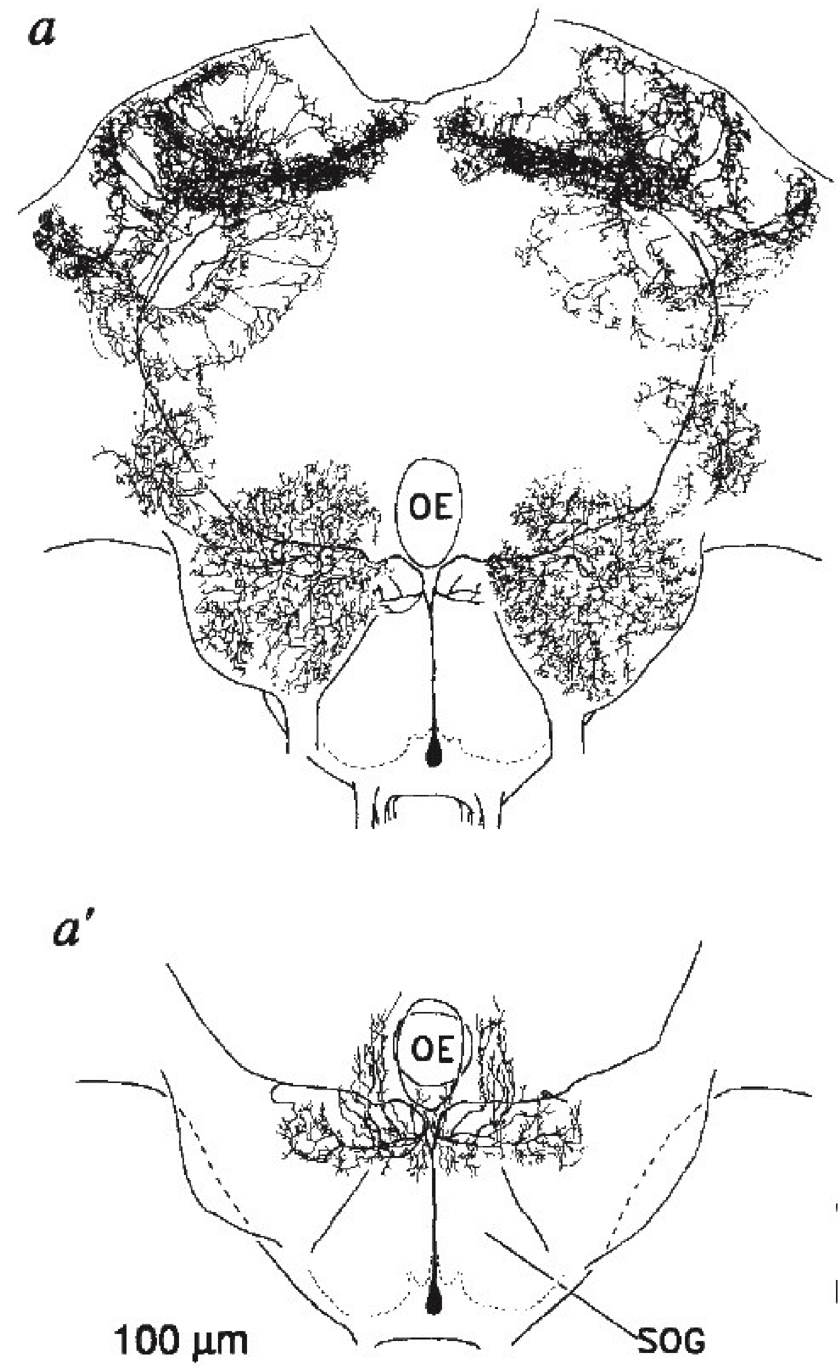
- Proboscis extension reflex (PER) was recorded as an electromyogram from the M17 muscle involved in the reflex.
- Neurons were tested for responsiveness to the US.



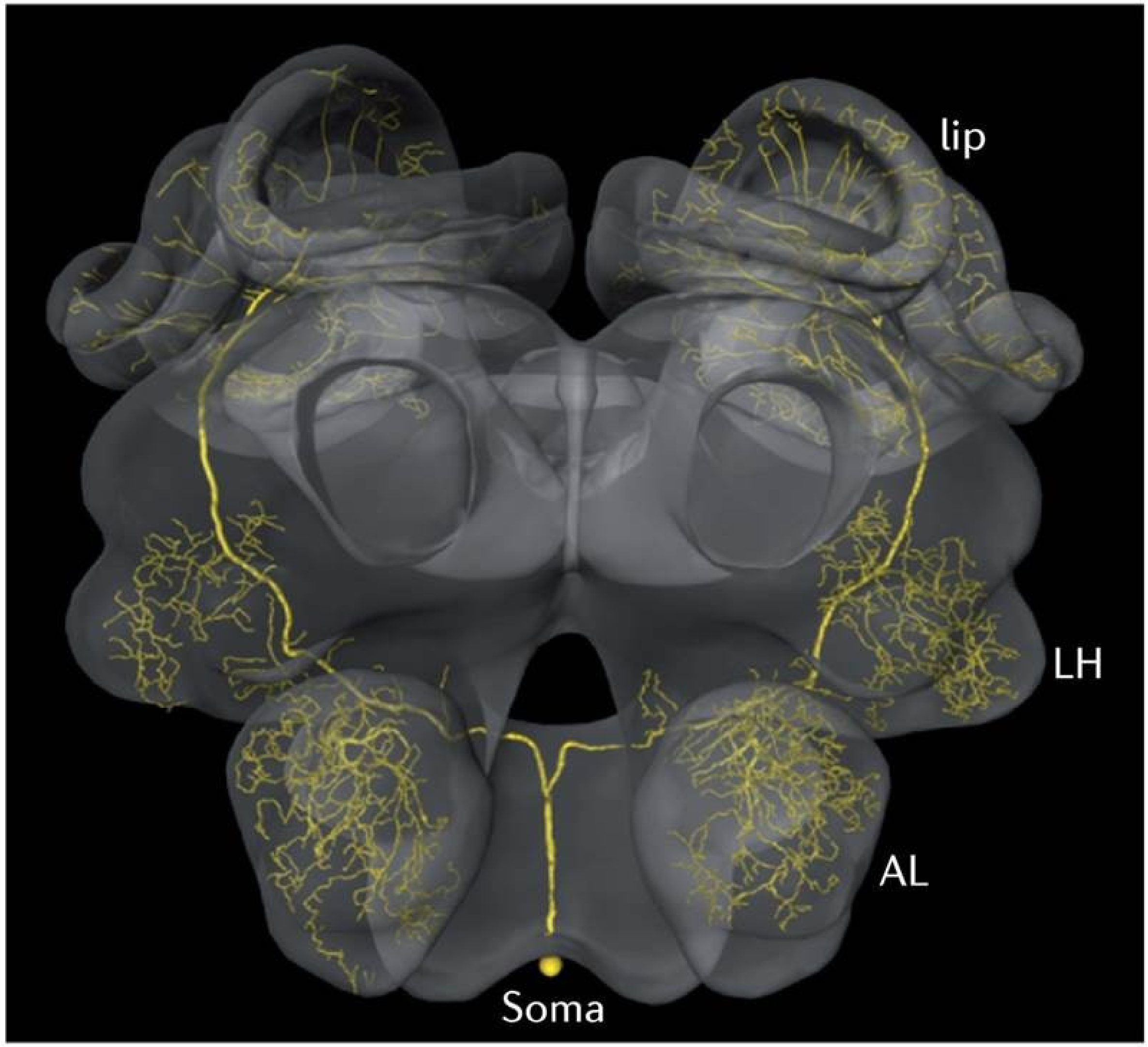
VUMmx1 Responds to US

- Unique morphology: arborizes in the suboesophageal ganglion (SOG) and projects widely in regions involved in odor (CS) processing
- Responds to sucrose with a long burst of action potentials which outlasts the sucrose US.
- Neurotransmitter is octopamine: related to dopamine.

OE = Oesophagus



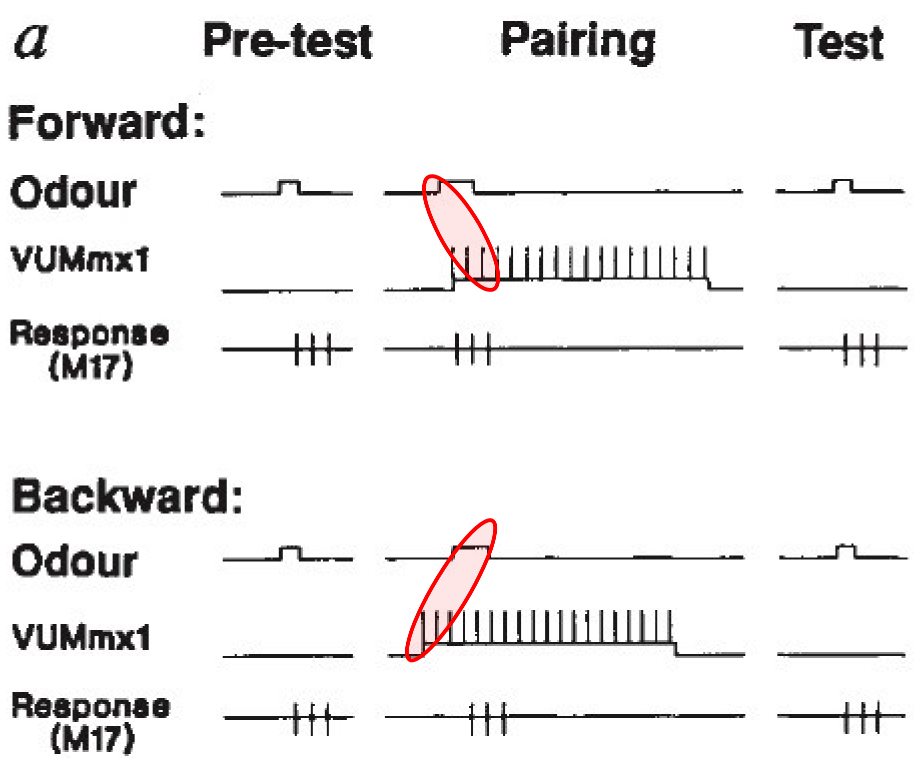
VUMmx1



Nature Reviews | Neuroscience

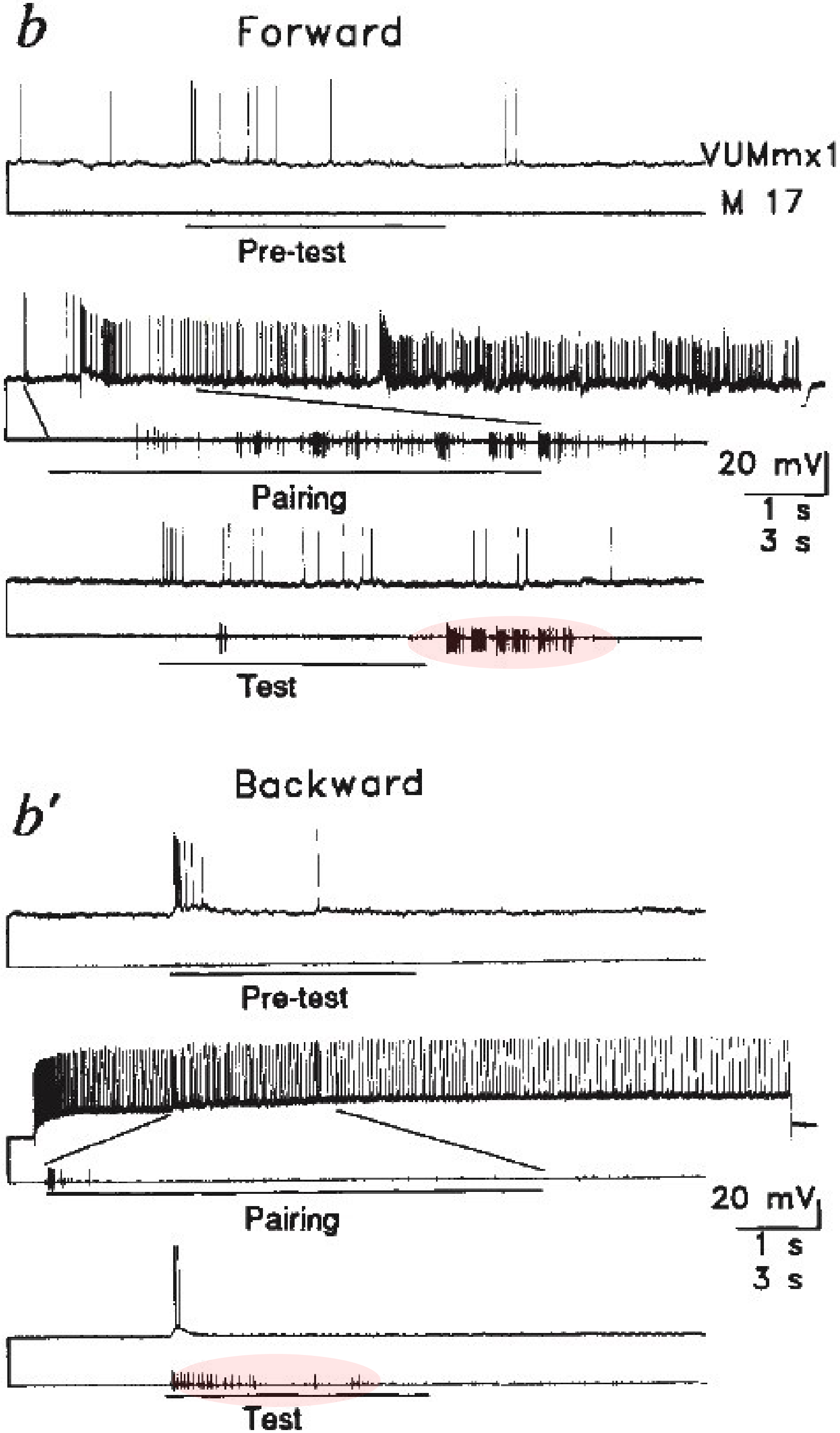
Stimulating VUMmx1 Simulates a US

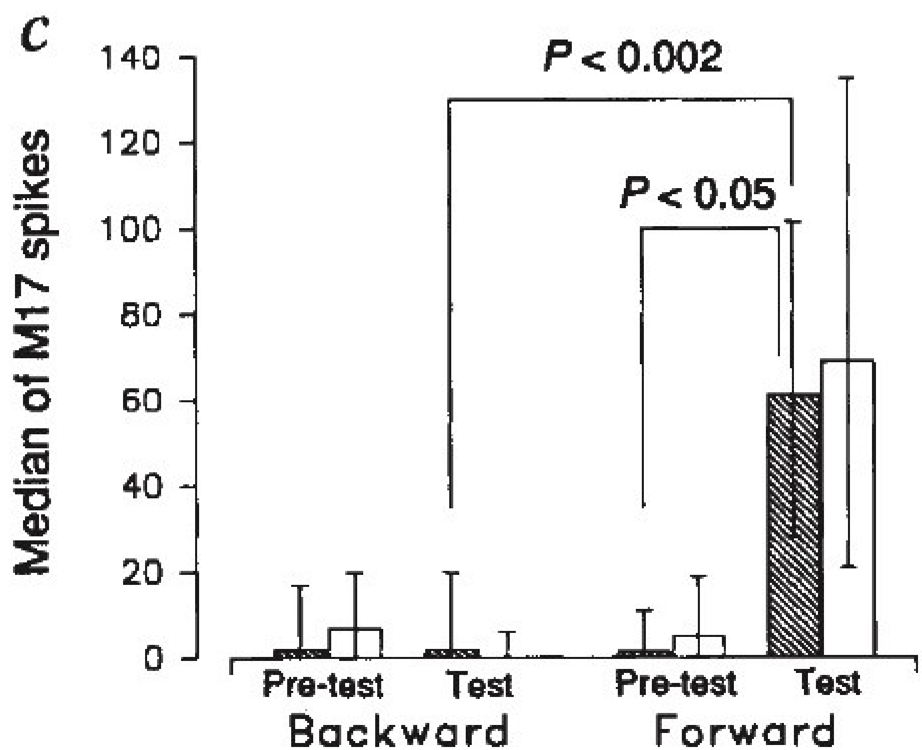
- Introduce CS then inject depolarizing current into VUMmx1 in lieu of applying sucrose.
- Try both forward and backward conditioning paradigms.



Schematic diagram.

Not real data!

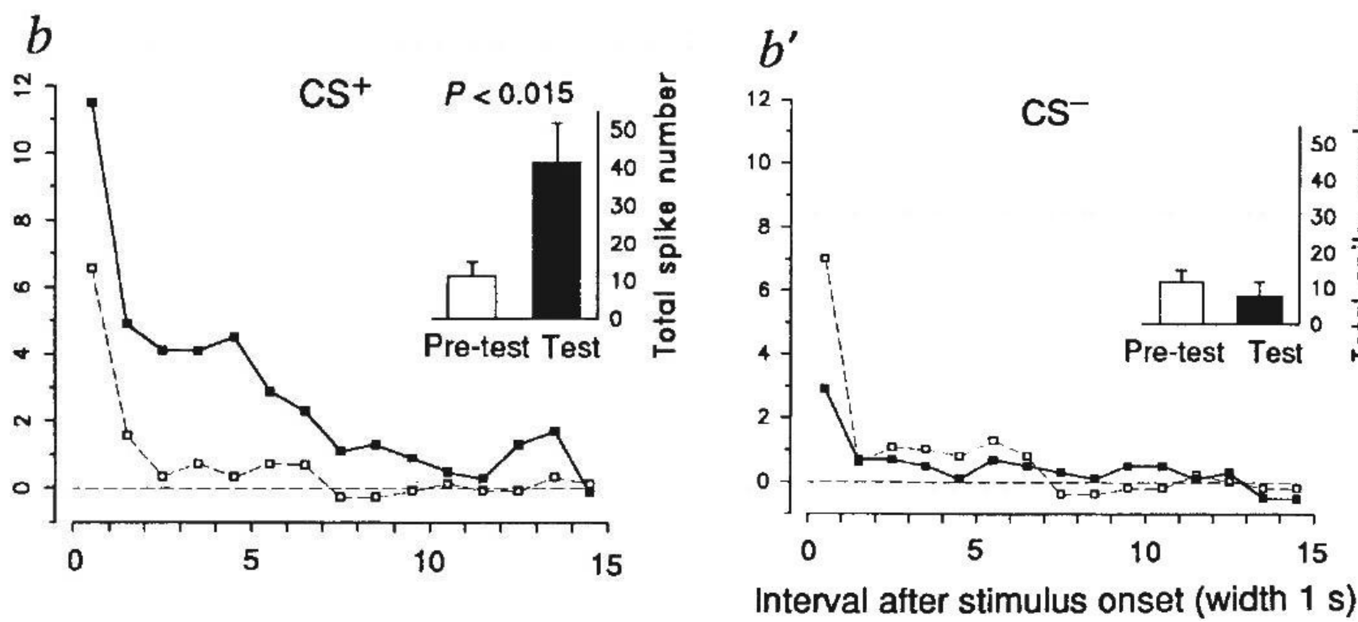




Open bars: sucrose US
Shaded bars: VUMmx1 stimulation

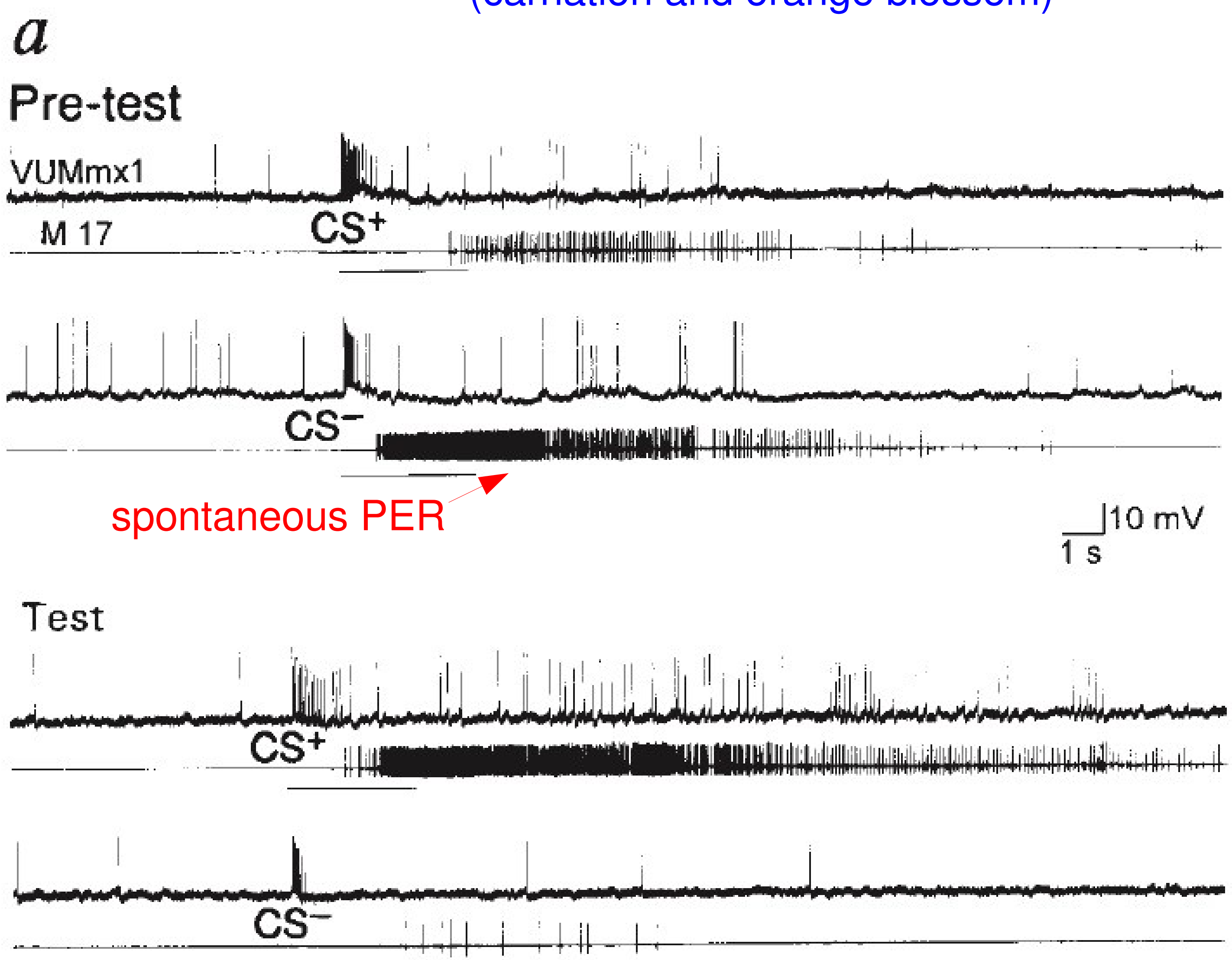
Learning Effects of VUMmx1 Stimulation

- After learning, the odor alone stimulates VUMmx1 activity.
- Temporal contiguity effect: forward pairing causes a larger increase in spiking than backward pairing.
- Differential conditioning effect:
  - Differentially conditioned bees respond strongly to an odor (CS+) specifically paired with the US, and significantly less to an unpaired odor (CS–).



Differential Conditioning of Two Odors

(carnation and orange blossom)



Discussion

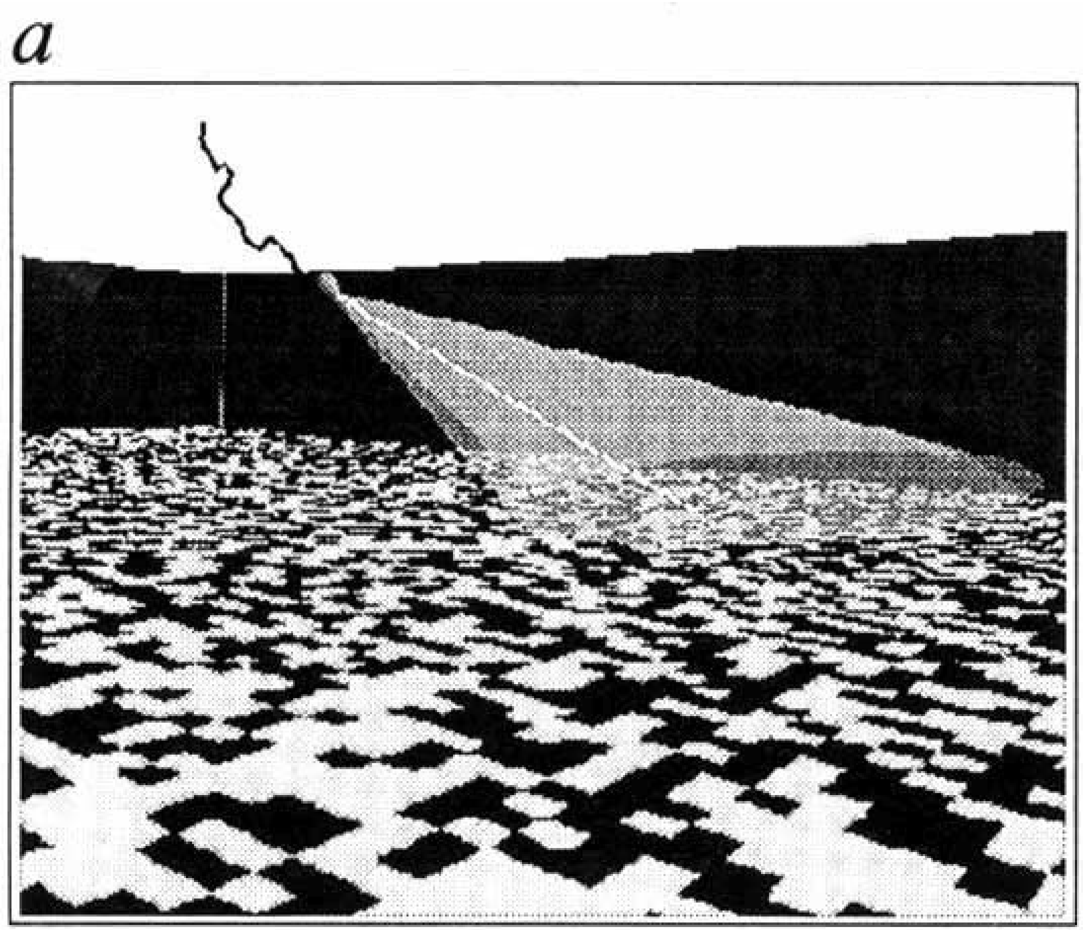
- Main claims:
  - VUMmx1 mediates the US in associative learning
  - A learned CS also activates VUMmx1.
  - Physiology is compatible with structures involved in complex forms of learning.
- Questions:
  - Is VUMmx1 the only neuron mediating the US?
    - Serial homologue of VUMmx1 has almost identical branching pattern.
    - Response to electrical stimulation is less than response to sucrose, so perhaps other neurons also contribute to the US signal.
  - Can VUMmx1 mediate other conditioning phenomena, e.g., blocking, overshadowing, extinction?
  - Do different stimuli induce similar responses?

Bee Foraging

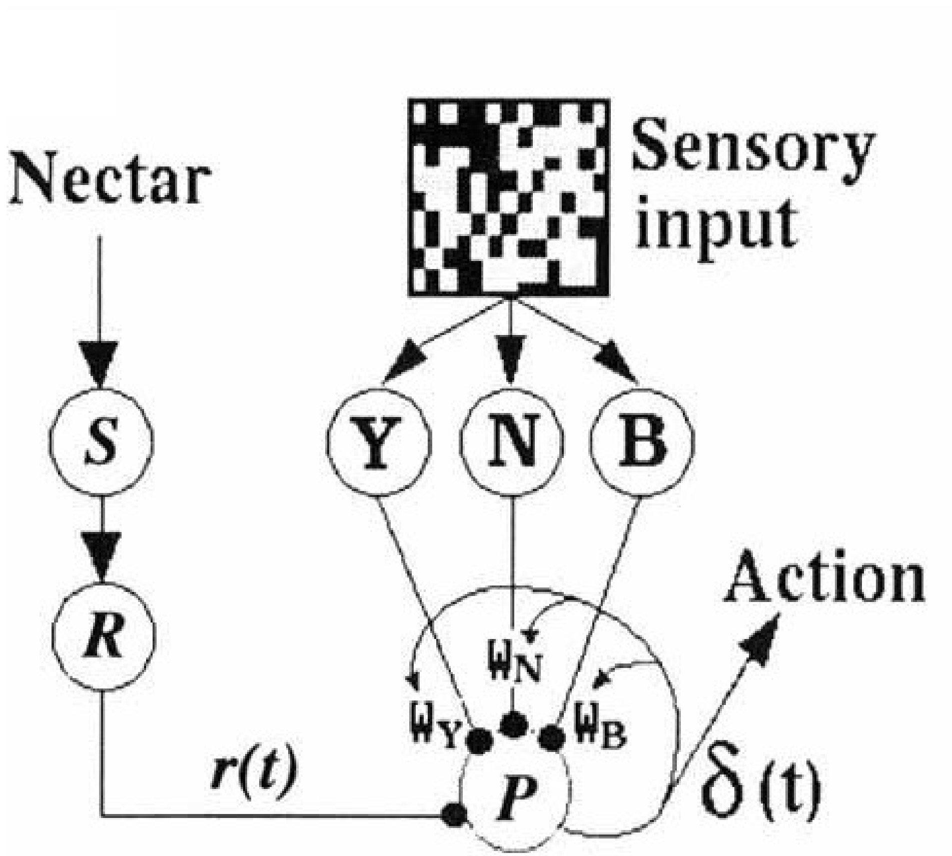
- Real's (1991) experiment:
  - Bumblebees foraged on artificial blue and yellow flowers.
  - Blue flowers contained 2 μl of nectar.
  - Yellow flowers contained 6 μl in one third of the flowers and no nectar in the remaining two thirds.
  - Blue and yellow flowers contained the same *average* amount of nectar.
- Results:
  - Bees favored the constant blue over the variable yellow flowers even though the mean reward was the same.
  - Bees forage equally from both flower types if the mean reward from yellow is made sufficiently large.

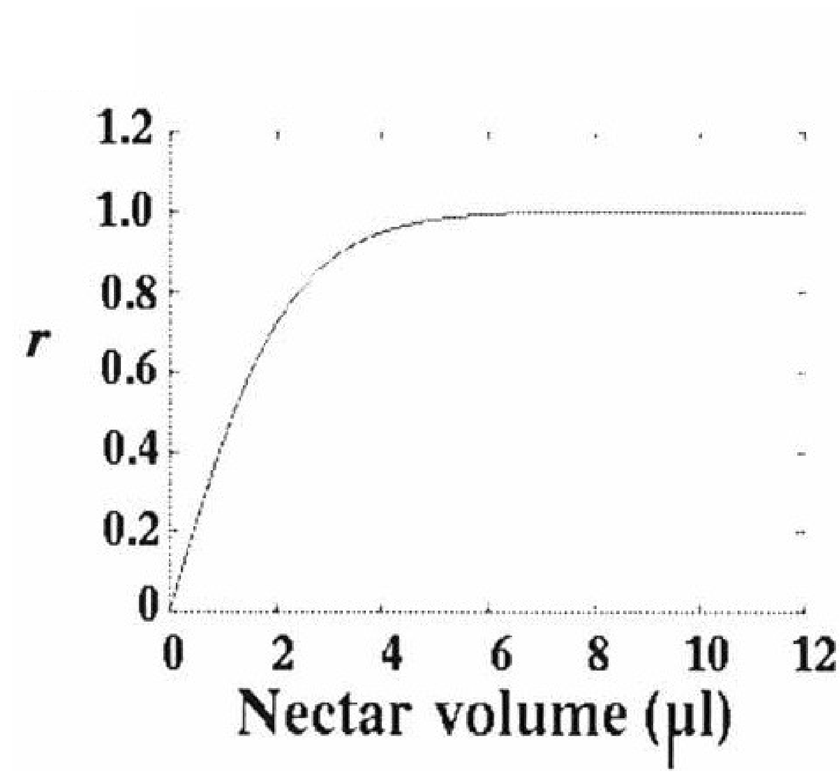
Montague, Dayan, and Sejnowski (1995)

- Model of bee foraging behavior based on VUMmx1.
- Bee decides at each time step whether to randomly reorient.



Neural Network Model





S: sucrose sensitive neuron; R: reward neuron;
P: reward predicting neuron; δ: prediction error signal

TD Equations

$$\delta(t) = r(t) + \gamma V(t) - V(t-1)$$

Let $\gamma = 1$: no discounting

¿

$$\begin{aligned}\delta(t) &= r(t) + V(t) - V(t-1) \\ &= r(t) + \dot{V}(t)\end{aligned}$$

$$V(t) = \sum_i w_i x_i(t)$$

$$\begin{aligned}\dot{V}(t) &= \sum_i w_i \left[x_i(t) - x_i(t-1) \right] \\ &= \sum_i w_i \dot{x}_i(t)\end{aligned}$$

$$\delta(t) = r(t) + \sum_i w_i \dot{x}_i(t)$$

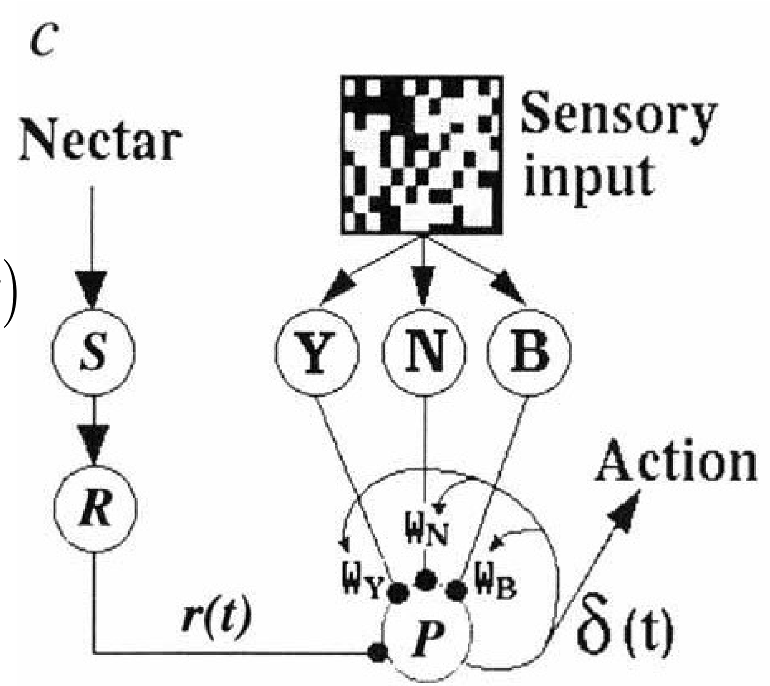
Bee Foraging Model

x_Y, x_B, x_N encode change in scene

$$\dot{V}(t) = w_b x_b(t) + w_y x_y(t) + w_n x_n(t)$$

$$\delta(t) = r(t) + \dot{V}(t)$$

$$\Delta w_i(t) = \lambda x_i(t-1) \cdot \delta(t)$$



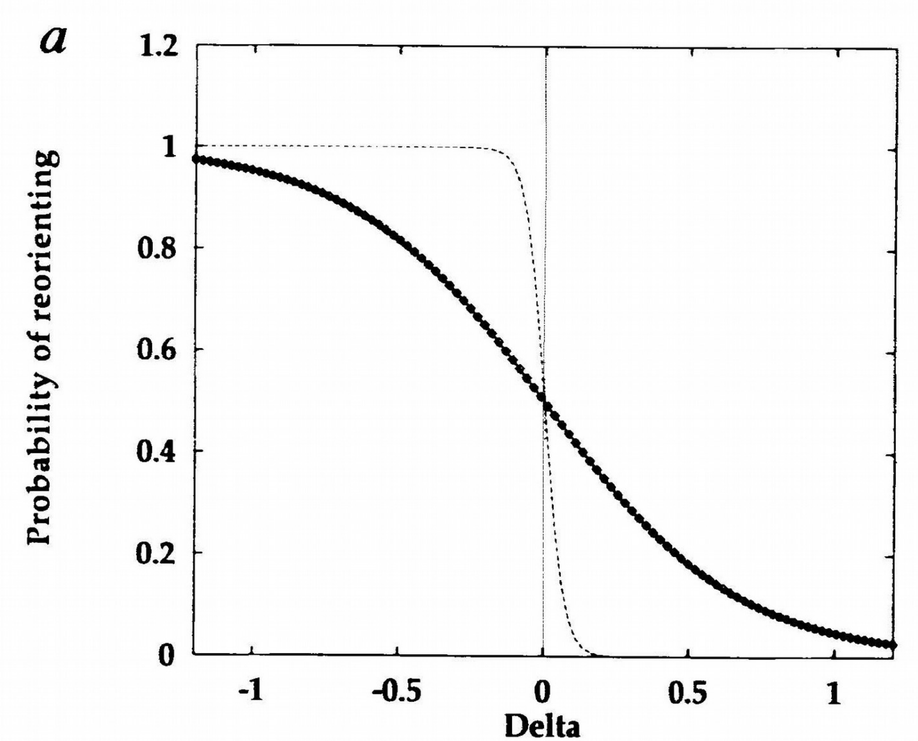
Parameters

w_B and w_Y are adaptable; w_N fixed at -0.5

Probability of reorienting: $P_r(\delta(t)) = \dfrac{1}{1+\exp(m\cdot\delta(t)+b)}$

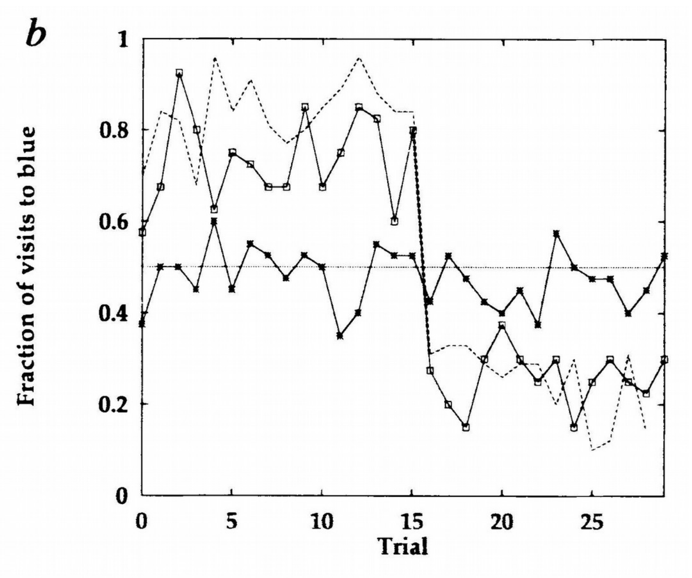
Learning rate $\lambda = 0.9$

Volume of nectar reward determined by empirically derived utility curve.

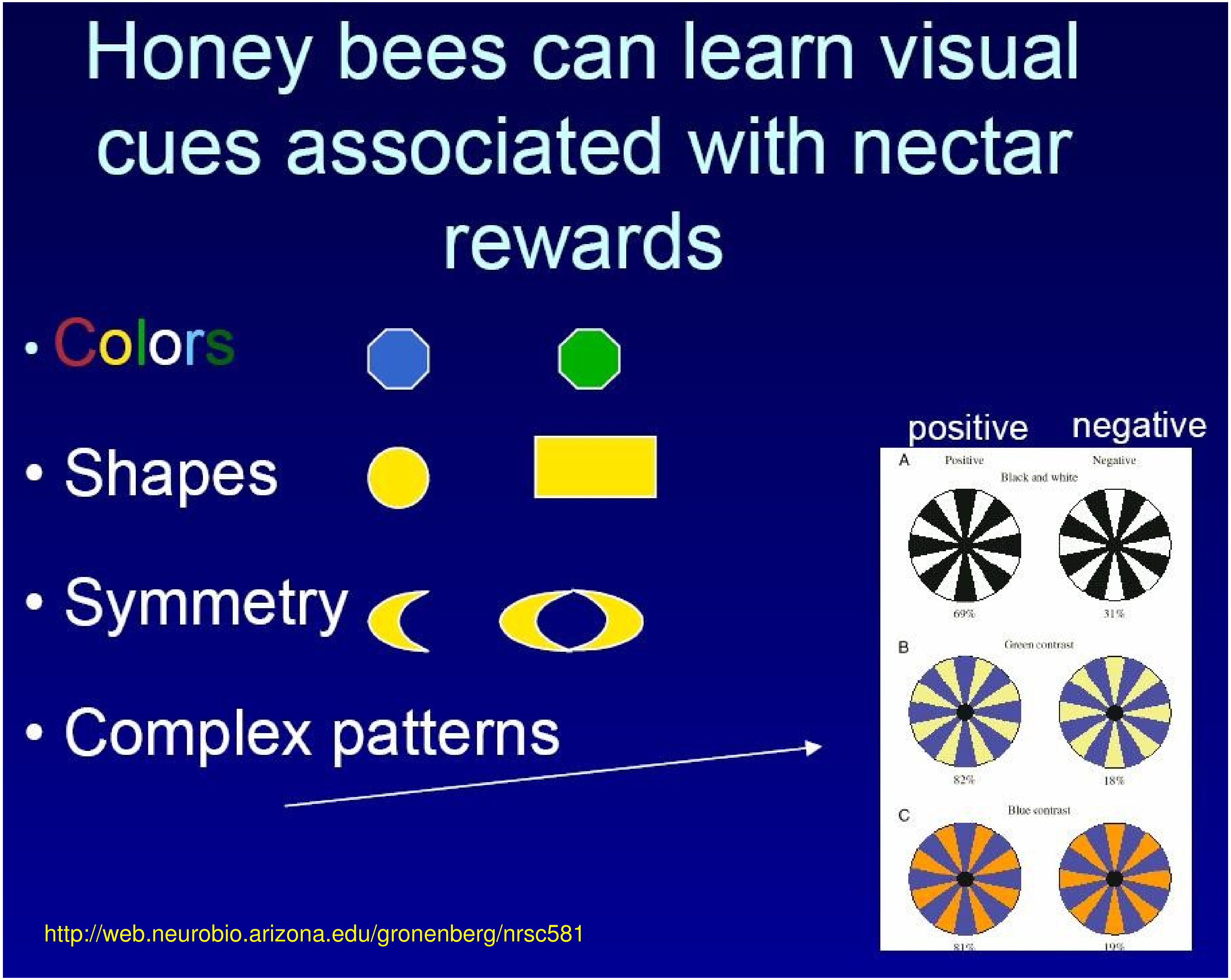


Theoretical Idea

- Unit P is analogous to VUMmx1.
- Nectar r(t) represents the reward, which can vary over time.
- At each time t, δ(t) determines the bee's next action: continue on present heading, or reorient.
- Weights are adjusted on encounters with flowers: they are updated according to the nectar reward.
- Model best matches the bee when λ = 0.9.
- Graph shows bee response to switch in contingencies on trial 15.

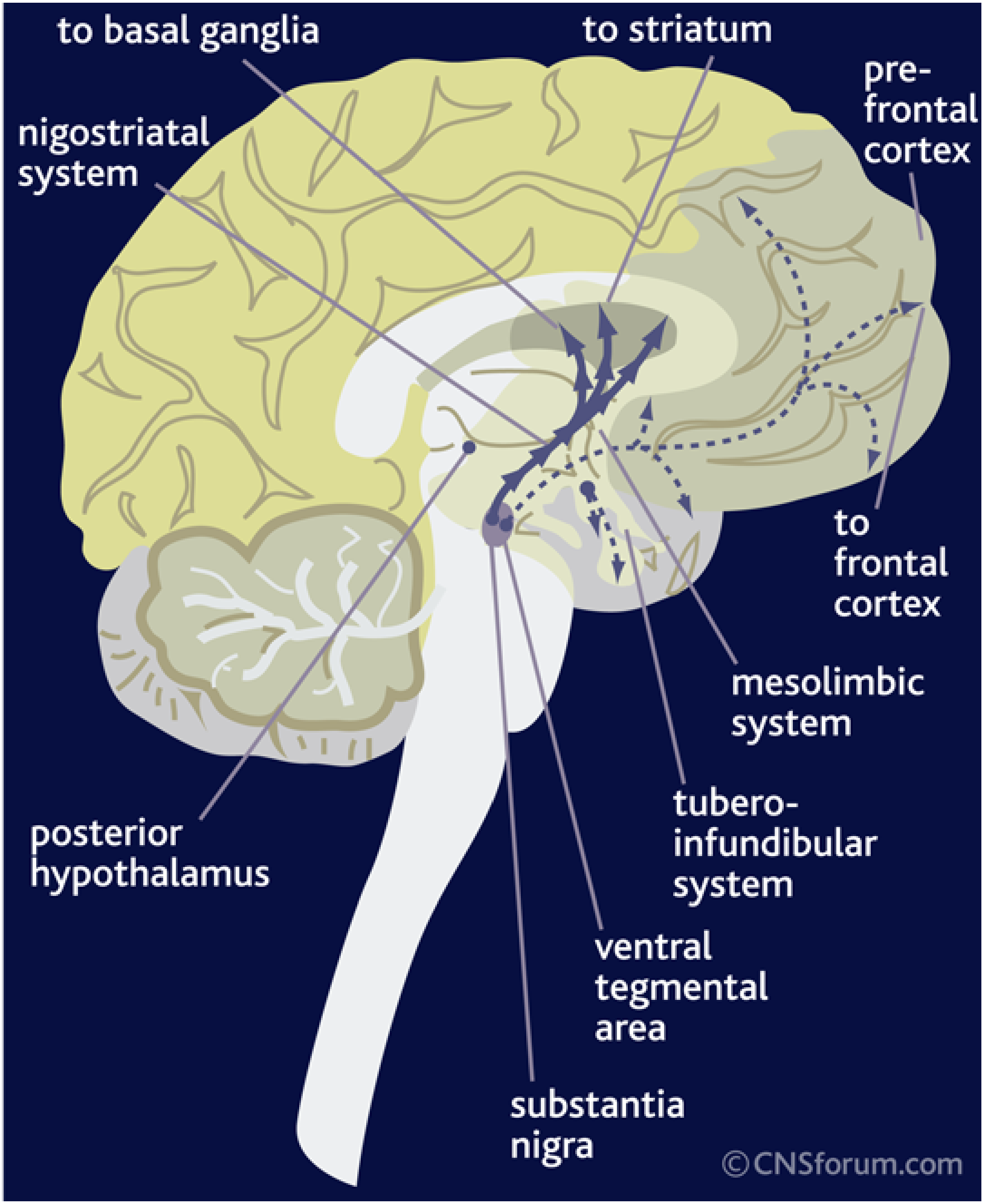


An Aside: Honeybee Operant Learning



Dopamine

- Involved in:
  - Addiction
  - Self-stimulation
  - Learning
  - Motor actions
  - Rewarding situations

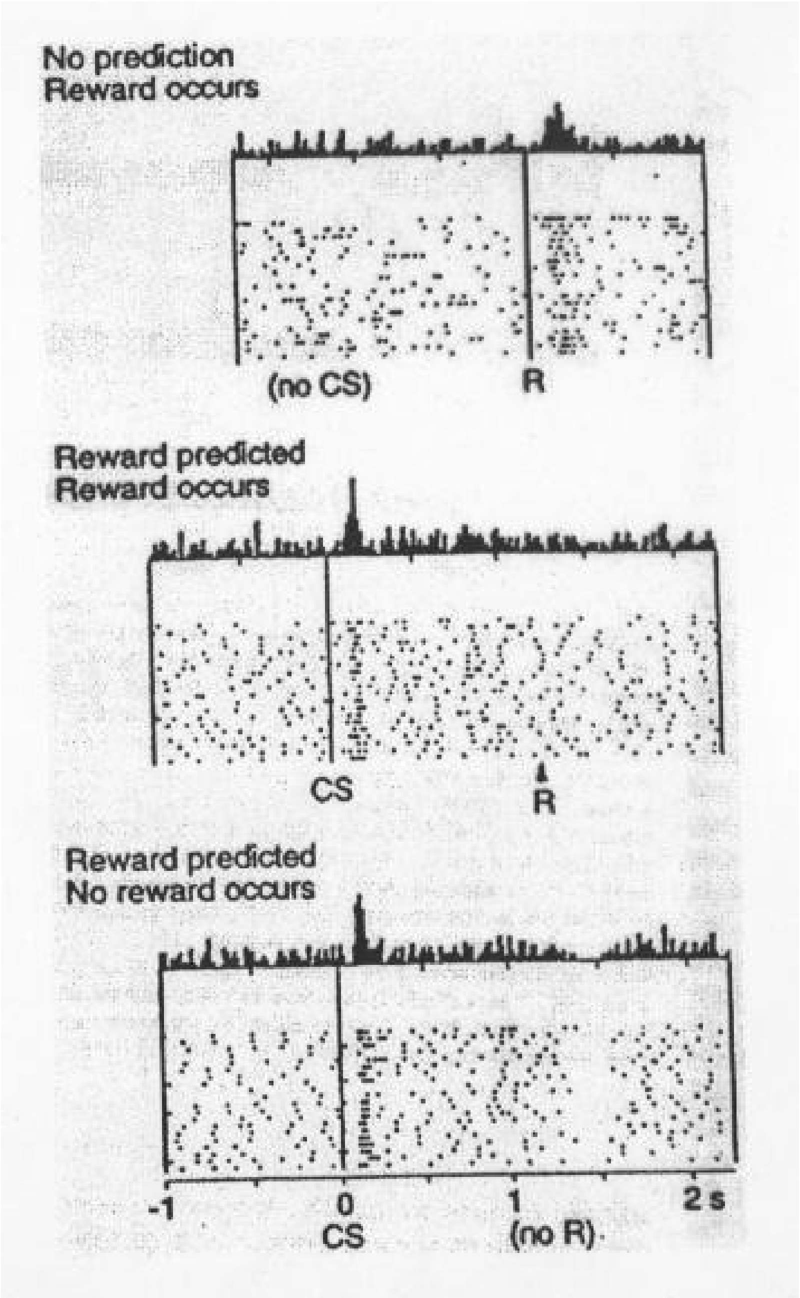


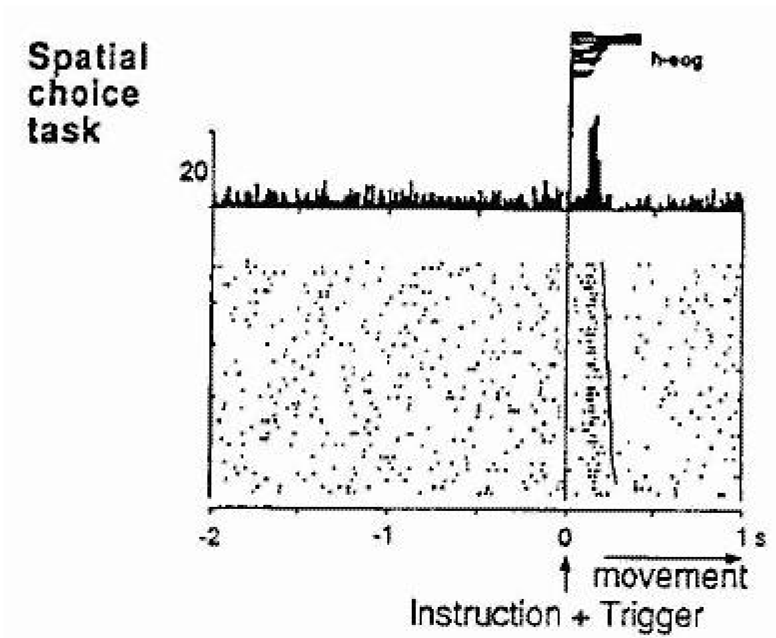
Responses of Dopamine Neurons in Macaques

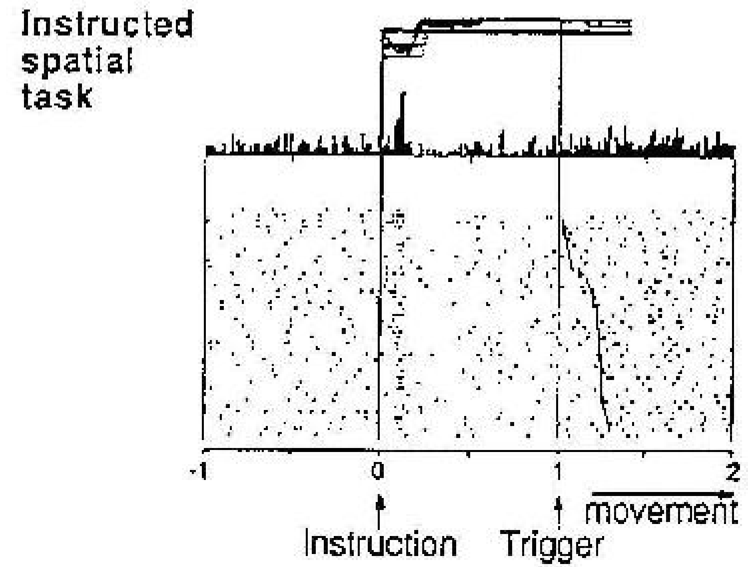
- Burst for unexpected reward

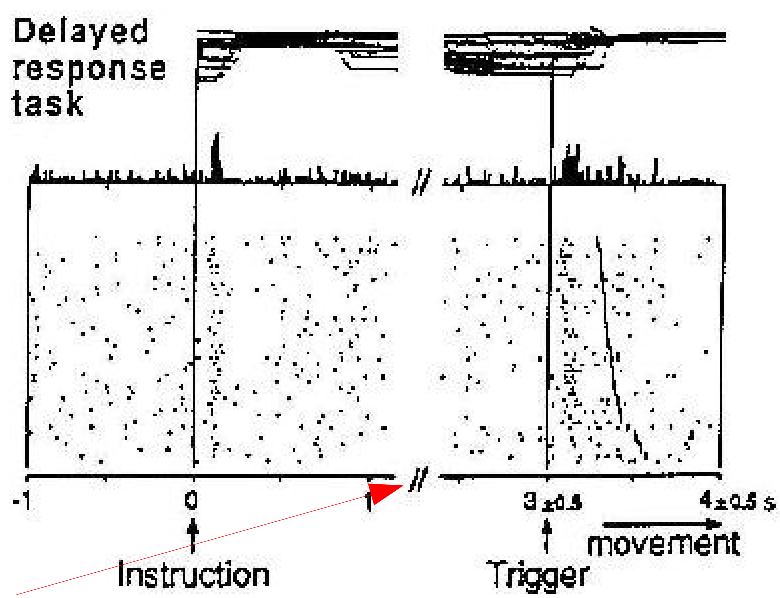
- Response transfers to reward predictors

- Pause at time of missed reward



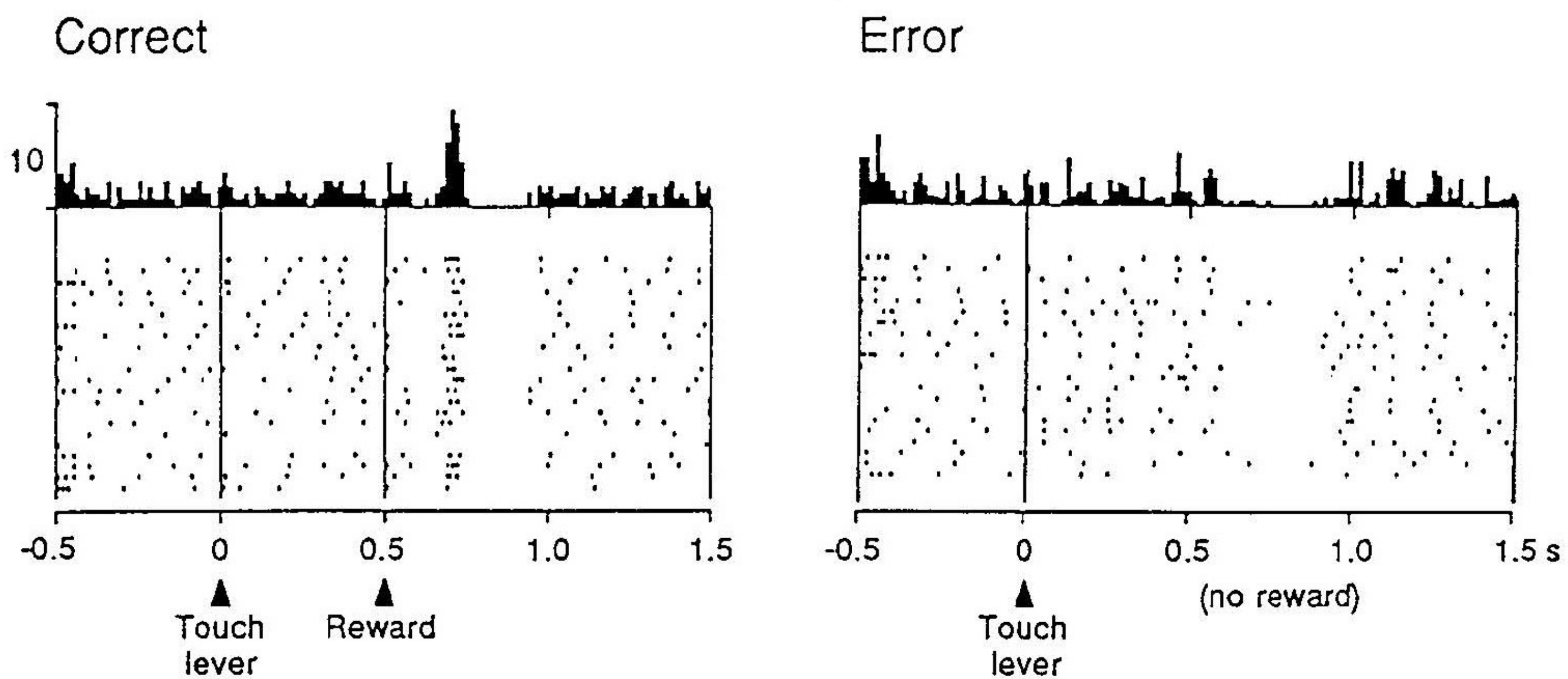




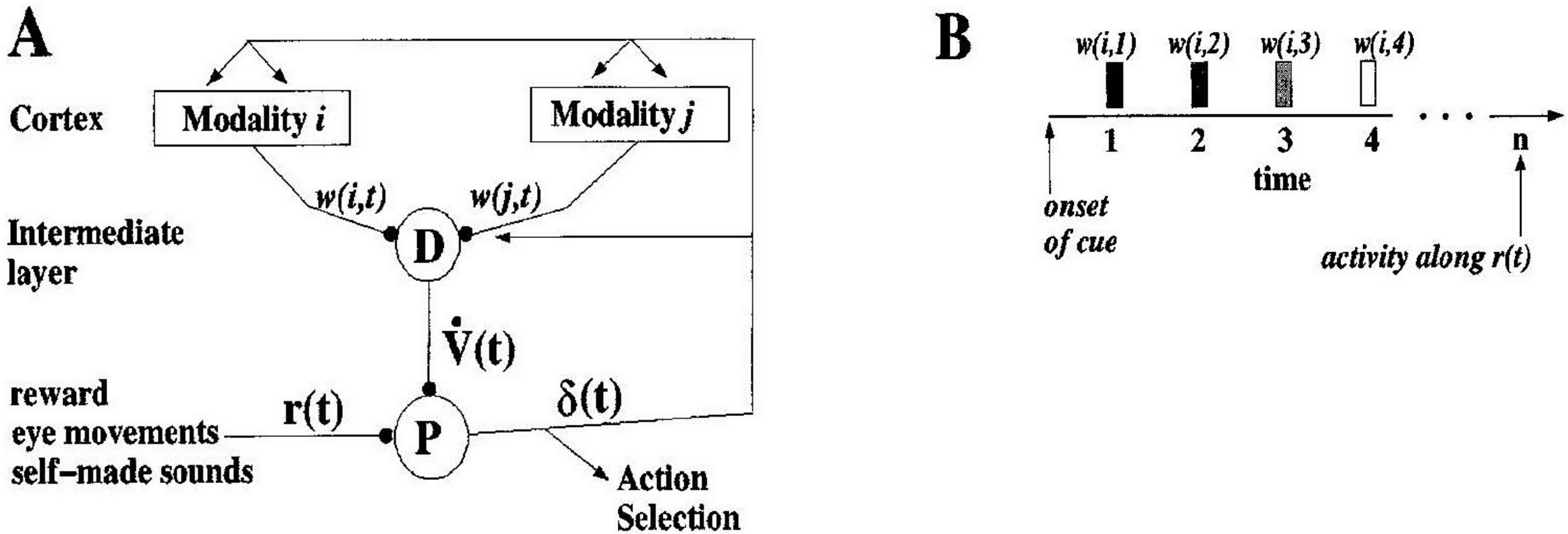


1.5 to 3.5 second delay

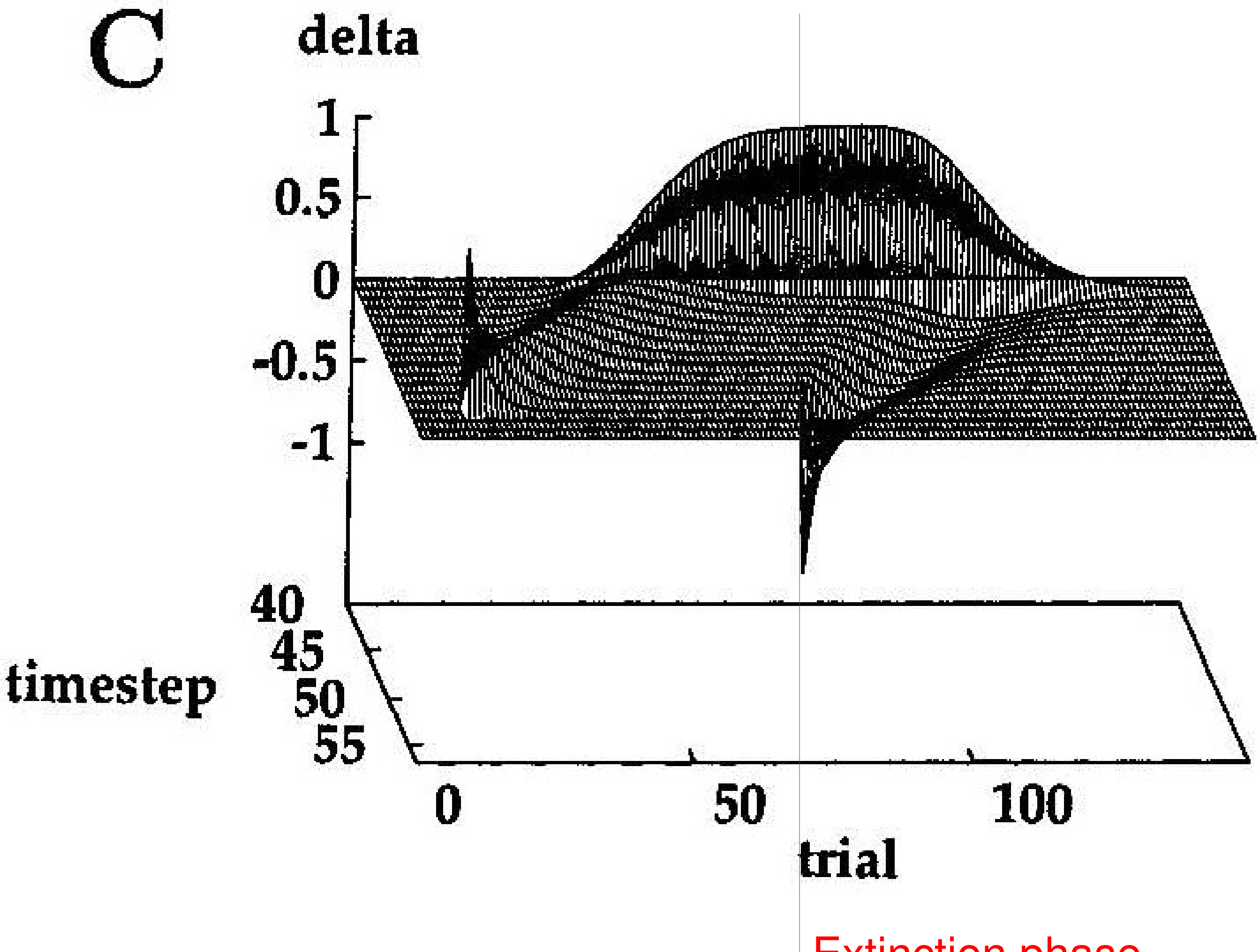
Correct and Error Trials



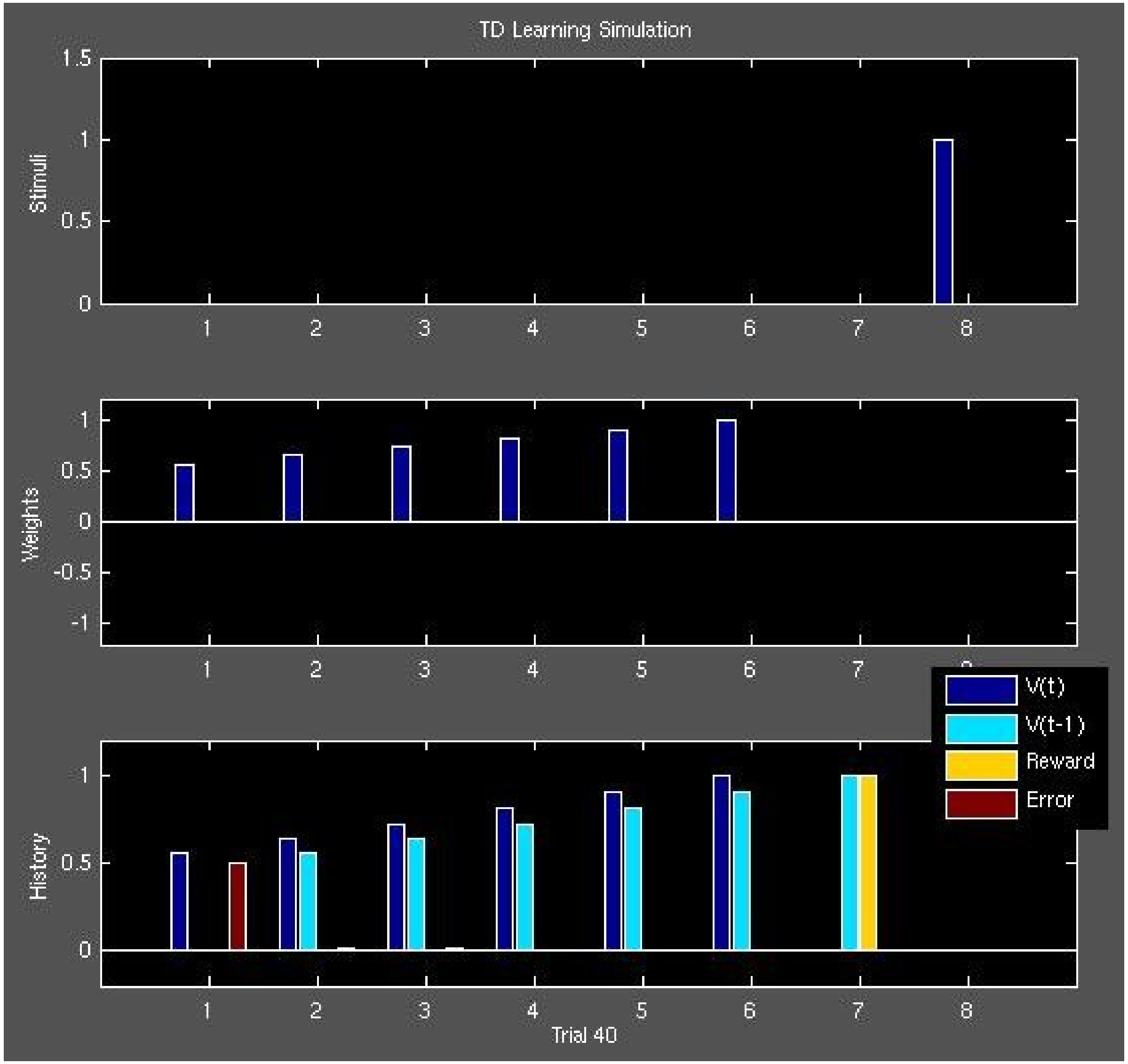
Predictive Hebbian Learning Model



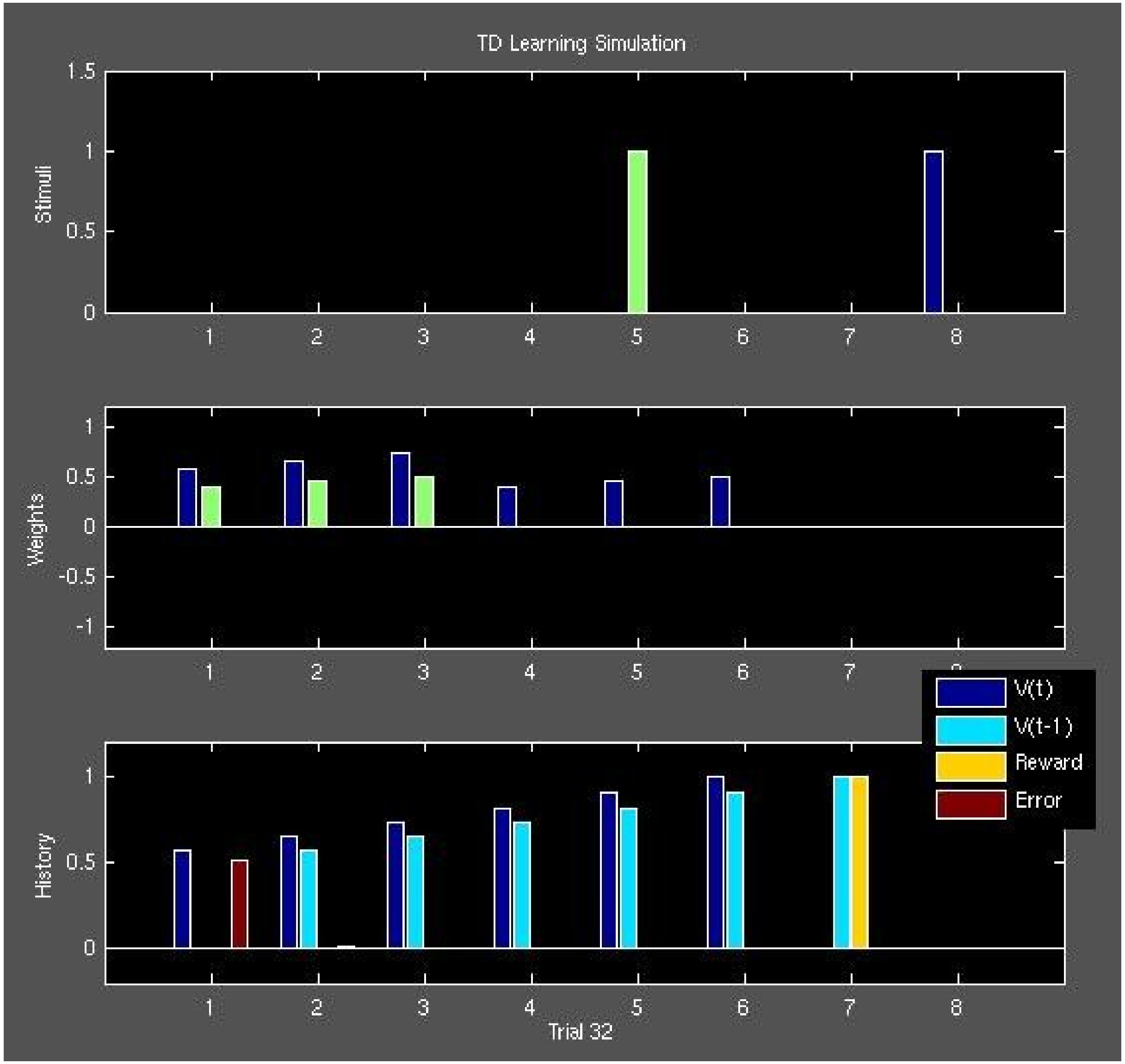
Model Behavior



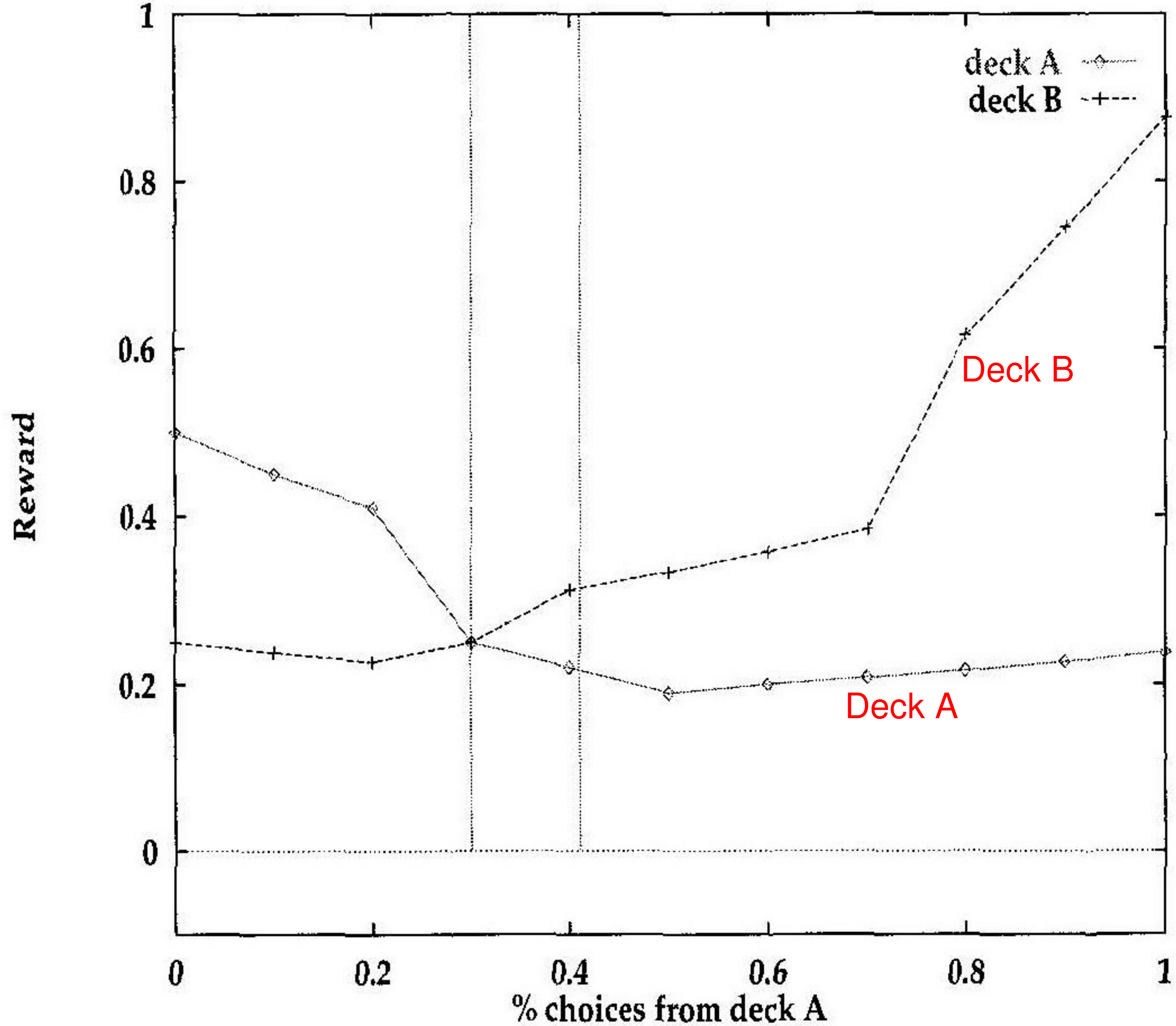
TD Simulation 1



TD Simulation 2

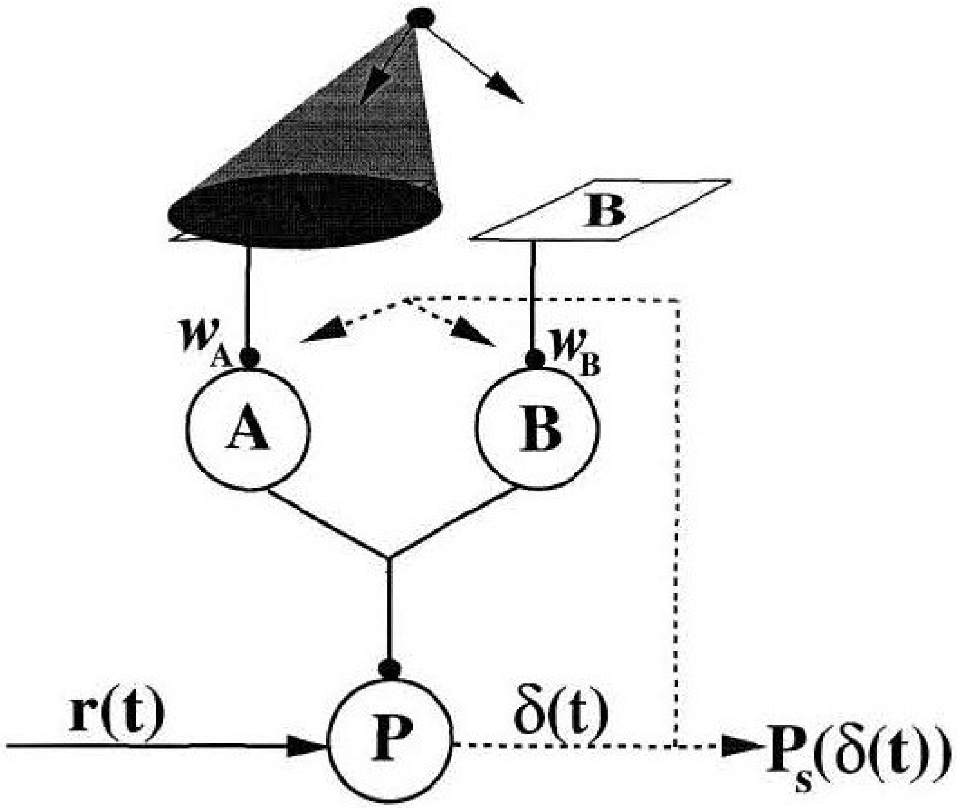


Card Choice Task



Magnitude of reward is a function of the % choices from deck A in the last 40 draws. Optimal strategy lies to the right of the crossover point, but human subjects generally get stuck around the crossover point

Card Choice Model



"Attention" alternates between decks A and B. Change in predicted reward determines P_s, the probability of selecting the current deck. The model tends to get stuck at the crossover point, as humans do.

Conclusions

- Specific neurons distribute a signal that represents information about future expected reward (VUMmx1; dopamine neurons).
- These neurons have access to the precise time at which a reward will be delivered.
  - Serial compound stimulus makes this possible.
- Fluctuations in activity levels of these neurons represent errors in predictions about future reward.
- Montague et al. (1996) present a model of how such errors could be computed in a real brain.
- The theory makes predictions about human choice behaviors in simple decision-making tasks.