Classical Conditioning V: Opposites and Opponents

PSY/NEU338: Animal learning and decision making:
Psychological, computational and neural perspectives

where were we?

- Classical conditioning = prediction learning
- Key experiment: blocking
- Rescorla-Wagner model
- Second order conditioning
- Temporal-Difference learning model
- The prediction error theory of dopamine

$$V^{new}(S) = V^{old}(S) + \eta[r(S') + V^{old}(S') - V^{old}(S)]$$

results of the 5 minute paper

Pace:

- perfect (9), challenging but doable (7), too fast (2), could move faster (3)

What worked for you?

- summary of where we are, repetition
- dopamine: firing patterns, network
- examples from real life
- matlab simulation
- writing out equations on the board rather than from slides
- working in groups to figure out new model; multiple choice questions
- precepts!

What was unclear?

- variables in TD learning rule unnecessarily complex, math notation confusing (t vs T etc)
- deriving TD model on the board top-down
- going through TD simulation with numbers in class
- implementational level, brain areas

outline for today...

- fMRI of prediction errors
- more classical conditioning: it is not all steaks and bells…
  - ➡ excitatory versus inhibitory conditioning
  - ➡ appetitive versus aversive conditioning
- opponent process model
- more fMRI of prediction errors

functional magnetic resonance imaging (fMRI)

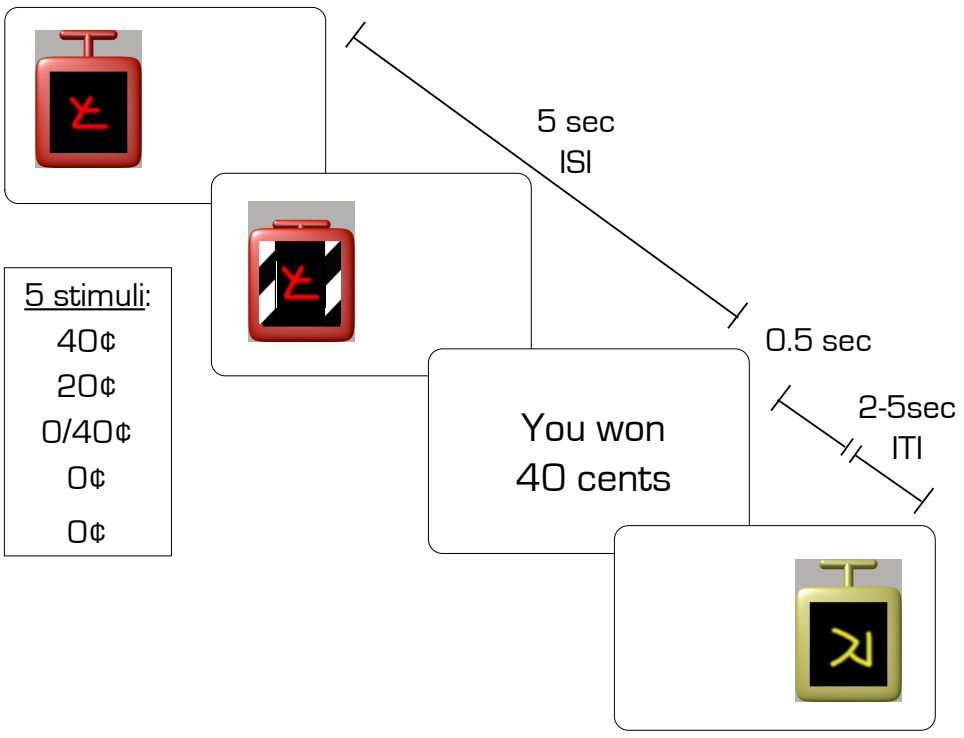
- measure BOLD ("blood oxygenation level dependent") signal
- oxygenated vs de-oxygenated hemoglobin have different magnetic properties
- detected by big superconducting magnet

Idea:

- Brain is functionally modular
- Neural activity uses energy & oxygen
- Measure brain usage, not structure

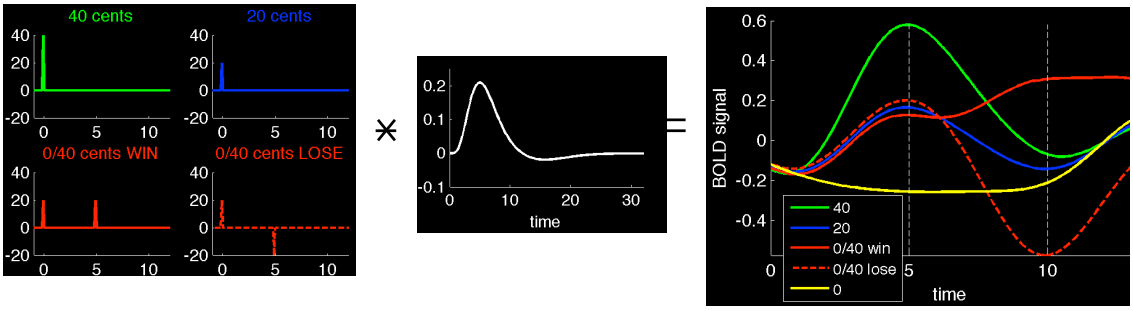
- Spatial resolution: ~3mm 3D "voxels"
- temporal resolution: 5-10 seconds

imaging prediction errors in humans

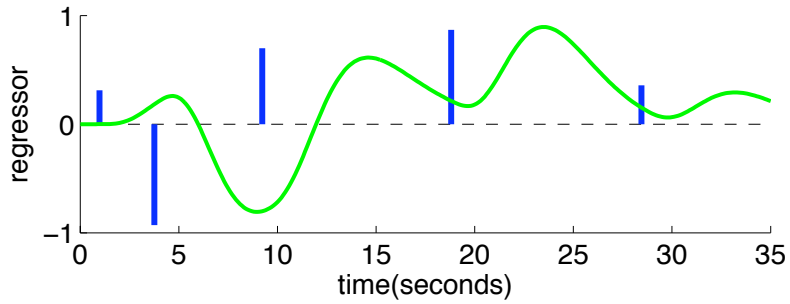


searching for prediction error signals in humans

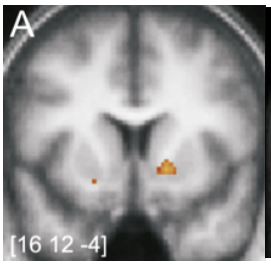
What would a prediction error look like?



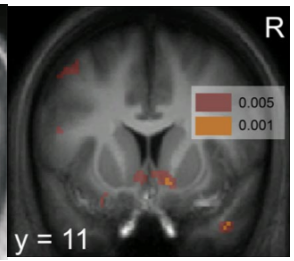
searching for prediction error signals in humans



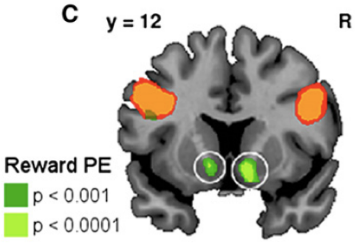
imaging prediction errors in humans



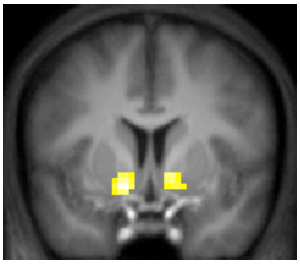
Wittmann et al., 2008



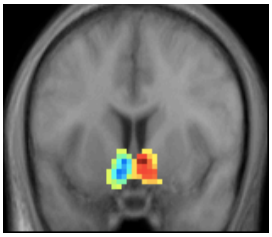
Gershman et al., 2009

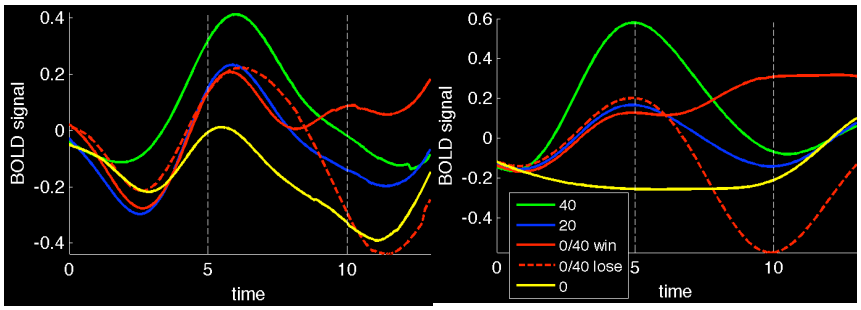


Glascher et al., 2010



(in my lab)





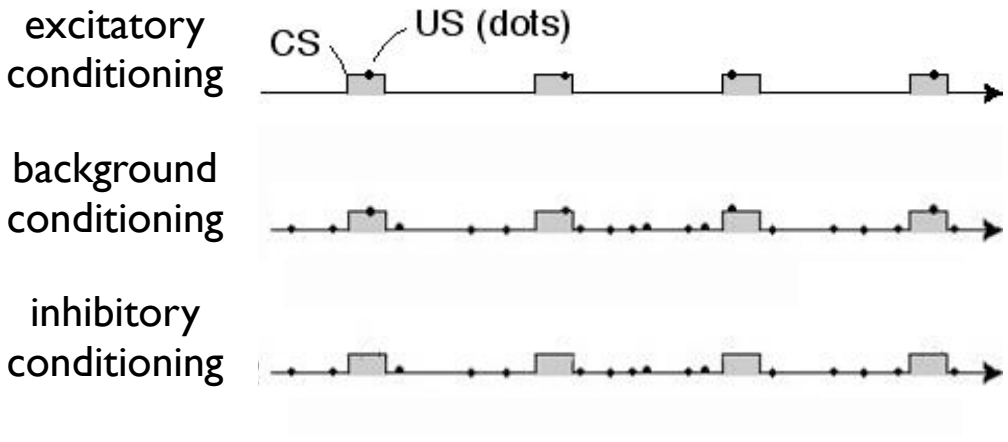
why is this useful?

All models are wrong,
some models are useful

outline for today...

- fMRI of prediction errors
- more classical conditioning: it is not all steaks and bells…
  - ➡ excitatory versus inhibitory conditioning
  - ➡ appetitive versus aversive conditioning
- opponent process model
- more fMRI of prediction errors

excitatory vs. inhibitory conditioning



how can we measure an inhibitory stimulus?

1. directly: CER, withdrawal,
   (only possible for some CRs, of course)
2. summation test
   (with another classically or instrumentally conditioned stimulus)
3. retardation test
   (must be with a US of the same motivational "class")

common requirement: pass both "tests" (2+3)

how can we create an inhibitory stimulus?

1. tone → food
   tone+light → no food
   (e.g. green light + policeman at intersection)
2. food
   light → no food
   (also: backward conditioning)
3. A+ ; B- alternating
   (differential inhibition)

which of these can RW/TD explain?

how can we create an inhibitory stimulus?

1. tone → food
   tone+light → no food
   (e.g. green light + policeman at intersection)

A. can be explained by RW and TD
B. can be explained by RW but not TD
C. can be explained by TD but not RW
D. can't be explained by the models we know so far

how can we create an inhibitory stimulus?

2. food
   light → no food
   (also: backward conditioning)

A. can be explained by RW and TD
B. can be explained by RW but not TD
C. can be explained by TD but not RW
D. can't be explained by the models we know so far

how can we create an inhibitory stimulus?

3. A+ ; B- randomly mixed
   (differential inhibition)

A. can be explained by RW and TD
B. can be explained by RW but not TD
C. can be explained by TD but not RW
D. can't be explained by the models we know so far

outline for today...

- fMRI of prediction errors
- more classical conditioning: it is not all steaks and bells…
  - ➡ excitatory versus inhibitory conditioning
  - ➡ appetitive versus aversive conditioning
- opponent process model
- more fMRI of prediction errors

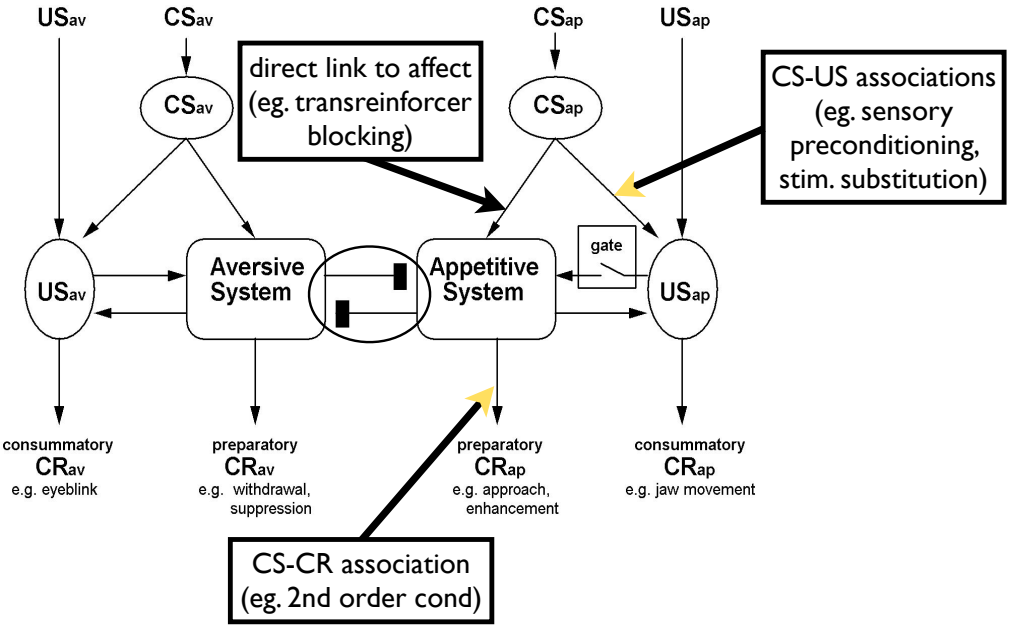
appetitive vs. aversive conditioning

- Pavlov: consummatory reflexes versus defense reflexes
- more general: appetitive USs (food, water) versus aversive (acid, shock) USs
- another (intuitive) distinction: satisfiers versus annoyers

Theory: two opponent motivational systems

Konorski: opponent (antagonistic) motivational systems

Idea: USs have sensory properties (determine type of CRs) and affective properties (determine reinforcing ability). The latter are only of two possible (antagonistic) types



Konorski (1976), Balleine+Dickinson (2002)

opponent (antagonistic) motivational systems: support

1. difficult to "activate" both systems simultaneously
   - counter conditioning is hard: CS→food, then CS→shock
   - aversive and appetitive USs "cancel" each other:
     CS(shock)→food
     UR to the shock disappears as food CR acquired and shock's aversiveness is reduced (tested through suppression)
2. CSs associated with one motivational system will increase behaviors dependent on this system, and suppress those dependent on the other (tested thru PIT)

summary so far...

	excitatory (+1)	inhibitory (-1)
appetitive (+1)	**+1 Hope**	**-1 Frustration**
aversive (-1)	**-1 Fear**	**+1 Relief**

appetitive motivation/affect system

aversive motivation/affect system

to know what CR a stimulus will cause, you must know what type of conditioning it received, and with what type of US

self-test questions

- What areas of the brain typically correlate with prediction errors in fMRI studies? explain why it makes sense that it would be these areas.
- You are designing an fMRI experiment in which you are interested in looking at prediction errors in the brain. Discuss two design features that you would use in order to maximize your power to detect prediction errors.
- A stimulus was conditioned with inhibitory conditioning. What type of CR would you expect to see?
- In Corbit & Solomon's opponent process model only the B (inhibitory) process is conditioned. How can you relate this to the phenomenon of drug overdosing?