Reference Dependence In the Brain

Mark Dean

Behavioral Economics G6943
Fall 2015

Introduction

- So far we have presented some evidence that sensory coding may be reference dependent
- In this lecture we will present evidence that important value systems in the brain are also reference dependent
- Will also show how neuroscientific measurement and economic theory can be merged to answer interesting questions

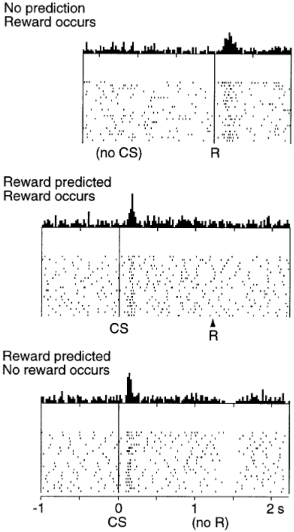
The Reward Prediction Error Hypothesis

- *Dopamine* is a neurotransmitter
  - Transmits information between brain cells
- Until relatively recently, assumed to be the brain's 'happy place'
- Turns out to be more complicated that that
- Now hypothesized to encode *Reward Prediction Error (RPE)*
  - Difference between *experienced* and *predicted* rewards
- If true, lots of interesting implications
  - RPE used in AI learning
  - Explicit reference dependence
  - Makes reference point observable

Early Evidence for RPE - Monkeys

Schultz et al. [1997]

- Dopamine fires only on receipt of *unpredicted* rewards
- Otherwise will fire at *first predictor* of reward
- If an expected reward is not received, dopamine firing will pause



Early Evidence for RPE - Humans

O'Doherty et al. [2003]

- Thirsty human subjects placed in fMRI scanner
- Shown novel visual symbols, which signalled 'neutral' and 'tasty' juice rewards
- Assumptions made to operationalize RPE
  - Reward: values of juice
  - Learning: through TD algorithm
- Resulting RPE signal then correlated with brain activity
- Positive correlation with activity in Ventral Striatum taken as supporting RPE hypothesis
  - Ventral Striatum rich in dopaminergic neurons

Problems with the Current Tests

- Several other theories for the role of dopamine
  - Salience hypothesis (e.g. Zink et al. 2003)
  - Incentive salience hypothesis (Berridge and Robinson, 1998)
  - Agency hypothesis (Redgrave and Gurney, 2006)
- These theories have been hard to differentiate
- Couched in terms of latent variable
  - ‘Rewards’, ‘Beliefs’, ‘Salience’, ‘Valence’ not directly observable
- Tests rely on ‘auxiliary assumptions’ - not central to the underlying theory
  - Experiments test both underlying theory *and* auxiliary assumptions
- Also different models tend to lead to very highly correlated predictions

An Axiomatic Approach

- Alternative: take an axiomatic approach to testing RPE hypothesis
  - A set of necessary and sufficient conditions on dopamine activity
  - Equivalent to the RPE model
  - Easily testable
- Similar to Samuelson's approach to testing utility maximization
  - Equivalent to the Weak Axiom of Revealed Preference
- Has several advantages
  - Provide a complete list of testable predictions of the RPE model
  - Non-parametric
  - Failure of particular axioms will aid model development

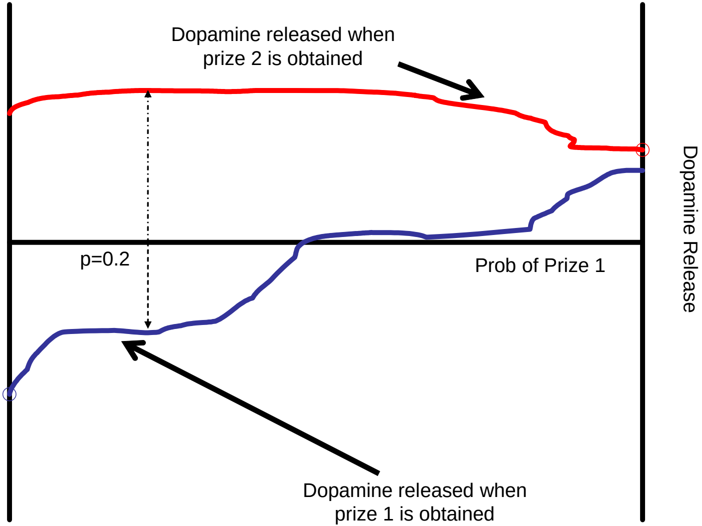
The Data Set

- Subjects receive prizes from lotteries:
  - Z: A space of prizes
  - Λ : Set of all lotteries on Z
- $\Lambda(z)$: Set of all lotteries whose support includes z
- e_z: Lottery that gives prize z with certainty

- Observable data is a Dopamine Release Function

$$\begin{aligned} \delta &: M \rightarrow \mathbb{R} \\ M &= \{(z,p) | z \in Z,\ p \in \Delta(z)\} \end{aligned}$$

$\delta(z,p)$ is dopamine activity when prize z is obtained from a lottery p


Dopamine released when
prize 2 is obtained
p=0.2
Prob of Prize 1
Dopamine Release
Dopamine released when
prize 1 is obtained

A Formal Model of RPE

The difference between how good an event was expected to be and how good it turned out to be

- Under what conditions can we find
  - A Predicted reward function: $b : \Lambda \to \mathbb{R}$
  - An Experienced reward function: $r : Z \to \mathbb{R}$
- Such that there is an aggregator function E
  - Represents the dopamine release function

$$\delta(z, p) = E(r(z), b(p))$$

  - Is increasing in experienced and decreasing in predicted reward.
  - Obeys basic consistency

$$r(z) = b(e_z)$$

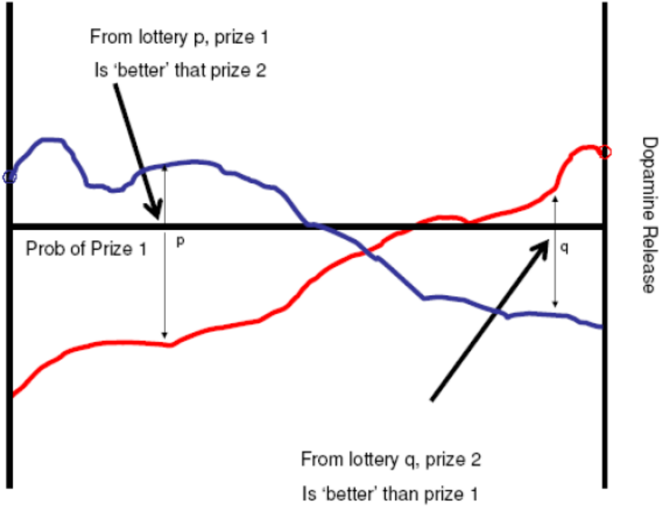
  - Treats 'no surprise' consistently:

$$E(x, x) = E(y, y)$$

  - Special case:

$$\delta(z, p) = r(z) - b(p)$$

Necessary Condition 1: Consistent Prize Ordering



Necessary Condition 1: Consistent Prize Ordering

- Consider two prizes, z and w
- Say that, when z is received from some lottery p, more dopamine is released than when w is received from p
- Implies higher 'reward' for z than w
- Implies that z should give more dopamine than w when received for *any* lottery q
- **Axiom A1: Coherent Prize Dominance**

$$\text{for all } (z,p),(w,p),(z,q),(w,q) \quad \in \quad M$$
$$\delta(z,p) > \delta(w,p) \Rightarrow \delta(z,q) > \delta(w,q)$$

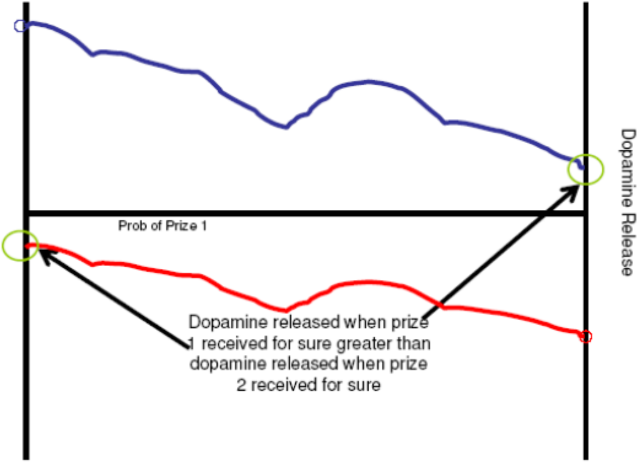
Necessary Condition 2: Coherent Lottery Dominance

Necessary Condition 2: Consistent Lottery Ordering

- Consider two lotteries p and q and a prize z which is in the support of p and q
- Say that more dopamine is released when z is obtained from p that when it is obtained from q
- Implies that predicted reward of p must be *lower* that that of q
- Implies that whenever the same prize is obtained from p and q the dopamine released should be higher from lottery p than from lottery q
- **Axiom A2: Coherent Lottery Dominance**

$$\text{for all } (z,p),(w,p),(z,q),(w,q) \in M$$
$$\delta(z,p) > \delta(z,q) \Rightarrow \delta(w,p) > \delta(w,q)$$

Necessary Condition 3: No Surprise Equivalence



Necessary Condition 3: Equivalence of Certainty

- ‘Reward Prediction Error’ is a comparison between predicted reward and actual reward
- If you know exactly what you are going to get, then there is no reward prediction error
- This is true whatever the prize we are talking about
- Thus, the reward prediction error of any prize should be zero when received for sure:
- **Axiom A3: No Surprise Equivalence**

$$\delta(z, e_z) = \delta(w, e_w) \ \forall \ z, w \in Z$$

A Representation Theorem

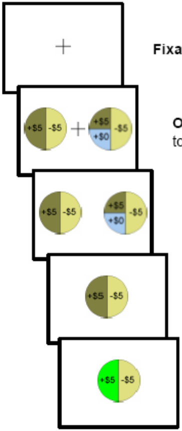
- In general, these conditions are necessary, but not sufficient for an RPE representation
- However, in the special case where we look only at lotteries with two prizes they are
- **Theorem 1:**
  If $|Z| = 2$, a dopamine release function δ satisfies axioms A1-A3 **if and only if** it admits an RPE representation
- Thus, in order to test RPE in case of two prizes, we need only to test A1-A3

Aim

- Generate observations of δ in order to test axioms
- Use a data set containing:
  - Two prizes: win \$5, lose \$5
  - Five lotteries: $p \in \{0, 0.25, 0.5, 0.75.1\}$
- Do not observe dopamine directly
  - Use fMRI to observe activity in the Nucleus Accumbens
  - Brain area rich in dopaminergic neurons

Experimental Design

Task design



Fixation: 12 seconds

Options: 3.6 seconds
to view options

Lottery 1 EV: $0
Lottery 2 EV: -$1.25

Choice selection: 1.2 seconds
to make a choice by button press
(fixation cross extinguished)

Choice: 8.4 seconds to
view the choice just made

Outcome: 3.6 seconds to
view outcome of choice
(outcome illuminated)

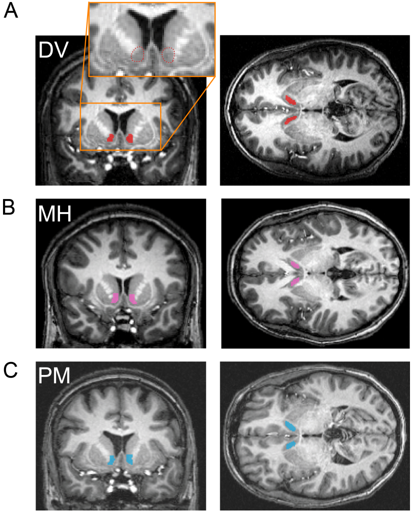
Trial length:

~29 seconds

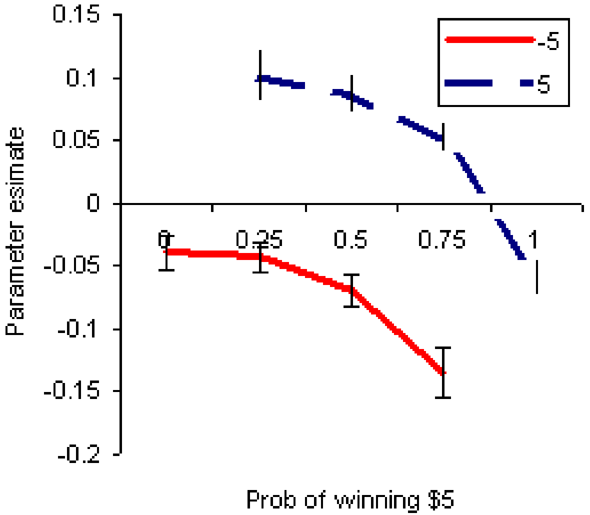
Experimental Details

- 14 subjects (2 dropped for excess movement)
- ‘Practice Session’ (outside scanner) of 4 blocks of 16 trials
- 2 ’Scanner Sessions’ of 8 blocks of 16 trials
- For Scanner Sessions, subjects paid \$35 show up fee, + \$100 endowment + outcome of each trial
- In each trial, subject offered one option from ‘Observation Set’ and one from a ‘Decoy Set’

- Need to determine which area of the brain is the Nucleus Accumbens
- Two ways of doing so:
  - Anatomical ROIs: Defined by location
  - Functional ROIs: Defined by response to a particular stimulus
- We concentrate on anatomical ROI, but use functional ROIs to test results

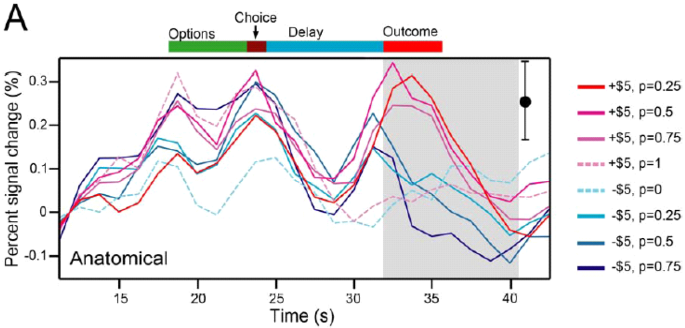

A
DV
B
MH
C
PM

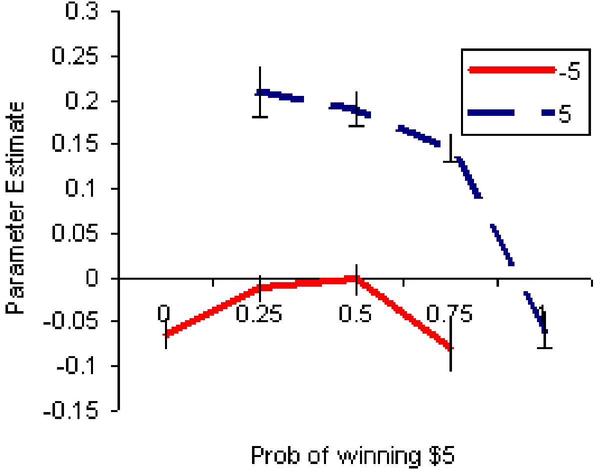
- We now need to estimate the function $\bar{\delta}$ using the data
- Use a between-subject design
  - Treat all data as coming from a single subject
- Create a single time series for an ROI
  - Average across voxels
  - Convert to percentage change from session baseline
- Regress time series on dummies for the revelation of each prize/lottery pair
  - $\bar{\delta}(x, p)$ is the estimated coefficient on the dummy which takes the value 1 when prize x is obtained from lottery p

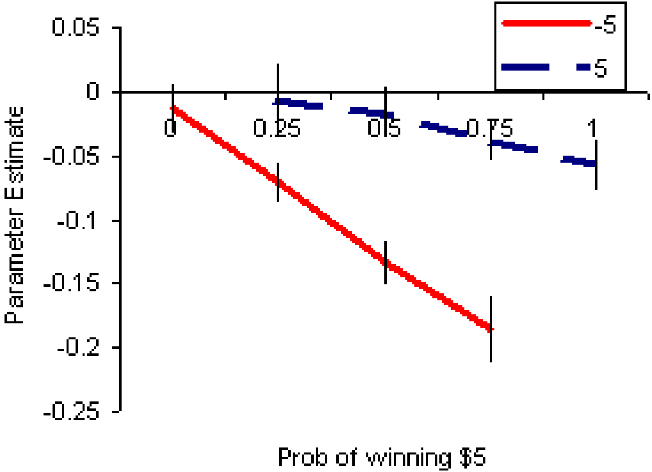

0.15
0.1
0.05
0
-0.05
-0.1
-0.15
-0.2
Parameter esimate
0
0.25
0.5
0.75
1
-5
5
Prob of winning $5

Results

- Axioms hold
- Nucleus Accumbens activity in line with RPE model
- Experienced and predicted reward ‘sensible’

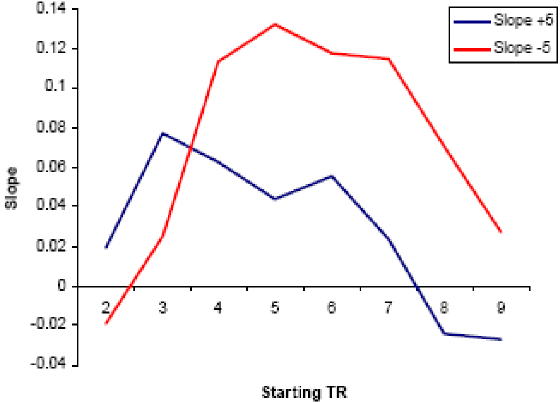

A
Options
Choice
Delay
Outcome
Percent signal change (%)
0.3
0.2
0.1
0
-0.1
Anatomical
15
20
25
30
35
40
Time (s)
+$5, p=0.25
+$5, p=0.5
+$5, p=0.75
+$5, p=1
-$5, p=0
-$5, p=0.25
-$5, p=0.5
-$5, p=0.75


0.3
0.25
0.2
0.15
0.1
0.05
0
-0.05
-0.1
-0.15
Parameter Estimate
0
0.25
0.5
0.75
1
-5
5
Prob of winning $5



-5
5
0.05
0
-0.05
-0.1
-0.15
-0.2
-0.25
0
0.25
0.5
0.75
1
Parameter Estimate
Prob of winning $5

Two Different Signals?



Key Results

- fMRI activity in Nucleus Accumbens does satisfy the necessary conditions for an RPE encoder
- However, this aggregate result may be the amalgamation of two separate signals
  - Vary in temporal lag
  - Vary in magnitude

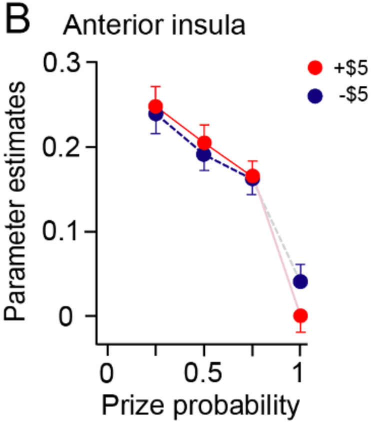
- Axioms + experimental results tell us we can assign numbers to events such that NAcc activity encodes RPE according to those numbers
- Can we use these numbers to make inferences about beliefs and rewards?
    - Are they 'beliefs' and 'rewards' in the sense that people usually use the words?
    - Can we find any 'external validity' with respect to other observables?
        - Behavior?
        - Obviously rewarding events?
    - Can we then generalize to other situations?

- New way of observing beliefs
- Makes ‘surprise’ directly observable
- Insights into mechanisms underlying learning
- Building blocks of ‘utility’

Conclusion

- We provide evidence that NAcc activity encodes RPE
- Can recover consistent dopaminergic ‘beliefs’ and ‘rewards’
- Potential for important new insights into human behavior and ‘state of mind’

What about other brain areas?

- 

- Anterior Insula seems to record magnitude of RPE