Behavioral and neural studies of human choice

Nathaniel Daw
New York University
IPAM summer school

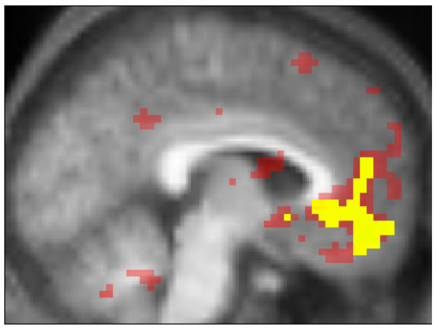
Bianca Wittman, Ben Seymour, Ray Dolan

Overview

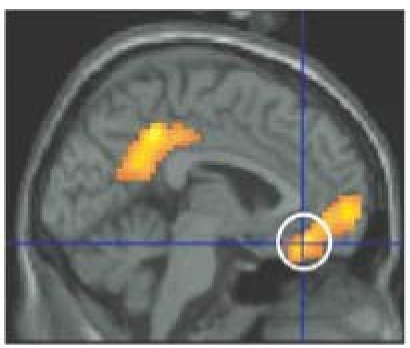
- Prediction errors and heuristics: novelty bonuses
- “Neuroeconomics”
  - Multiplayer interactions: iterated model-based reasoning
  - Methodology: Bayesian model fitting & comparison

General findings

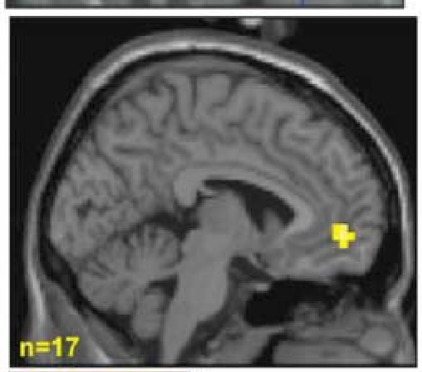
Variety of rewards or reward anticipation activates vmPFC/OFC, striatum (sometimes midbrain)



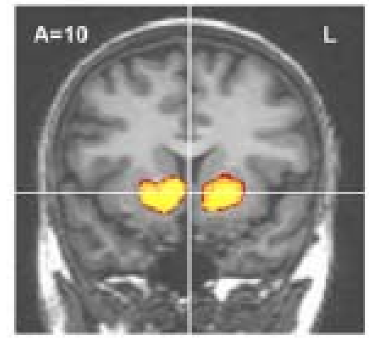
money
value predicted
(Daw et al 2006)



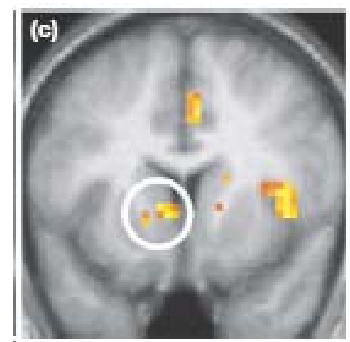
faces
attractiveness
(O'Doherty et al 2003)



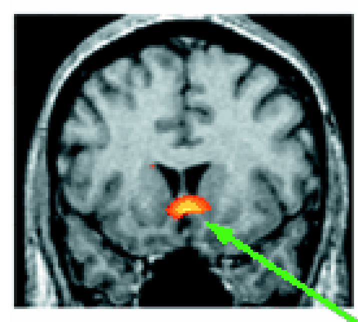
Coke or Pepsi
degree favored
(McClure et al. 2004)



money
gain vs loss
(Kuhnen & Knutson 2005)



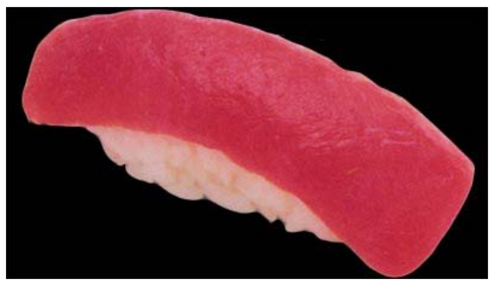
food odors
valued vs devalued
(Gottfreid et al 2003)



juice
unpredictable vs predictable
(Berns et al 2001)

Explore/exploit dilemma

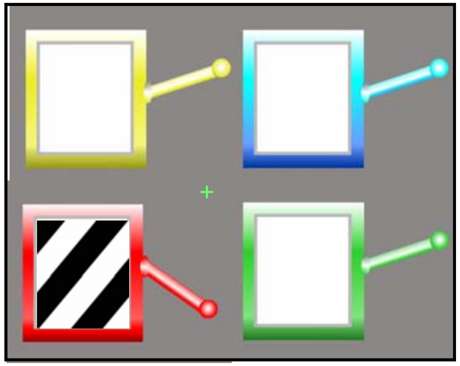
- In choice under uncertainty: how to balance exploiting known good options vs. sampling unfamiliar ones (short vs long term win; latter hard to quantify)

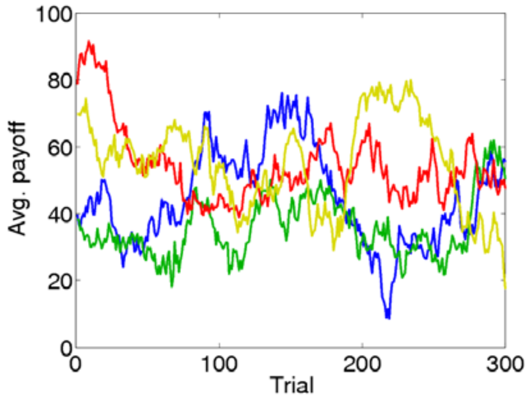


- Optimal solution prizes resolving uncertainty (how much so tractable in special case; Gittins 1972)
- In practice:
  - random exploration (e-greedy; softmax)
  - heuristic bonuses (uncertainty; novelty; inspired by Gittins)
  - these are behavioral/neural hypotheses

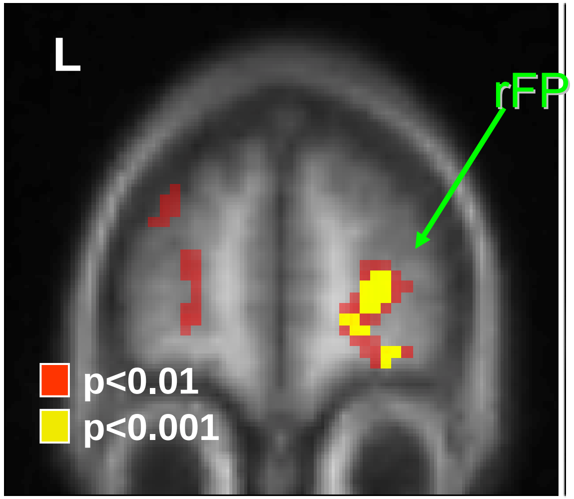
Explore/exploit in humans

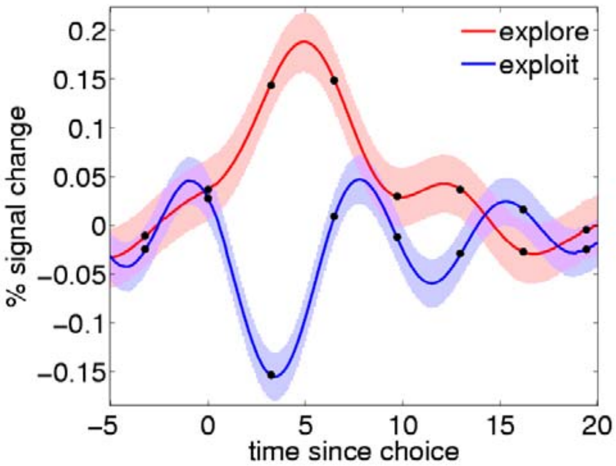
- 4-armed bandit; induce uncertainty with payoff diffusion





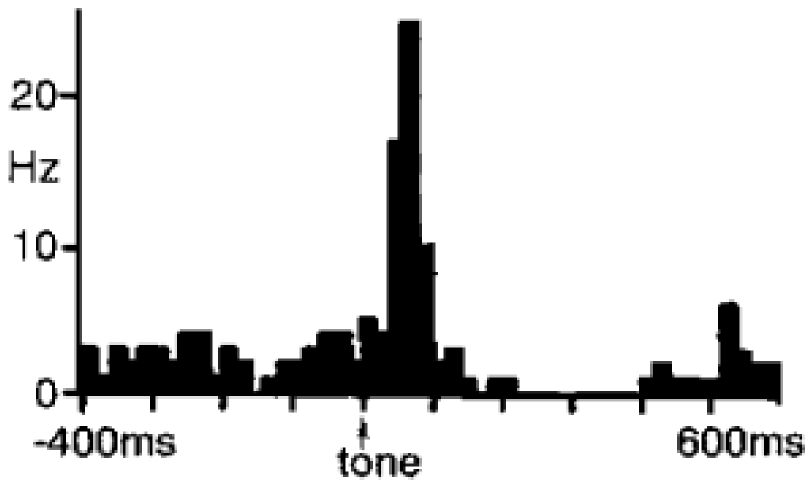
- Behavior: exploration not directed toward uncertainty (softmax)
- Neural: exploration activates control rather than reward areas



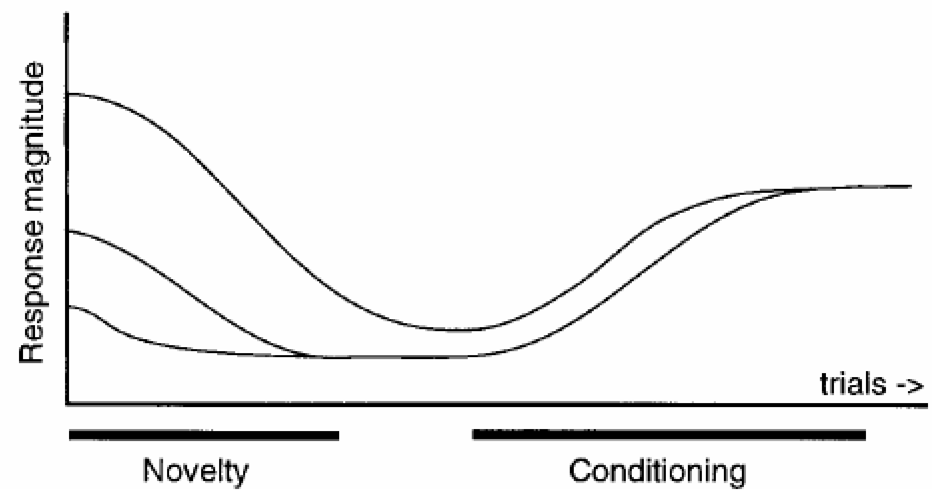


(Daw, O'Doherty et al 2006)

Dopamine novelty responses



Horvitz et al 1997

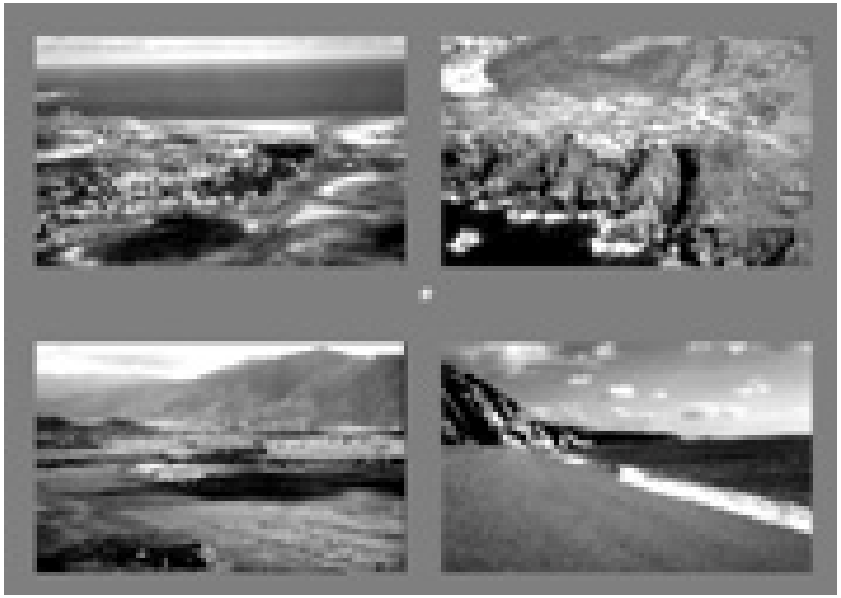


Schultz 1998

- Burst / pause response to salient novel stimuli
- Suggests novelty bonus (Kakade & Dayan 2002) but neural signal never linked to behavior
- "Optimistic initialization" (Ng & Russell 1999)

→ Manipulate novelty directly

Novelty task



- Each picture has unknown probability of reward
- New picture (with new probability) swapped in periodically

→ Half of pictures pre-exposed in a separate task
No difference in value or uncertainty in decision task

- Does perceptual novelty impact choice?

Behavioral model

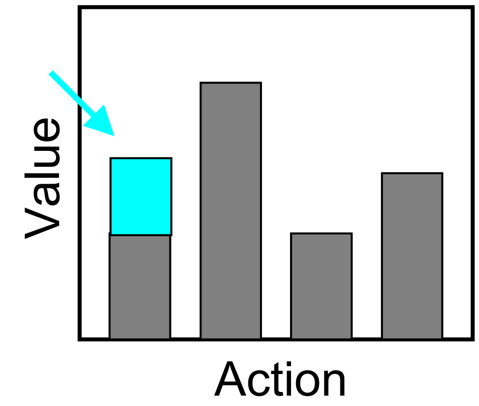
1. Estimate payoffs (TD)

$$V_{t+1}(c_t) = V_t(c_t) + \alpha(r_t - V_t(c_t))$$

2. Derive choice probabilities (softmax)

$$P_t(c) \propto \exp(\beta \cdot V_t(c))$$

Choose randomly according to these



Fit initial value of new image separately for novel vs preexposed

(Fit max likelihood parameters over data set)

Results

14 subs, random effects (summary statistics)

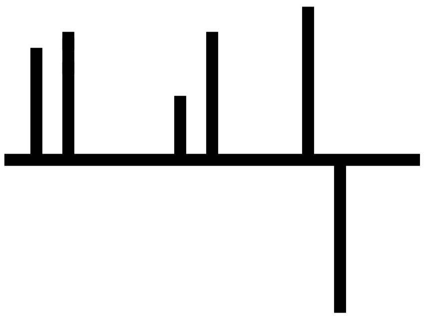
initial value, preexposed	= £ .37
initial value, novel	= £ .41
net novelty bonus	= £ .04 ± .01 (p=.01)

→ Behavioral evidence for novelty heuristic

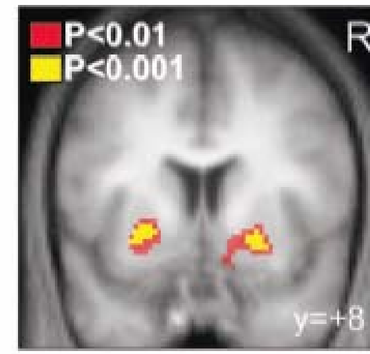
Imaging

Run model to generate per-subject, per trial TD error signals

Use (convolved with hemodynamic filter) as regressor for BOLD

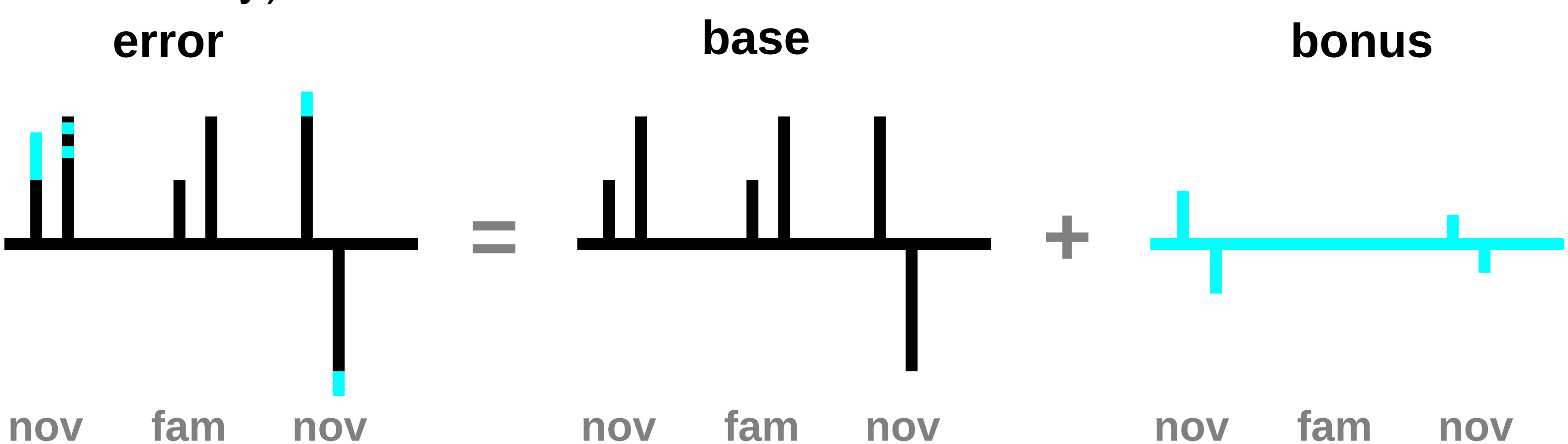


recall: striatal BOLD signal correlates with TD error



(O'Doherty et al 2004)

Decompose signal into base error plus error due to novelty bonus (exploit GLM additivity)

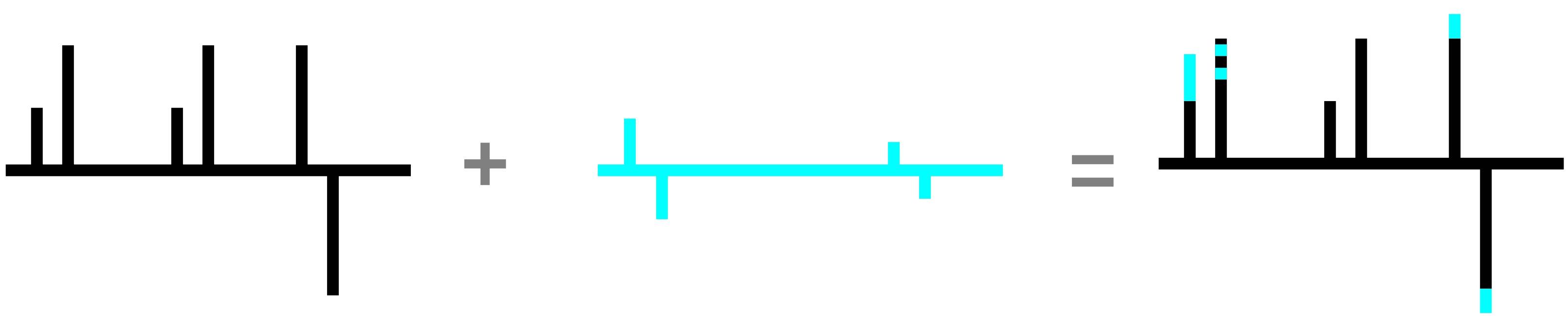


Results

Do striatal prediction errors incorporate bonus?

Base error **Bonus error** **Conj**

hi – I had to remove data from my friends and collaborators that hasn't been published yet. please contact me personally if you would like to see it.



(all survive small volume correction for striatum)

w/ Wittman, Seymour, Dolan, submitted

Timecourse

Do timecourses reflect characteristic pos-neg pattern of bonus?

Subtract timecourses: Choose novel – Choose preexposed

(illustrative, not conclusive)

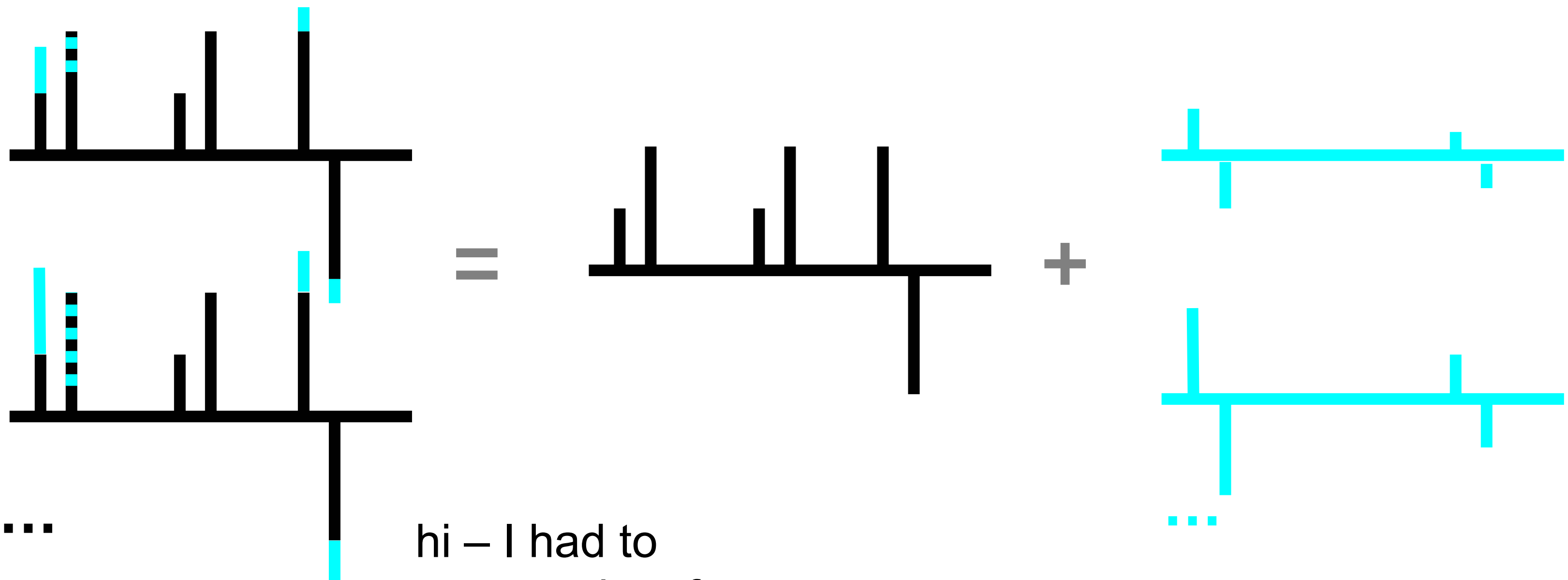
hi – I had to remove data from my friends and collaborators that hasn't been published yet. please contact me personally if you would like to see it.

time (secs) since reward
(uncorrected for hemodynamic lag)

w/ Wittman, Seymour, Dolan, submitted

2nd level results

Does bonus response reflect population variation in novelty seeking?



bonus response (& not base TD error) correlates with per-subject novelty bonus fit from choices

hi – I had to remove data from my friends and collaborators that hasn't been published yet. please contact me personally if you would like to see it.

also with novelty-seeking personality trait (TPQ)

w/ Wittman, Seymour, Dolan, submitted

Summary

Perceptual novelty manipulated holding true value constant

Behavior: novelty bonus heuristic

- Novelty as proxy for uncertainty
- Probably a good approximation in natural circumstances, but exploitable
- Characteristic suboptimality: quantifiable irrationality

Imaging: pattern of novelty effects on striatal error signals also characteristic of this scheme

- Neural signal tracks behavior over trials & subjects
- Support account of dopaminergic responses

Do not see frontal signals as in previous task:

- Less need for control?

Neuroeconomics

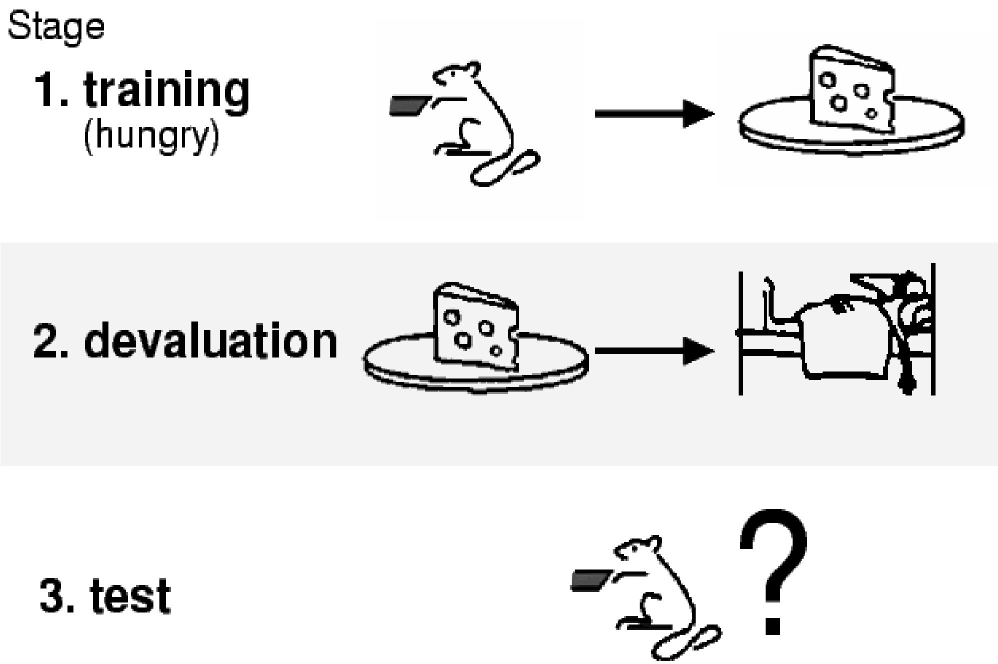
What can neuroscience/psychology learn from economists?

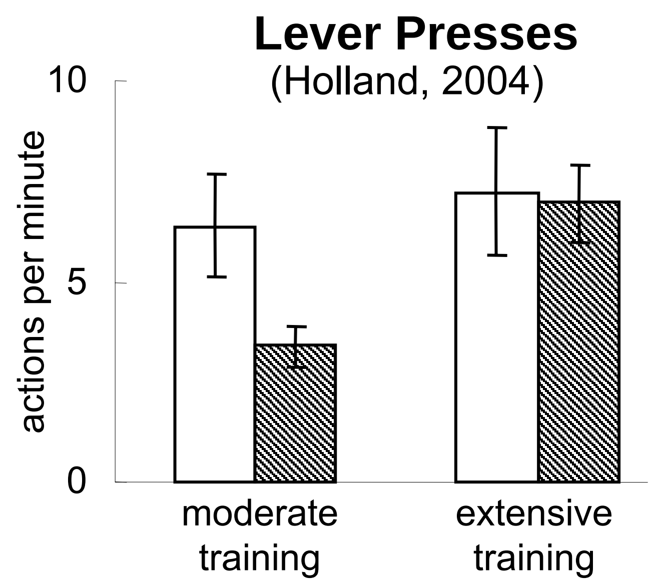
- Axiomatic utility theory and behavioral anomalies (Maloney, Fox)
  - Interest in multi-system models (Loewenstein; Sanfey; Kahneman) to explain these
- Multiplayer interactions, strategic reasoning (Zhang)
- Model estimation and comparison

p-beauty contest

- Write down an integer between 0 and 100, inclusive
- I will average all entries. The contestant who picks closest to 2/3 of the average wins the prize
- Prize split in case of tie

Behavioural results





Why is this called a p-beauty contest?

- Keynes (1936):

> It is not a case of choosing those [faces] which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees.

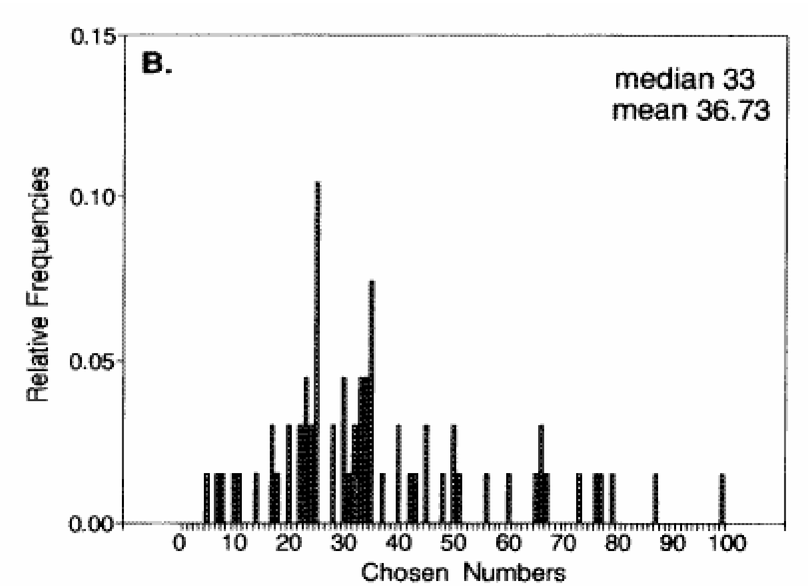
- Economists are fond of old quotes.

Model-based reasoning in multiplayer games

- eg rock-paper-scissors
  - could learn by S-R
  - or could model opponent & best-respond
  - or could take into account that opponent is probably modeling me best-responding … etc
  - this is much like multiple steps of dynamic programming
  - versions of this are involved in computing game theoretic equilibria, and in thinking about learning in games

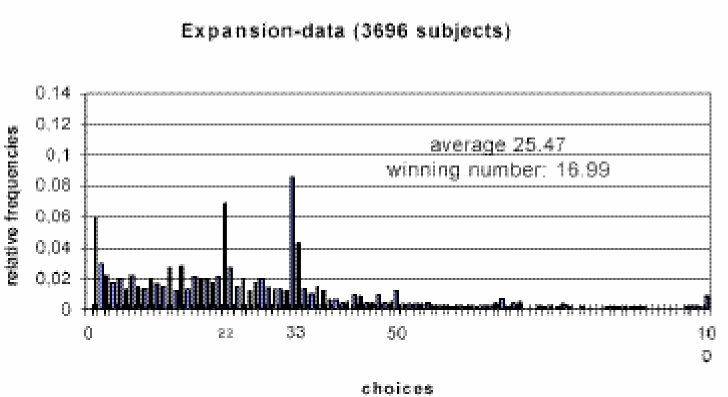
- p-bc etc try to measure this iteration

Results

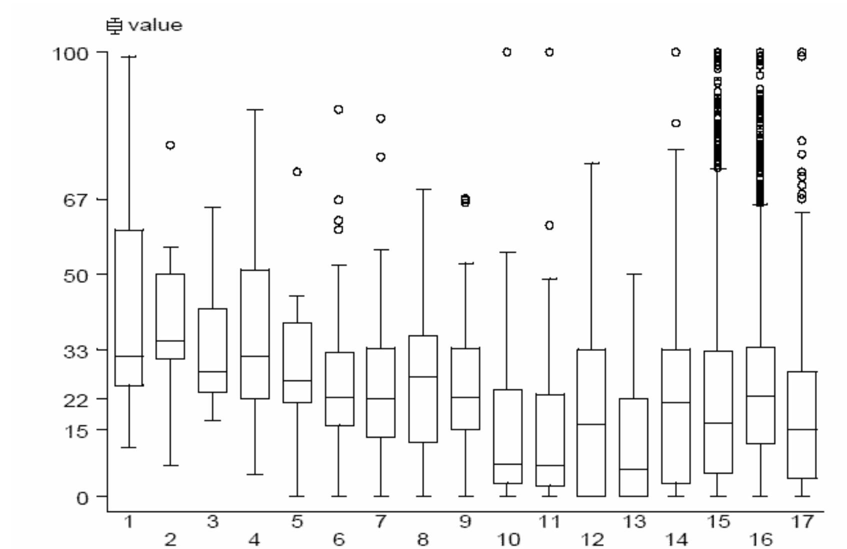


German undergrads - Nagel 1995

- Mean around 25-40; win around 16-27
  - Lower for more educated, more motivated, more time
- Suggests 0-3 rounds of iterated reasoning
  - Not predicted by TD of course
  - Realistic profit-maximizing response to dumb opponents?
  - Cognitive hierarchy model (Jun Zhang)



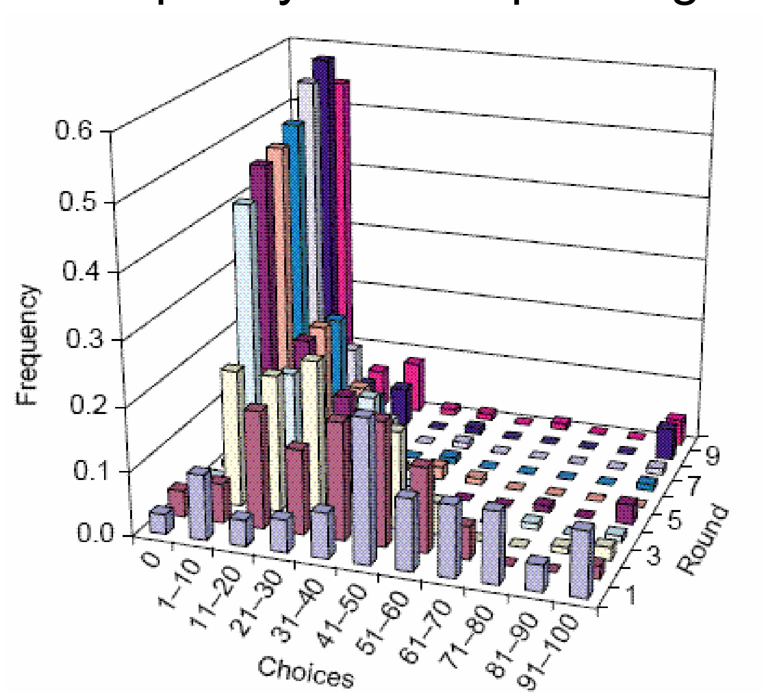
Spanish newspaper - Nagel et al. 1999



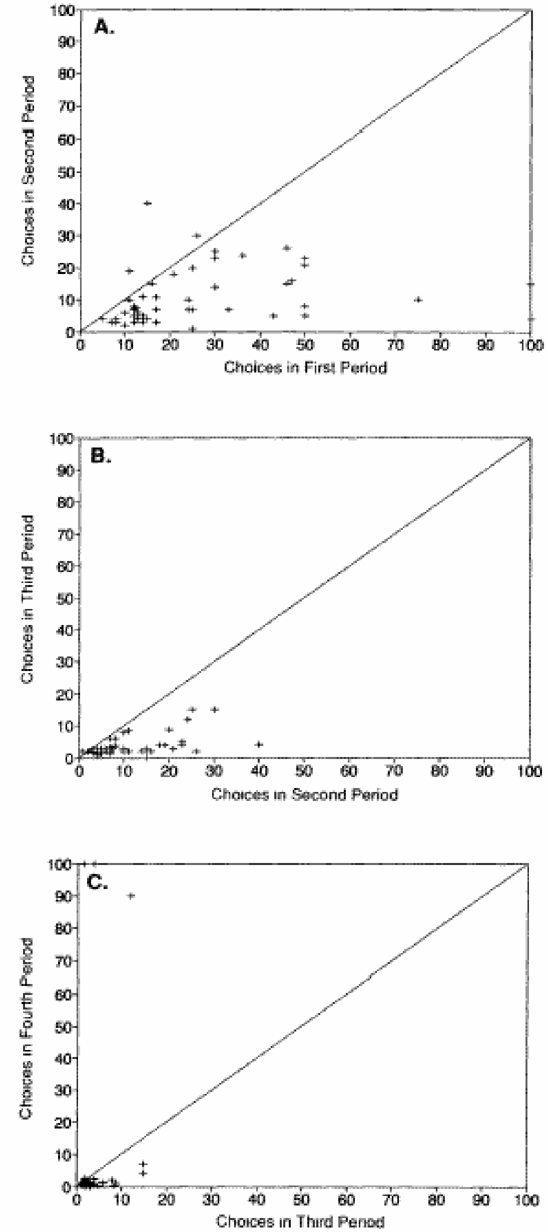
Many studies - Nagel et al. 1999

Learning

- With repeated runs, subjects quickly approach equilibrium
- This occurs before most could have won anything: again, can’t be simple TD
  - Counterfactual learning
  - Anticipatory best-responding?



Singaporean undergrads – Ho et al. 1998



German undergrads – Nagel 1995

Another example

Sequential bargaining game (Johnson et al 2001; see also Hedden & Zhang 2002)

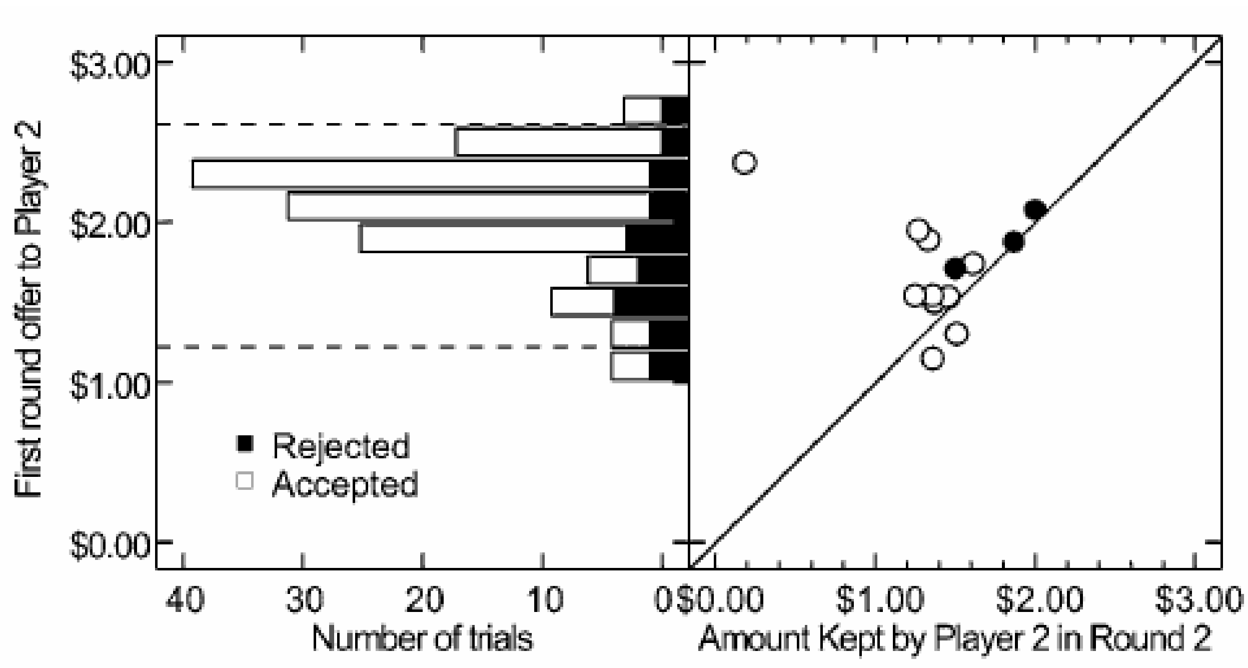
Player 1 offers player 2 part of $5.00

If refused, player 2 offers player 1 part of $2.50

If refused, player 1 offers player 2 part of $1.25

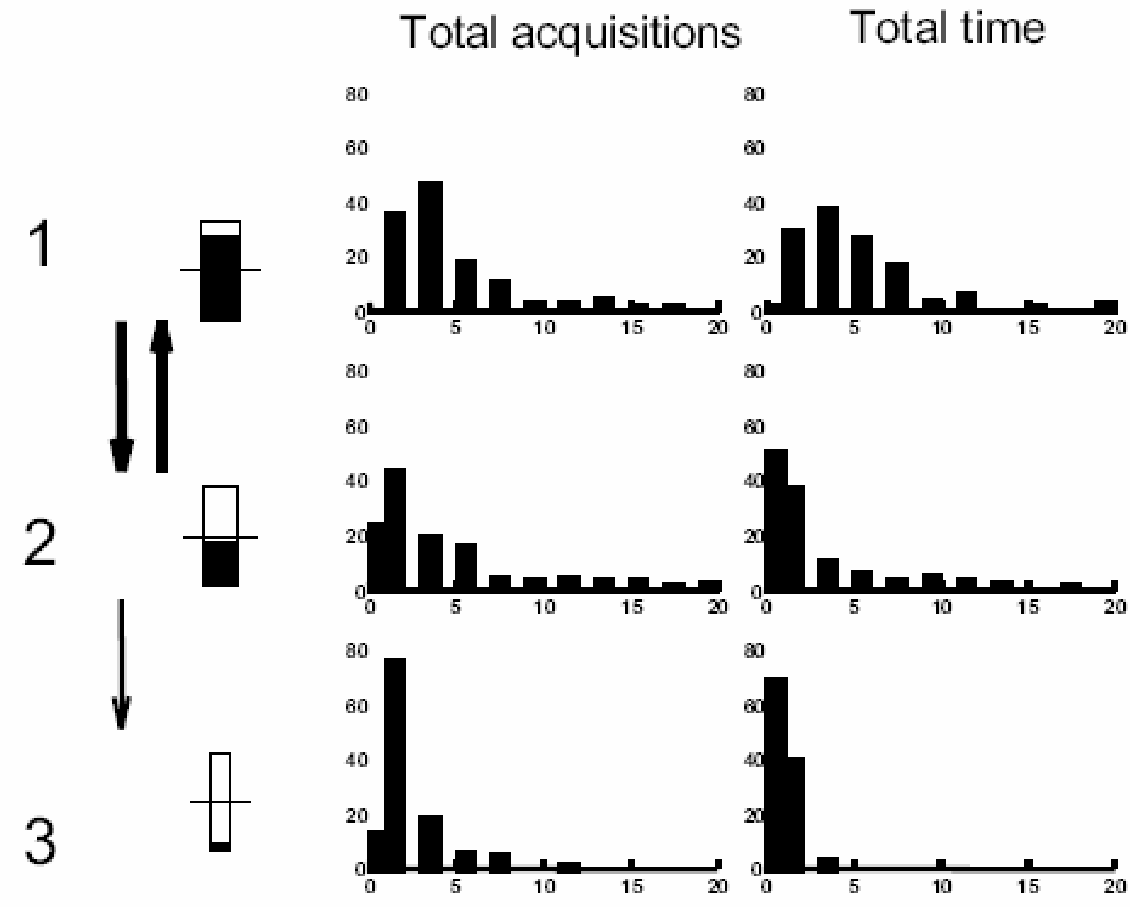
If refused, no one gets anything

(What are the equilibrium offers for self-interested, profit maximizing opponents?)



(Johnson et al 2001)

- Average first-round offer $2.11
- Offers rejected 10.8% of time
- Most counter-offers are not advantageous!



(Johnson et al 2001)

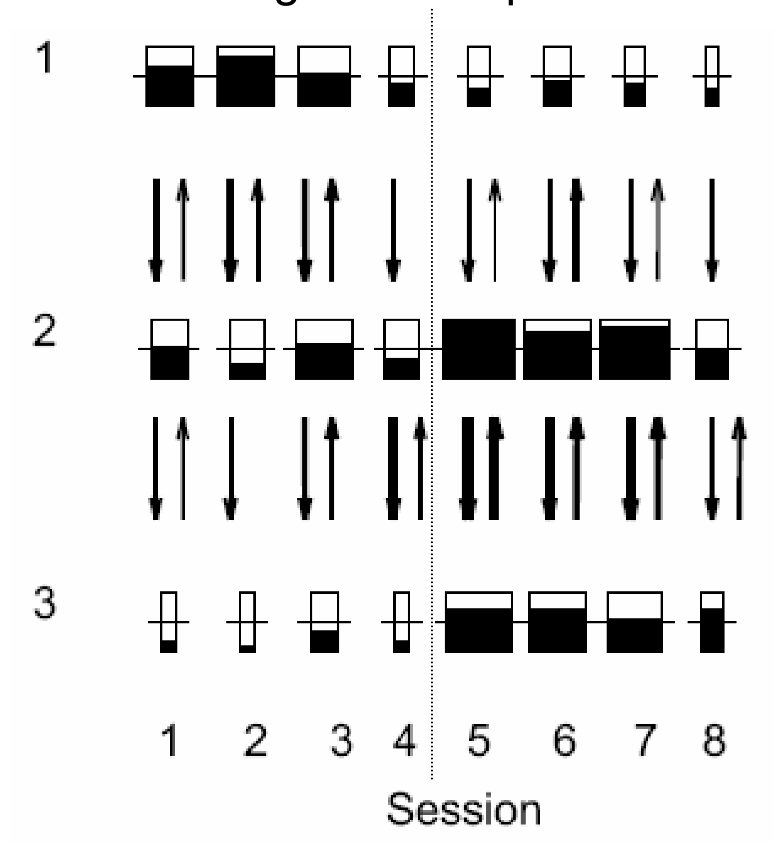
- Subjects mainly look at first-round payoffs
- Often do not even consult third round

Manipulations

Two hypotheses for cause of bad play:

- Social preferences (fairness)
  - → but offers still high when told they are playing a computer
- Cognitive demands
  - → but offers become perfect after a 5 min. lesson in game theory

Against computer

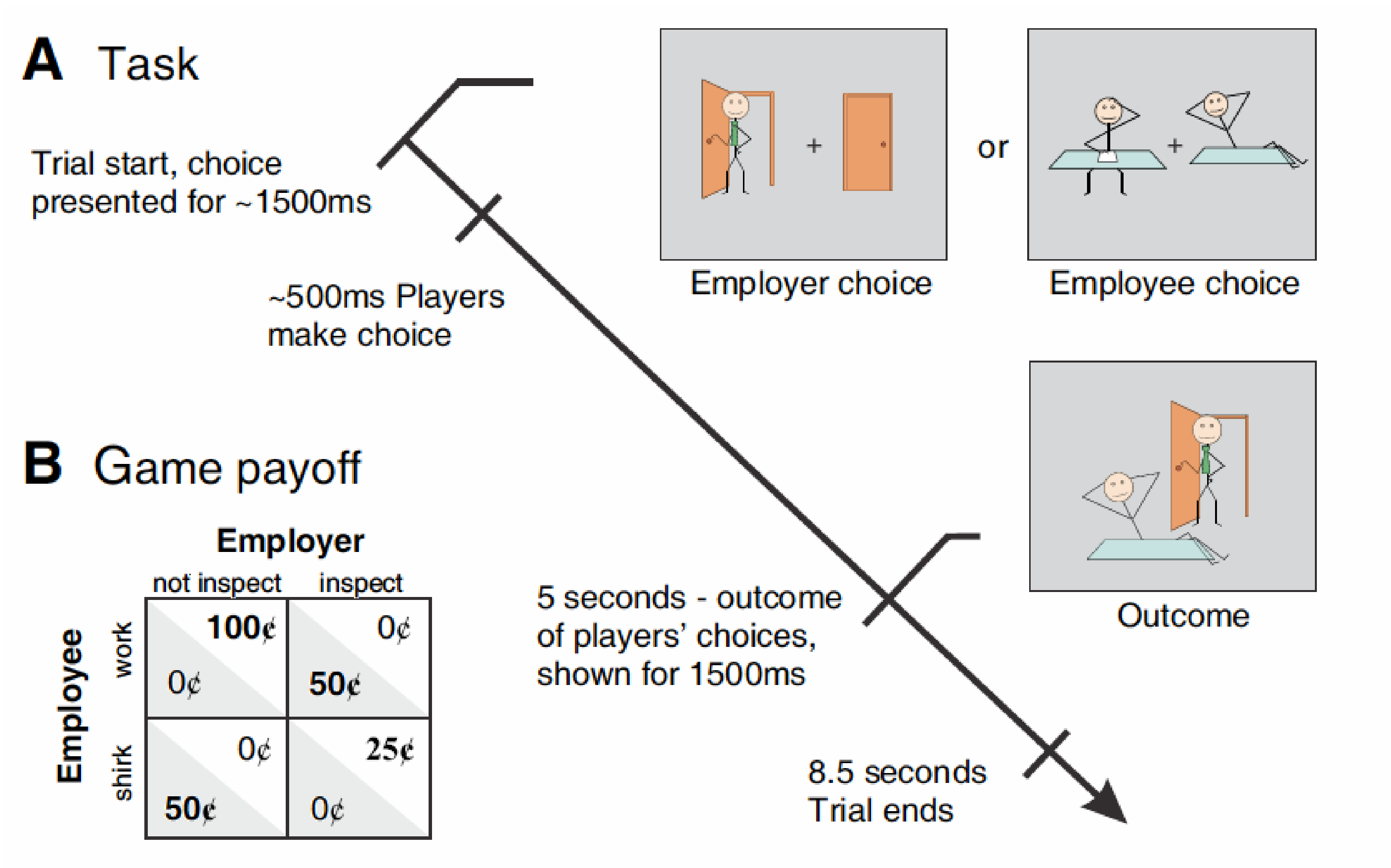


Pre training | Post training
Avg offer: $1.84 | Avg offer: $1.22

(Johnson et al 2001)

Neural example

"Work or shirk" game (Hampton et al, submitted)



Models

hi – I had to remove data from my friends and collaborators that hasn't been published yet. please contact me personally if you would like to see it.

$\delta_t = r_t - V_t$
predictable in 2 player game

$\delta^p = P - p$
model opponent's action
(compute value in expectation)

λ^p complicated function
model effects of my choice on opponent's estimate of me, assuming he is using preceding model

behavior best fit by influence learning model

(Hampton et al, submitted)

Neural activity

vmPFC value signals also best fit those predicted by influence model

hi – I had to remove data from my friends and collaborators that hasn't been published yet. please contact me personally if you would like to see it.

(Hampton et al, submitted)

Summary

- Model-based planning vs model-free RL also arises in multiplayer games; here, basic RL leaves one open to exploitation and strategic considerations demand modeling opponent's preferences
- Computing game theoretic equilibrium closely linked to sequential planning
- Behavioral evidence for various manifestations of this
- Neural evidence also, through visiualizing subjective value representations

Neuroeconomics

What can neuroscience/psychology learn from economists?

- Axiomatic utility theory and behavioral anomalies (Maloney, Fox)
  - Interest in multi-system models (Loewenstein; Sanfey; Kahneman) to explain these
- Multiplayer interactions, strategic reasoning (Zhang)
- Model estimation and comparison

Model estimation

What is a model?

- parameterized stochastic data-generation process

Model *m* predicts data *D* given parameters θ

posterior distribution over θ by Bayes' rule

$$P(\theta \mid D, m) \propto P(D \mid \theta, m) P(\theta \mid m)$$

Typically use a maximum likelihood point estimate instead

$$\arg\max_{\theta} P(D \mid \theta, m)$$

ie the parameters for which data are most likely.

Can still study uncertainty around peak: interactions, identifiability

This is good for what?

- parameters may measure something of interest
  - loss aversion, novelty sensitivity
- allow to quantify & study subjective representations
  - expected value, prediction error
- compare groups
- compare models

Example: RL

eg D is ordered list of choices c_t, rewards r_t

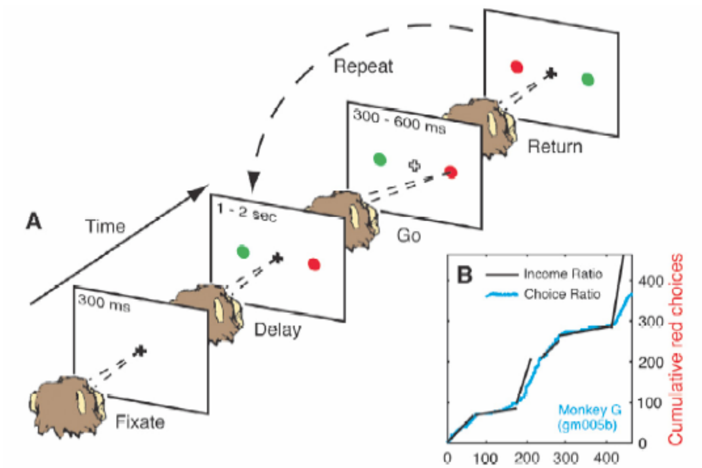
$$P(D \mid \theta, m) = \prod_t P(c_t \mid c_{1..t-1}, r_{1..t-1}, \theta, m)$$

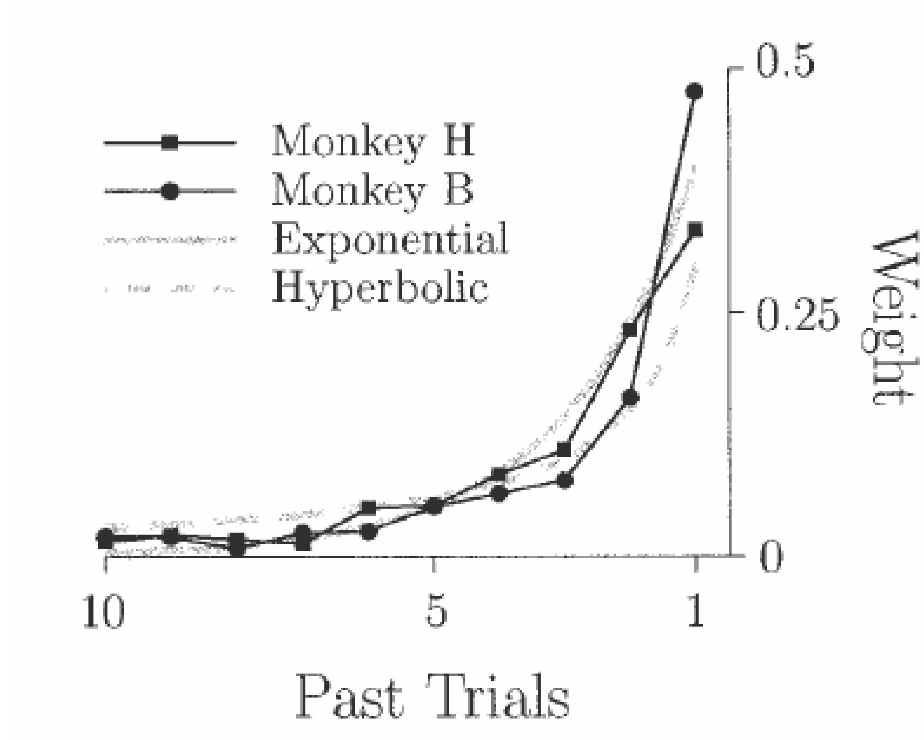
for eg

$$P(c_t \mid c_{1..t-1}, r_{1..t-1}, \theta, m) \propto \exp(V_t(c_t))$$

where V might be learned by a TD model or is a weighted sum of regressors

Animal example



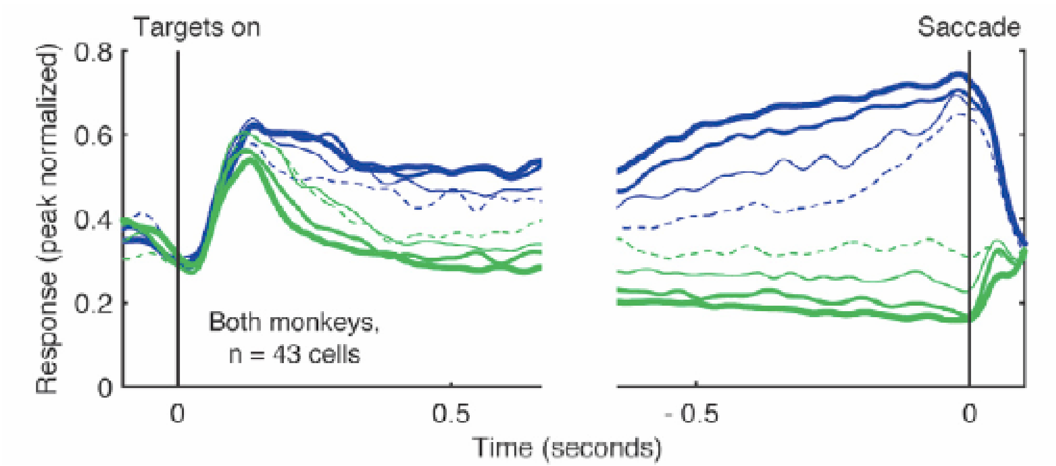


Monkey repeatedly choosing between two alternatives; sigmoidal regression: rewards→choices (Lau et al. 2005; cf Corrado et al. 2005)

Suggests exponential filter (Kalman filter, etc)

Example, continued

Sugrue et al. (2004): primate LIP neurons:



Random effects

- How to cope with multiple subjects?
  - hierarchical ('random effects') model

$$P(D \mid \theta, m) = \prod_s \int d\theta_s P(\theta_s \mid \theta) \prod_t P(c_{s,t} \mid c_{s,1..t-1}, r_{s,1..t-1}, \theta_s, m)$$

  - subject drawn randomly from population distribution, data drawn randomly from subject's distribution
  - θ_s might be, e.g., a regression weight or a learning rate
  - $P(\theta_s \mid \theta)$ might be a Gaussian, or a mixture
  - θ might be the mean and variance, over the population, of the regression weights

  - essentially all fMRI analyses work this way

Example

weight

hi – I had to remove data from my friends and collaborators that hasn't been published yet. please contact me personally if you would like to see it.

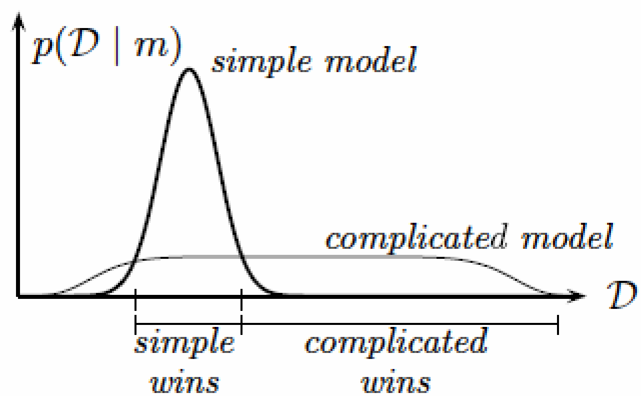
money shocks choices

Model comparison

$$P(m \mid D) \propto P(D \mid m)P(m)$$

$$P(D \mid m) = \int d\theta_m P(D \mid \theta_m, m)P(\theta_m \mid m)$$

In principle: 'automatic Occam's razor'



In practice: approximate integral as max likelihood + penalty: Laplace, BIC etc. Frequentist version: likelihood ratio test

Or: holdout set; difficult in sequential case

Good example ref: Ho & Camerer

Summary

- Bayesian inference for data analysis
- Models quantify subjective factors; allow studying their neural representations
  - link multiple types of data (behavior, physiology)
- Models as hypotheses, can be quantitatively compared
  - average fit over raw data, not fit to average data

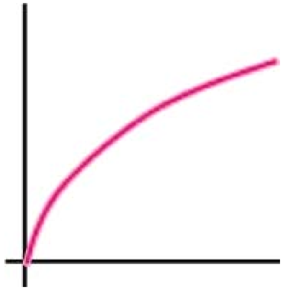
Levels of analysis

Marr's (1982) famous hierarchy:

Computation

interpretation: why?

Bayesian decision theory



Algorithm

eg TD sampling: heuristics, shortcuts

$$\delta_t = r_t + \hat{V}_{s(t+1)} - \hat{V}_{s(t)}$$

Implementation

simulation: how?

eg dopamine