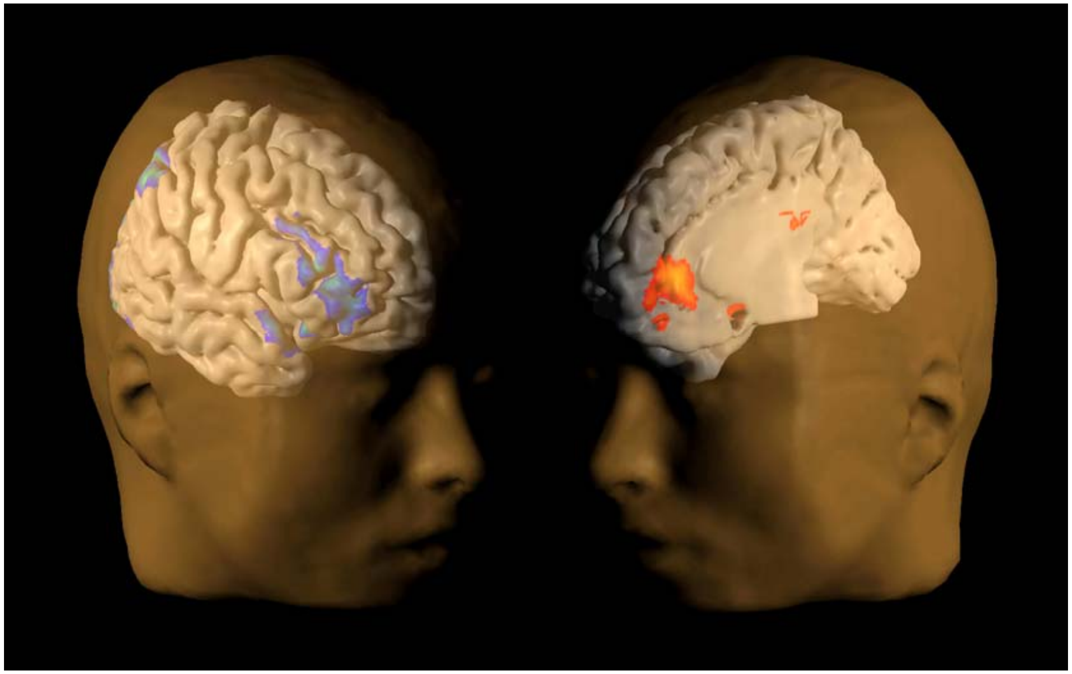
Neuroeconomics
Laibson Lecture 2



David Laibson

January 2010
AEA

This deck contains “hidden slides” that were not presented in course.
View in “slideshow” mode to avoid these hidden slides.

Outline:

- Neuroeconomics: definition
- Multiple Systems Hypothesis

Neuroeconomics: definition.

Definition: Neuroeconomics is the study of the *biological microfoundations* of *economic cognition*.

- *Biological microfoundations* are neurochemical mechanisms and pathways, like brain systems, neurons, genes, and neurotransmitters.
- *Economic cognition* is cognitive activity that is associated with economic perceptions, beliefs and decisions, including mental representations, emotions, expectations, learning, memory, preferences, decision-making, and behavior.

The “field”

- Less than a decade old
- About 200 neuroscientists and economists are active
- Roughly a 2:1 mix of neuroscientists and economists
- This contrast’s with behavioral economics, where it’s a one-sided game (mostly economists).
- In neuroeconomics, there is a mix of behavioral economists (Camerer, Fehr, Laibson, Loewenstein, Prelec), classical economists (Dickhaut, McCabe, Rustichini), and some people who are hard to categorize (Bernheim, Rangel).
- You don’t need to believe people are irrational to be a neuroeconomist, though as we’ll see believe it helps.
- Annual conference and a “society.”

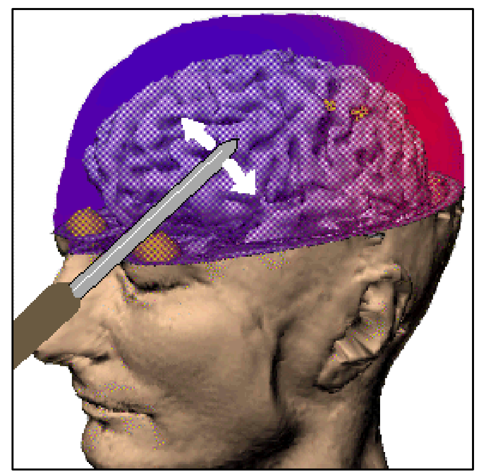
Neuroeconomics: methods

- Animal behavioral studies (e.g., addicting rats to cocaine, loss aversion in monkeys; Chen et al 2006)
- Studies of children (e.g. Mischel et al 1975)
- Studies of children with autism (e.g. Sally et al 2001).
- External physiological measurement (e.g., pupil dilation, voice tone, facial expression, skin conductance, heart rate)
- Cognitive load (remember this seven digit number and do a task; e.g. Shiv and Fedorhikin 1999)

More methods...

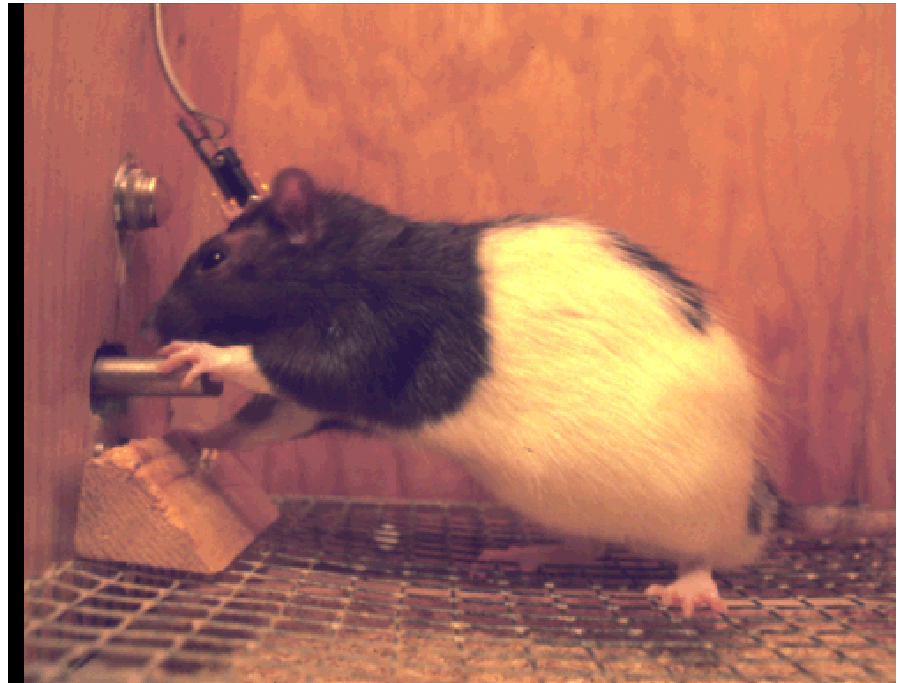
Lesions, localized damage, gene knockout…

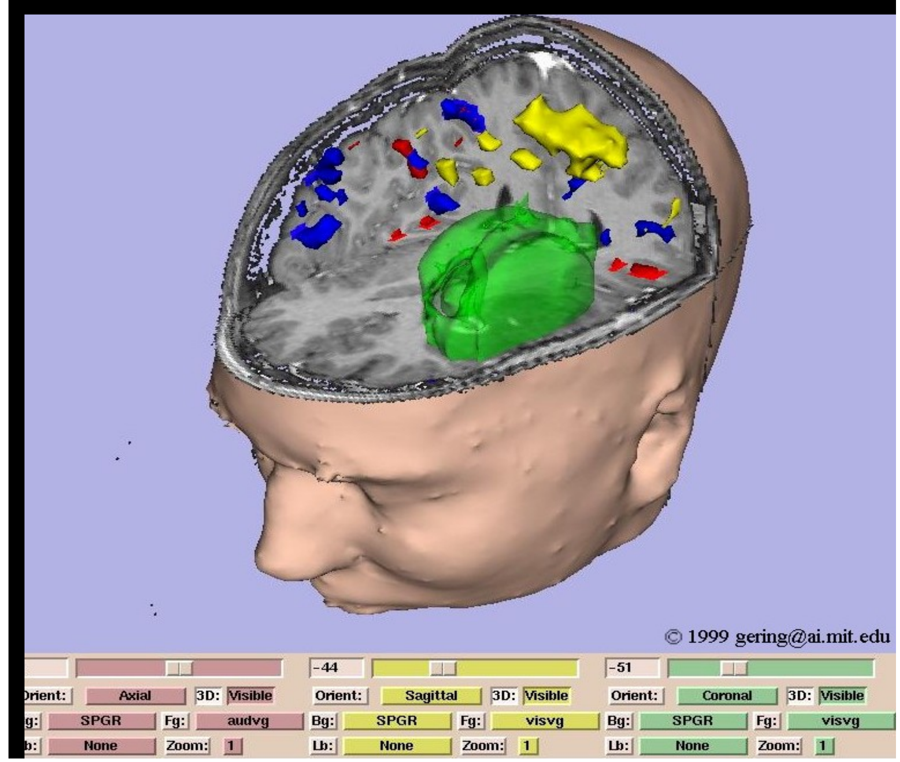
- e.g., experimental destruction of both amygdalas in an animal tames the animal, making it sexually inactive and indifferent to danger like snakes or other aggressive members of its own species
- e.g., humans with lesions of the amygdala lose affective (i.e. emotional) meaning
- e.g., knocking out the gene that makes a key protein for amygdala function makes rats relatively fearless
- e.g., hippocampus removal prevents experiences from being encoded in long-term memory



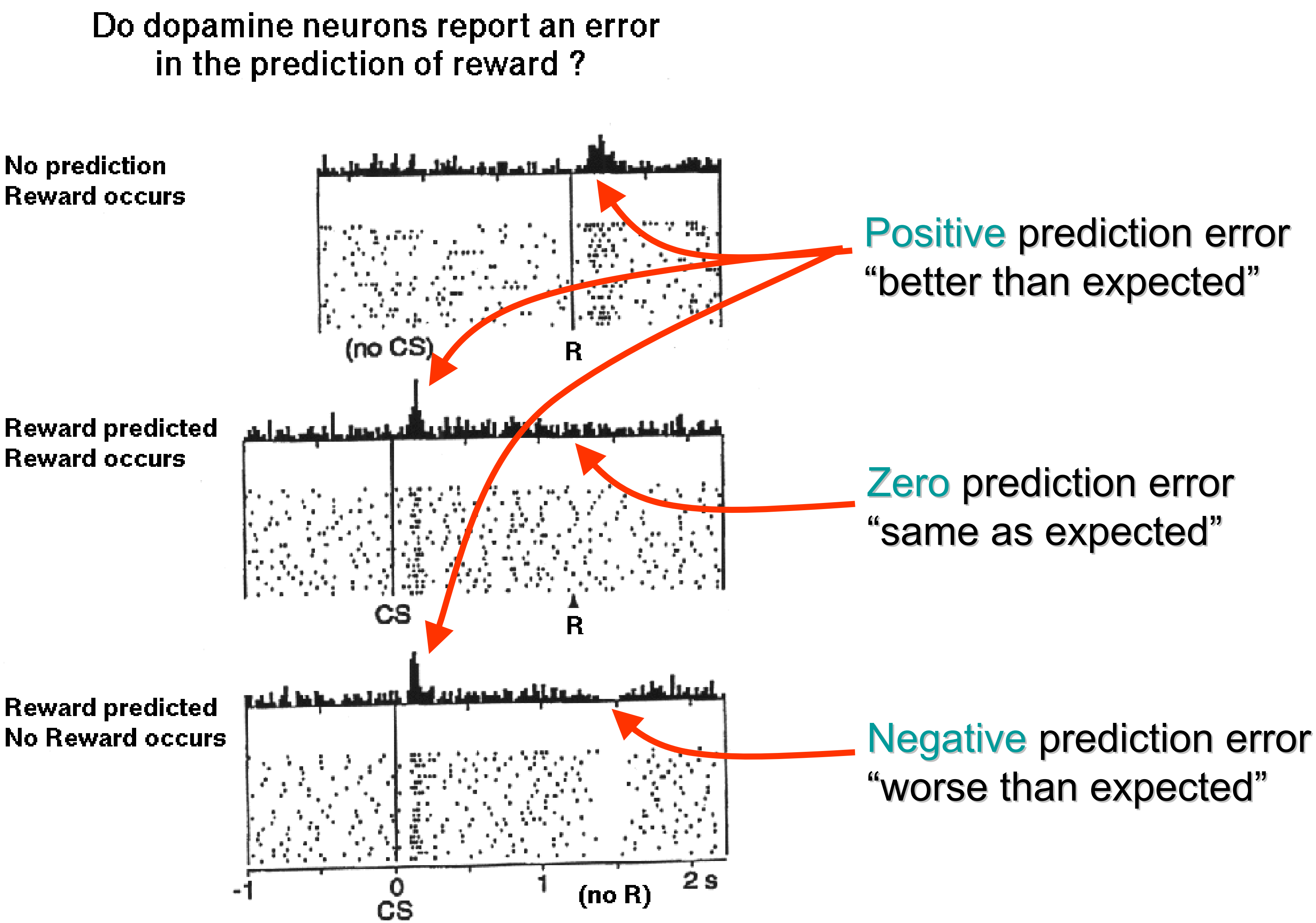
More methods

- Brain stimulation (e.g., electrical stimulation of the amygdala elicits violence and aggressivity; at special loci, electrical brain stimulation is highly reinforcing)
- Single neuron measurement (e.g., track high frequency dopamine release in animal models, Schultz et al, Glimcher)
- Transcranial magnetic stimulation (TMS)
- EEG, PET, fMRI… (taking pictures of the *active* brain; e.g., McCabe, Houser, Ryan, Smith, and Trouard 2001).





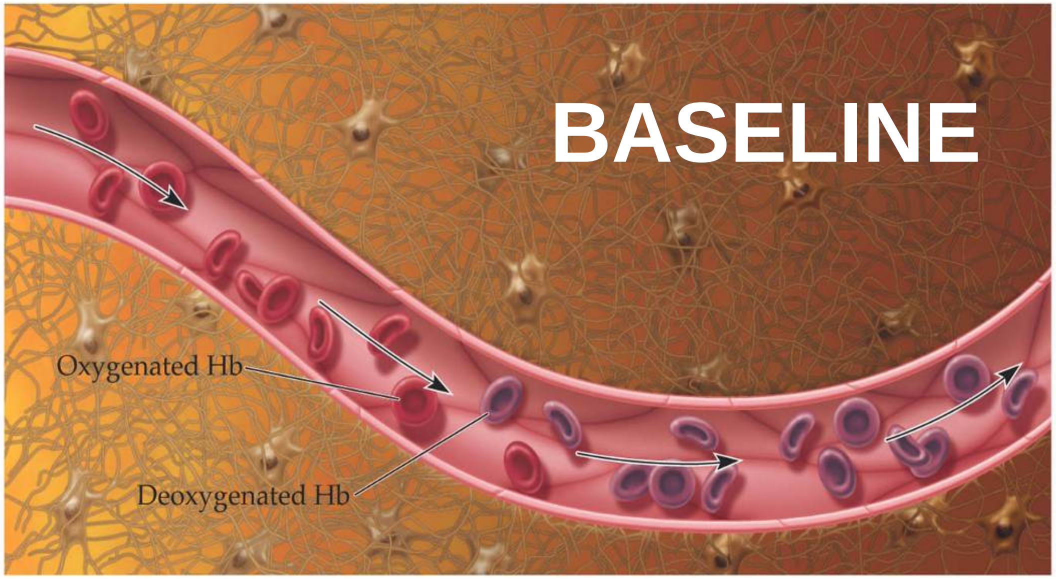
Schultz, Dayan, Montague (1997) *Science* 275:1593-1599

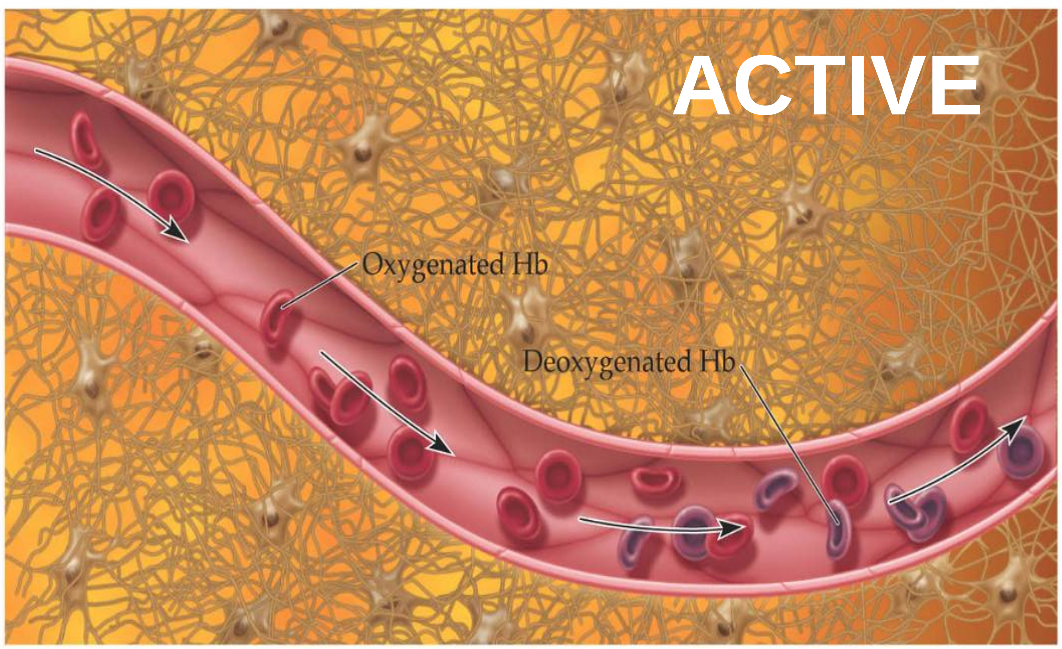


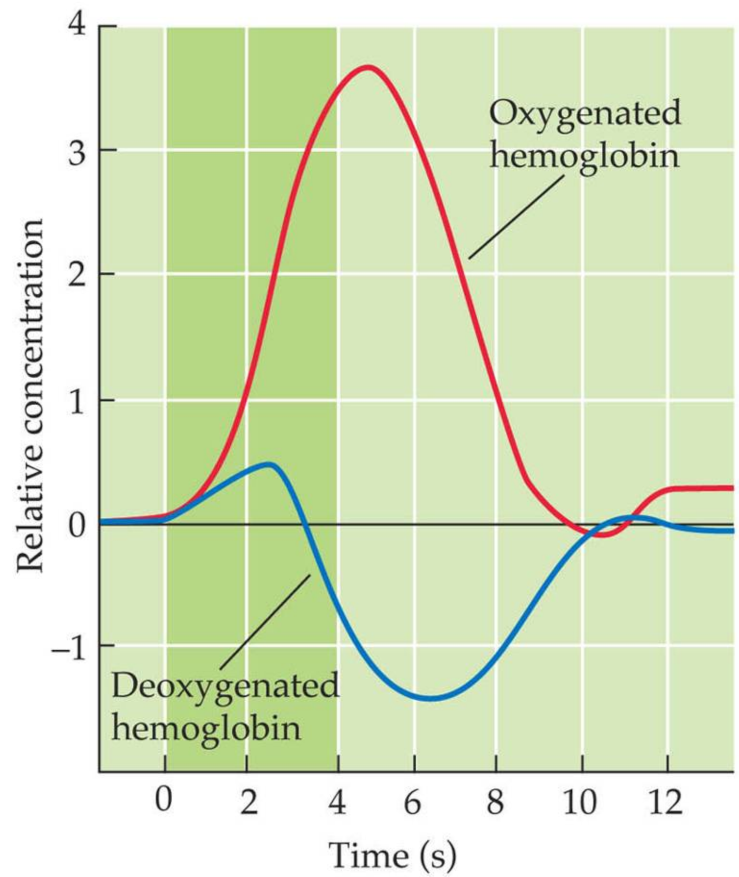
Source: Scott Huettel

The fMRI Blood-Oxygenation-Level-Dependent (BOLD) Response

Increased neuronal activity results in increased MR (T_2) signal*

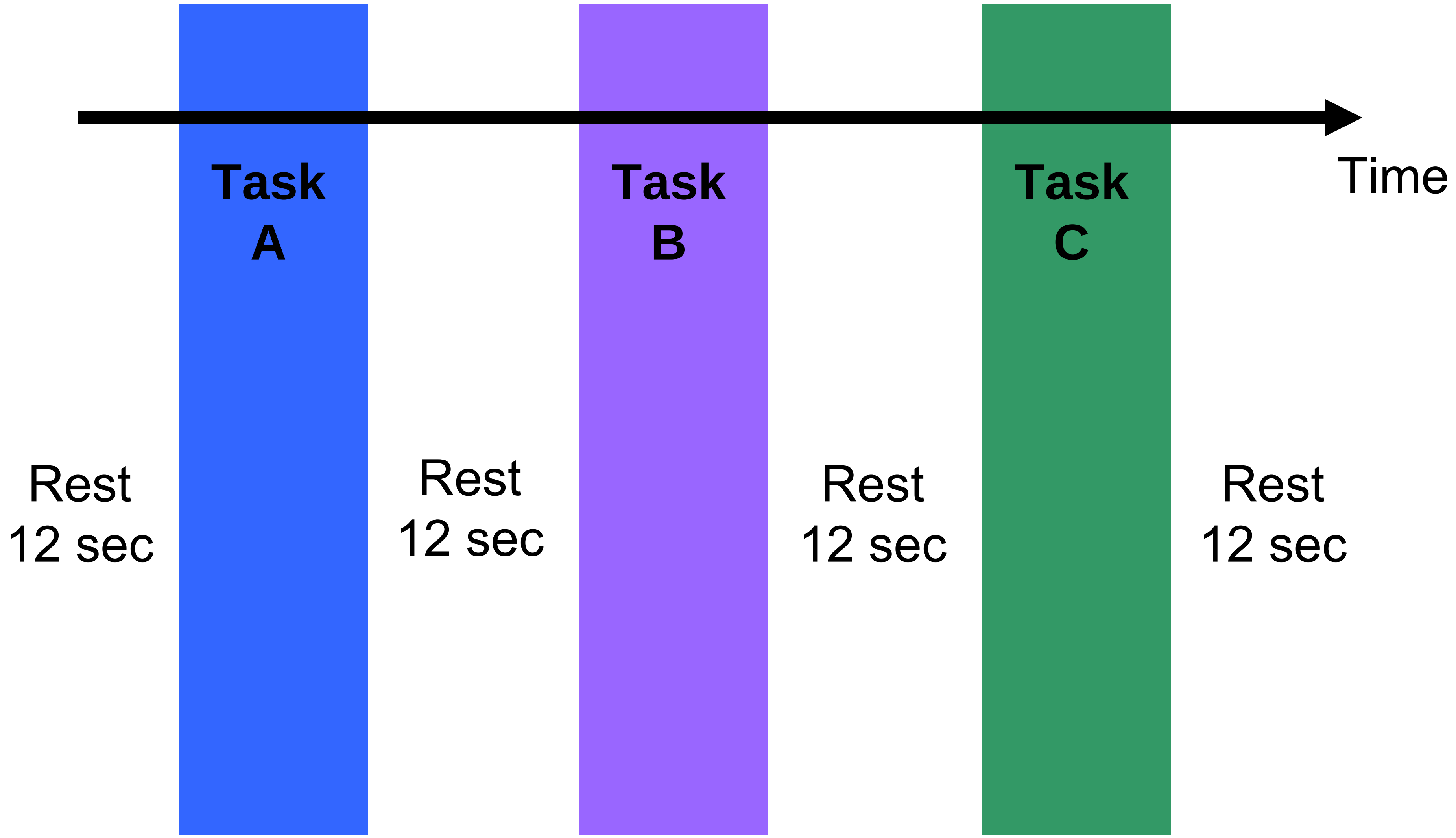






Source: Scott Huettel

Basic experimental design



Basic econometric methodology

- Divide brain into 10,000 voxels (cubes 3 mm on edge)
- Measure blood flow at the voxel level (BOLD signal)
- Run “regressions” (general linear model) relating BOLD signal to covariates:

  $BOLD_{v,i,t}$ = FE_i + $controls_t$ + task $dummy_t$
- Indexes for voxel (v), subject (i), and time (t)
- Controls: time in scanner, lagged reward event, etc.
- event dummy: decision, experience, event, etc.
- Analogous method: “contrast”

$$\text{Contrast at voxel } v = \sum_{i \in I} \left(BOLD_{v,i,t} - BOLD_{v,i,t'} \right)$$

Unusual features of neuroeconomic experiments

- Extensive *pre-scanner* piloting (scanner time is $$$)
- Only need about 20-40 scanned subjects (not that expensive after all – about $20,000 for scanner costs)
- Claustrophobia and noise issues
- Try to implement within-subject designs
- Tasks can’t take too long (max about 45 minutes)

Continued…

- Can do almost anything in the scanner
  - Play games
  - Play real-time games with other subjects (so-called hyperscans)
  - Eat
  - Experience pain
- Analysis is subject to methodological criticisms (lots of hypothesis tests) – use conservative thresholds, such as 5+ contiguous voxels and $\alpha = 0.001$ or smaller
- Most economists need neuroscience collaborators to implement neuroimaging experiments

Why should economists study the brain?

1. Because we care about the brain for its own sake? (No.)
2. Because we hope to improve our measurements of utility? (Maybe.)
3. Because the study of the brain will direct and catalyze the development of new models, speeding up the rate of progress in model development. (Yes.)
4. Because neuroscience will provide new empirical methods that will sometimes provide new empirical tests. (Yes.)
5. Because we will eventually be able to use neuroscience measurements to help people better understand and manage themselves. (Yes.)

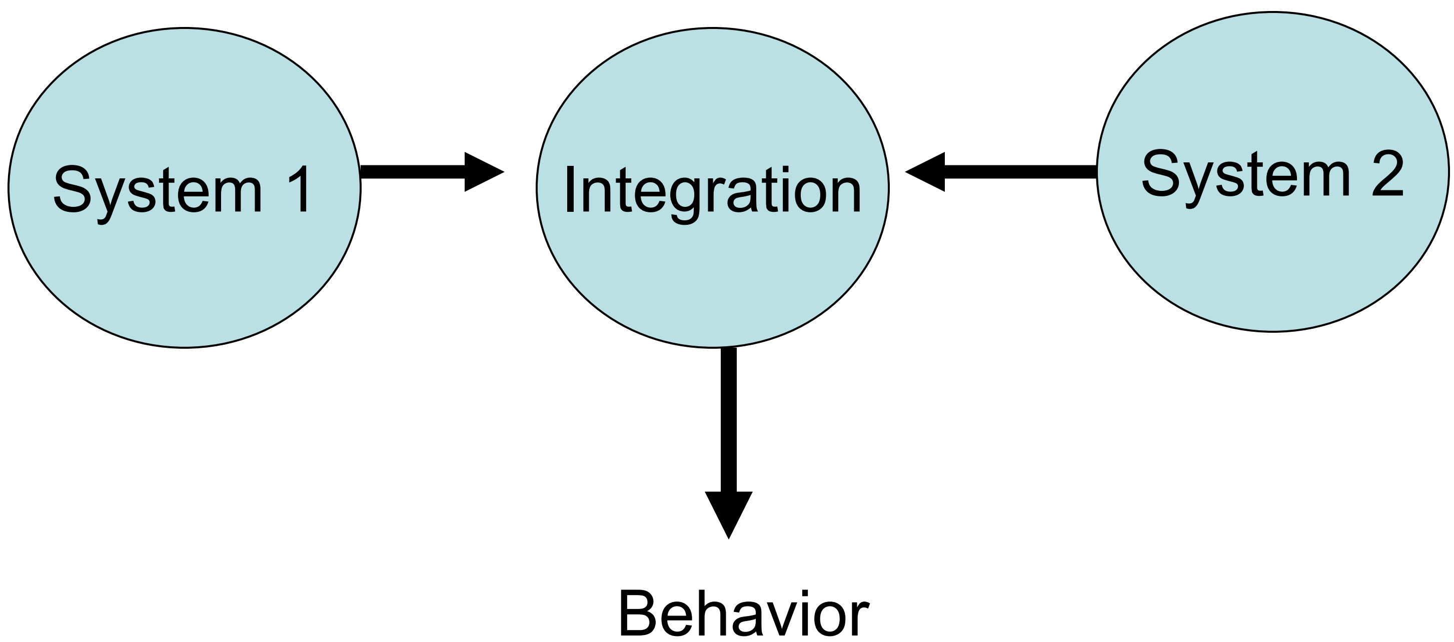
An example:
The Multiple Systems Hypothesis

- **Statement of Hypothesis**
- **Variations on a theme**
- **Caveats**
- **Illustrative predictions**
  - **Cognitive load manipulations**
  - **Willpower manipulations**
  - **Affect vs. analytic manipulations**
  - **Cognitive Function**
  - **Development**
  - **Neuroimaging**
- **Directions for future research**

Statement of Multiple Systems Hypothesis (MSH)

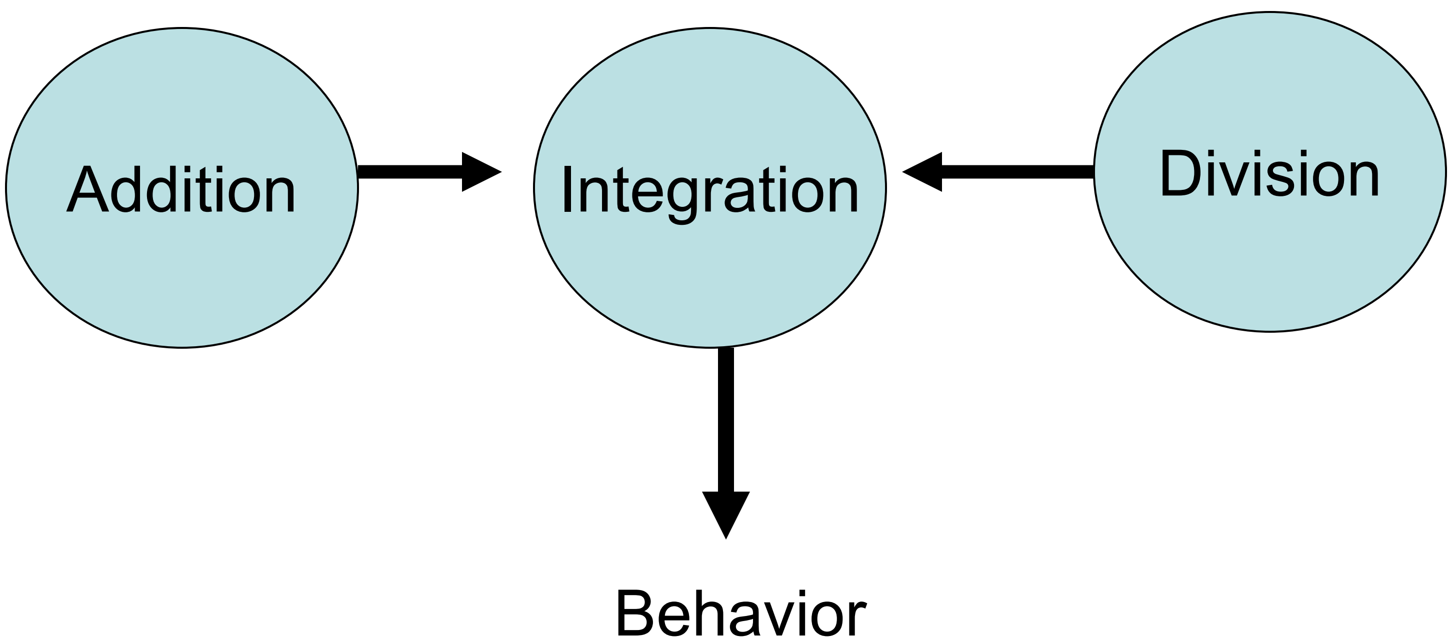
- The brain makes decisions (e.g. constructs value) by integrating signals from multiple systems
- These multiple systems process information in qualitatively different ways and in some cases differentially weight attributes of rewards (e.g., time delay)

An (oversimplified) multiple systems model



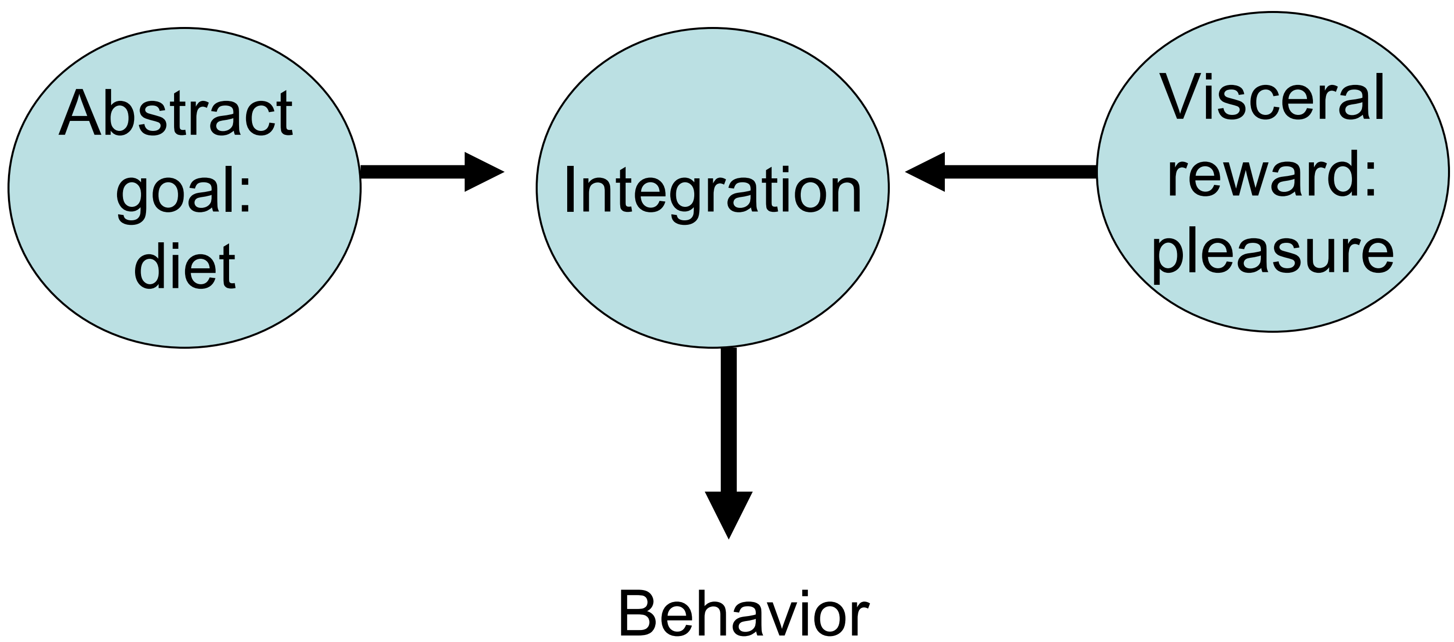
An uninteresting example

What is 6 divided by 3?



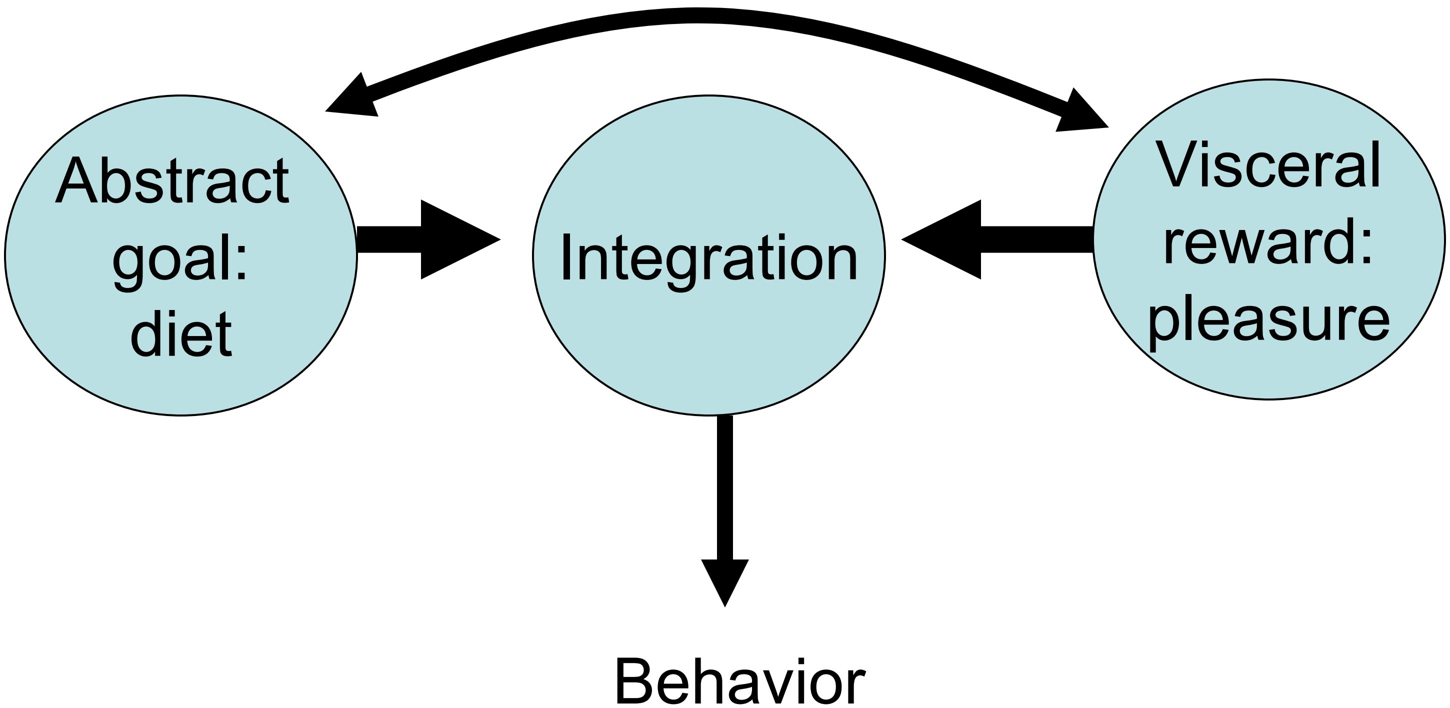
A more interesting example

Would you like a piece of chocolate?



A more interesting example

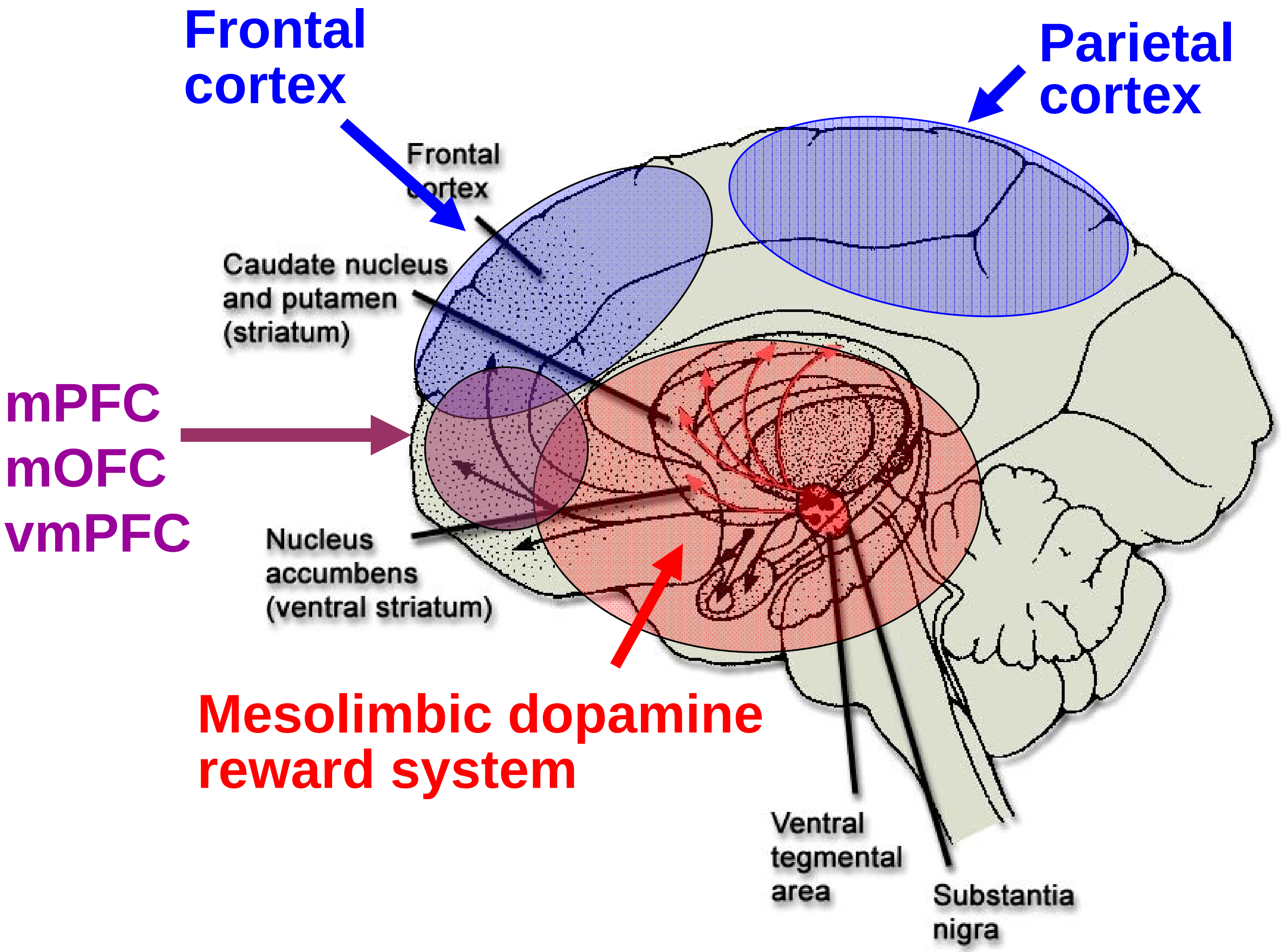
Would you like a piece of chocolate?



Variations on a theme

- Interests vs passions (Smith)
- Superego vs Ego vs Id (Freud)
- Controlled vs Automatic (Schneider & Shiffrin, 1977; Benhabib & Bisin, 2004)
- Cold vs Hot (Metcalfe and Mischel, 1979)
- System 2 vs System 1 (Frederick and Kahneman, 2002)
- Deliberative vs Impulsive (Frederick, 2002)
- Conscious vs Unconscious (Damasio, Bem)
- Effortful vs Effortless (Baumeister)
- Planner vs Doer (Shefrin and Thaler, 1981)
- Patient vs Myopic (Fudenburg and Levine, 2006)
- Abstract vs Visceral (Loewenstein & O'Donoghue 2006; Bernheim & Rangel, 2003)
- PFC & parietal cortex vs Mesolimbic dopamine (McClure et al, 2004)

Affective vs. Analytic Cognition



Commonalities between classification schemes

Affective system

- fast
- unconscious
- reflexive
- myopic

Analytic system

- slow
- conscious
- reflective
- forward-looking
- (but still prone to error: heuristics may be analytic)

Relationship to quasi-hyperbolic model

- Hypothesize that mesolimbic dopamine system is impatient.
- Hypothesize that the fronto-parietal system is patient
- Here's one implementation of this idea:

$$U_t = u_t + \beta\,[\delta u_{t+1} + \delta^2 u_{t+2} + \delta^3 u_{t+3} + \ldots]$$

$$(1/\beta)U_t = (1/\beta)u_t + \delta u_{t+1} + \delta^2 u_{t+2} + \delta^3 u_{t+3} + \ldots$$

$$(1/\beta)U_t = \underbrace{(1/\beta - 1)u_t}_{\text{limbic}} + \underbrace{[\delta^0 u_t + \delta^1 u_{t+1} + \delta^2 u_{t+2} + \delta^3 u_{t+3} + \ldots]}_{\text{fronto-parietal cortex}}$$

Relationship to quasi-hyperbolic model

- Hypothesize that the fronto-parietal system is patient
- Hypothesize that mesolimbic system is impatient.
- Then integrated preferences are quasi-hyperbolic

		now	t+1	t+2	t+3	
PFC		1	1	1	1	…
Mesolimbic		1	0	0	0	…
Total		2	1	1	1	…
Total normed		1	1/2	1/2	1/2	…

Caveats

- $N \geq 2$
- The systems do not have well-defined boundaries (they are densely interconnected)
- Maybe we should not say “system,” but should instead say “multiple processes”
- Some systems may not have a value/utility representation
  - Making my diet salient is not the same as assigning utils/value to a Devil Dog
- If you look downstream enough, you’ll find what looks like an integrated system

Predictions

- **Cognitive Load Manipulations**
  - Shiv and Fedorikhin (1999), Hinson, Jameson, and Whitney (2003)
- **Willpower manipulations**
  - Baumeister and Vohs (2003)
- **Affect vs. analytic manipulations**
  - Rodriguez, Mischel and Shoda (1989)
- **Cognitive Function**
  - Benjamin, Brown, and Shapiro (2006), Shamosh and Gray (forth.)
- **Developmental Dynamics**
  - Green, Fry, and Myerson (1994), Krietler and Zigler (1990)
- **Neuroimaging Studies**
  - Tanaka et al (2004), McClure et al (2004), Hariri et al (2006), McClure et al (2007), Kabel and Glimcher (2007), Hare, Camerer, and Rangel (2009)

Cognitive Load Should Decrease Self-regulation

Shiv and Fedorikhin (1999)

- Load manipulated by having people keep either a 2-digit or 7-digit number in mind during experiment
- Subjects choose between cake or fruit-salad

Processing burden	% choosing cake
Low (remember only 2 digits)	41%
High (remember 7 digits)	63%

Cognitive Load Should Increase Discounting

Hinson, Jameson, and Whitney (2003)

- Task: Subjects choose one reward from a set of two or more time-dated rewards.
- Some subjects are under cognitive load: hold 5-digit string in memory

	One month discount rate*
Control (2 items):	26.3%
Treatment (2 items + load):	49.8%
Treatment (3 items):	48.4%

*Discount rate for 1 month is ln(1+k), where discount function is 1/(1+kt).

Willpower should be a domain general and exhaustable resource

Muraven, Tice and Baumeister (1998)

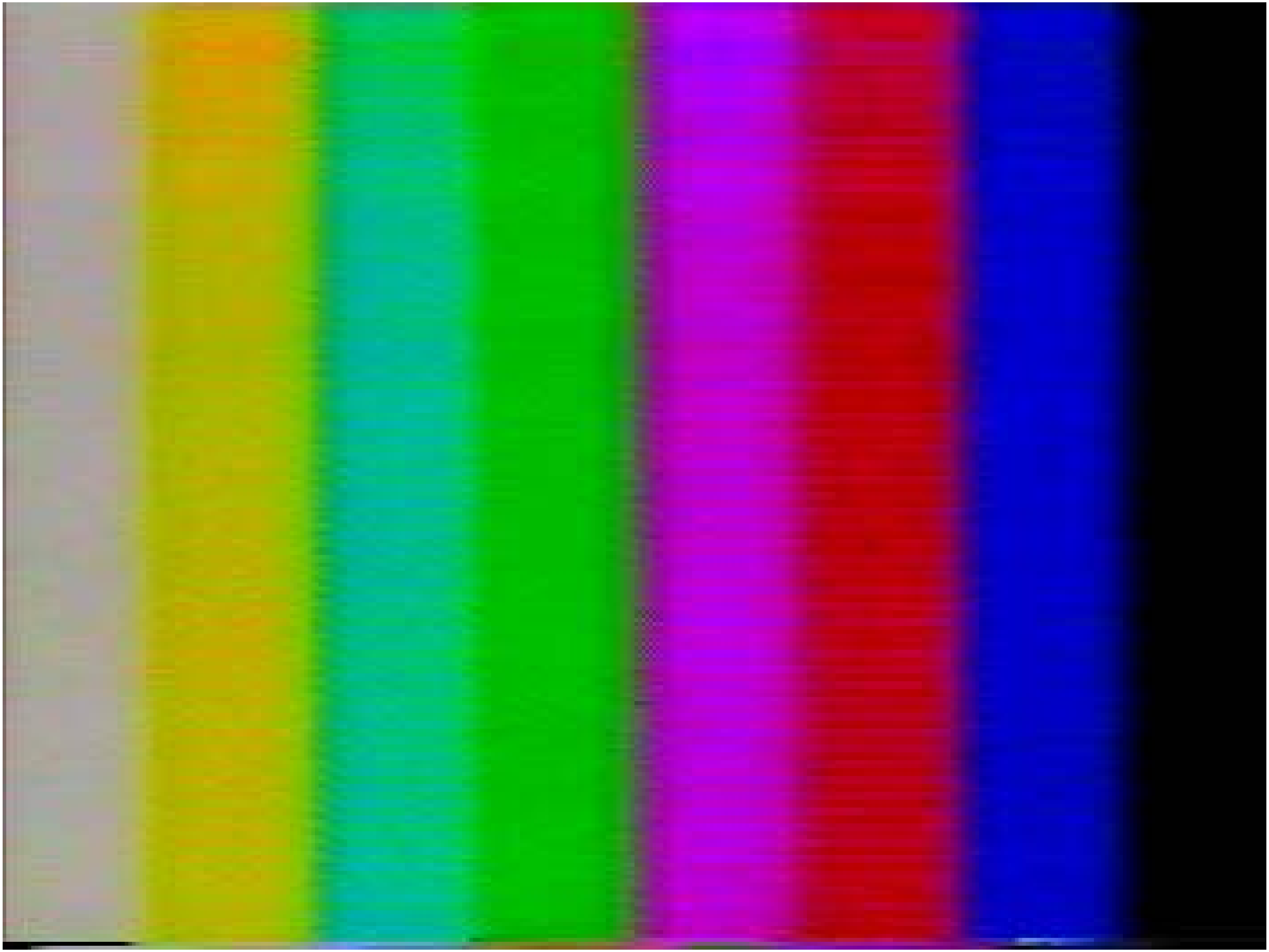
- All subjects watch upsetting video clip
  - Treatment subjects are asked to control emotions and moods
  - Control subjects are not asked to regulate emotions
- All subjects then try to squeeze a handgrip as long as possible
- Treatment group gives up sooner on the grip task
- Interpretation: executive control is a scarce resource that is depleted by use
- Many confounds and many variations on this theme

Affect Augmentation/Reduction Should Predictably Change Patience

Rodriguez, Mischel and Shoda (1989)

- Task: Children *try* to wait 15 minutes, to exchange a smaller immediate reward for a larger delayed reward.
- Manipulations:
  - Control
  - Affect Augmentation: exposure to rewards
  - Affect Reduction: represent the delayed reward abstractly (pretzels are logs, marshmallows are clouds)
- Results:
  - Ability to wait goes down after affect augmentation
  - Ability to wait goes up after affect reduction

Delay of Gratification Paradigm



Individuals with greater analytic intelligence should be more patient

Shamosh and Gray (forthcoming)

- Meta-analysis of the relationship between delay discounting and analytic intelligence
  - e.g. g, IQ, math ability
- In 24 studies, nearly all found a negative relationship
- Quantitative synthesis across 26 effect sizes produce a correlation of -0.23

- See also new work by Burks et al. (2009) which finds this pattern in a sample of truck drivers.

Shamosh and Gray (forthcoming)

“Delay discounting paradigms appear to require some of the specific abilities related to both working memory and intelligence, namely the active maintenance of goal-relevant information in the face of potentially distracting information, as well as the integration of complex or abstract information.”

1. Manage affect: “controlling one’s excitement over the prospect of receiving an immediate reward”
   – Shift attention away from affective properties of rewards
   – Transform reward representation to make them more abstract
2. Deploy strategies
3. Recall past rewards and speculate about future rewards

Individuals with greater analytic intelligence should be more patient

Benjamin, Brown, and Shapiro (2006)

- Choice task: “500 pesos today vs. 550 pesos in a week”
- One standard deviation increase in a subject’s math test score is associated with a 9.3% increase in the subject’s likelihood of choosing patiently

Patience should track PFC development

- PFC undergoes development more slowly than other brain regions
- Changes in patience track timing of PFC development
  - Green, Fry, and Myerson (1994)
  - Krietler and Zigler (1990)
  - Galvan (2008)
- Cross-species evidence is also supportive
  - Monkeys have a 50% discount for juice rewards delayed 10 seconds (Livingstone 2009)

Neural activity in frontoparietal regions should be differentiated from neural activity in mesolimbic dopamine regions

- Not clear.
- At least five relevant studies:
  - McClure et al (2004) – supportive
  - McClure et al (2007) – supportive
  - Kable and Glimcher (2007) – critical
  - Hariri et al (2007) – supportive
  - Hare, Camerer, and Rangel (2009) – integration
- And one other study that is related (and supportive):
  - Tanaka et al (2004): different behavioral task

Tanaka et al (2004)

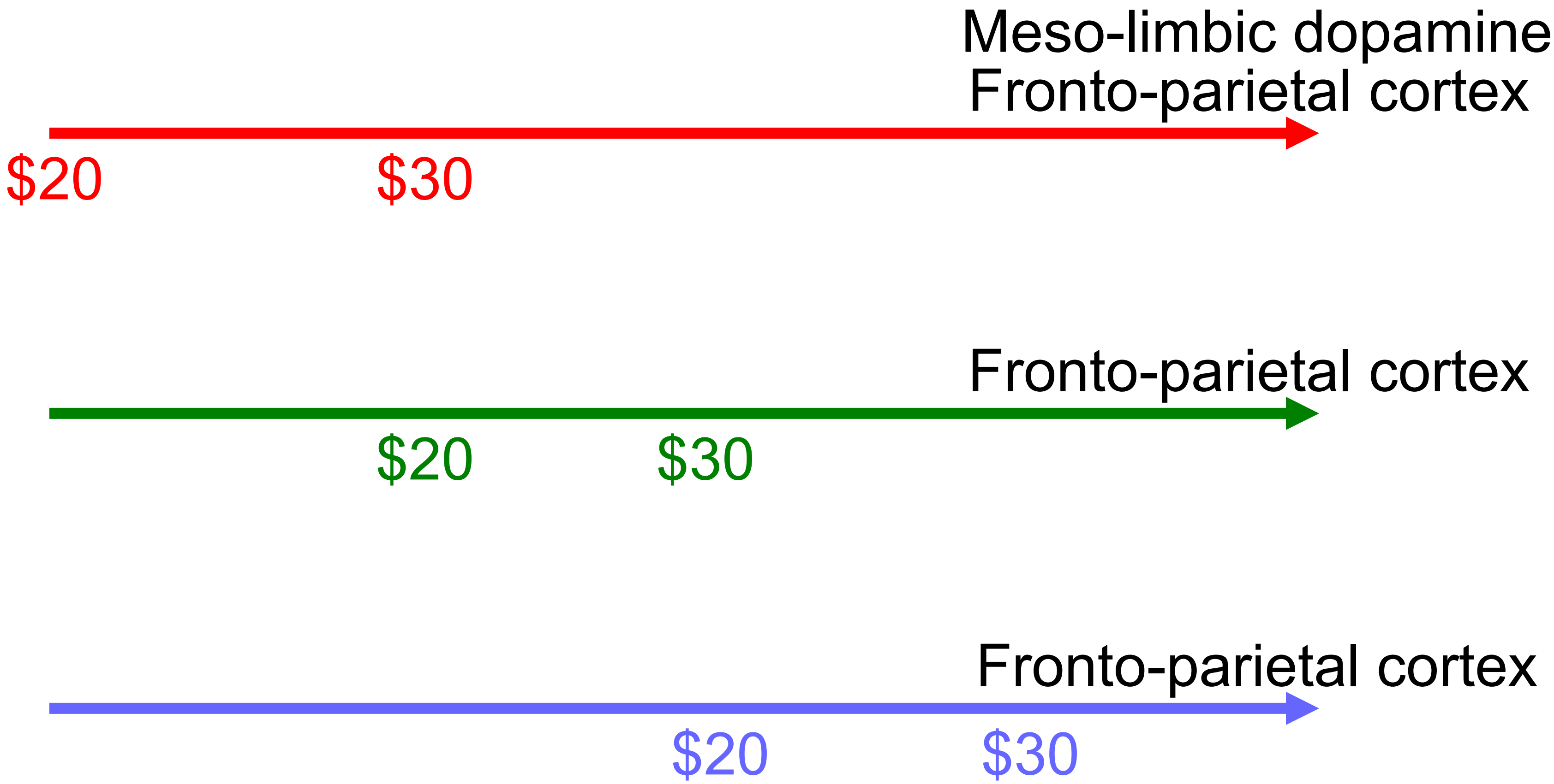
- LONG condition: optimal to take a small payoff hit now to get a delayed payoff benefit later in time.
- SHORT condition: optimal to take a small payoff now.
- They contrast activation in the SHORT condition to activation in a control 'NO' condition in which there was no reward: They get activation in the OFC, insula, occipitotemporal area, the striatum, the globus pallidus and the medial cerebellum.
- Next, they contrast activation in the LONG and SHORT conditions. In this contrast an increased activity was observed in VLPFC and the DLPFC, and the premotor cortex (like us) , the Inferior parietal cortex (IPC), the striatum, the dorsal raphe nucleus, the lateral cerebellum , the PCC, and the subthalamic nucleus.

- In sum, there is a reasonably good overlap of our Beta areas and their SHORT system activation and our Delta areas and their LONG system activation.

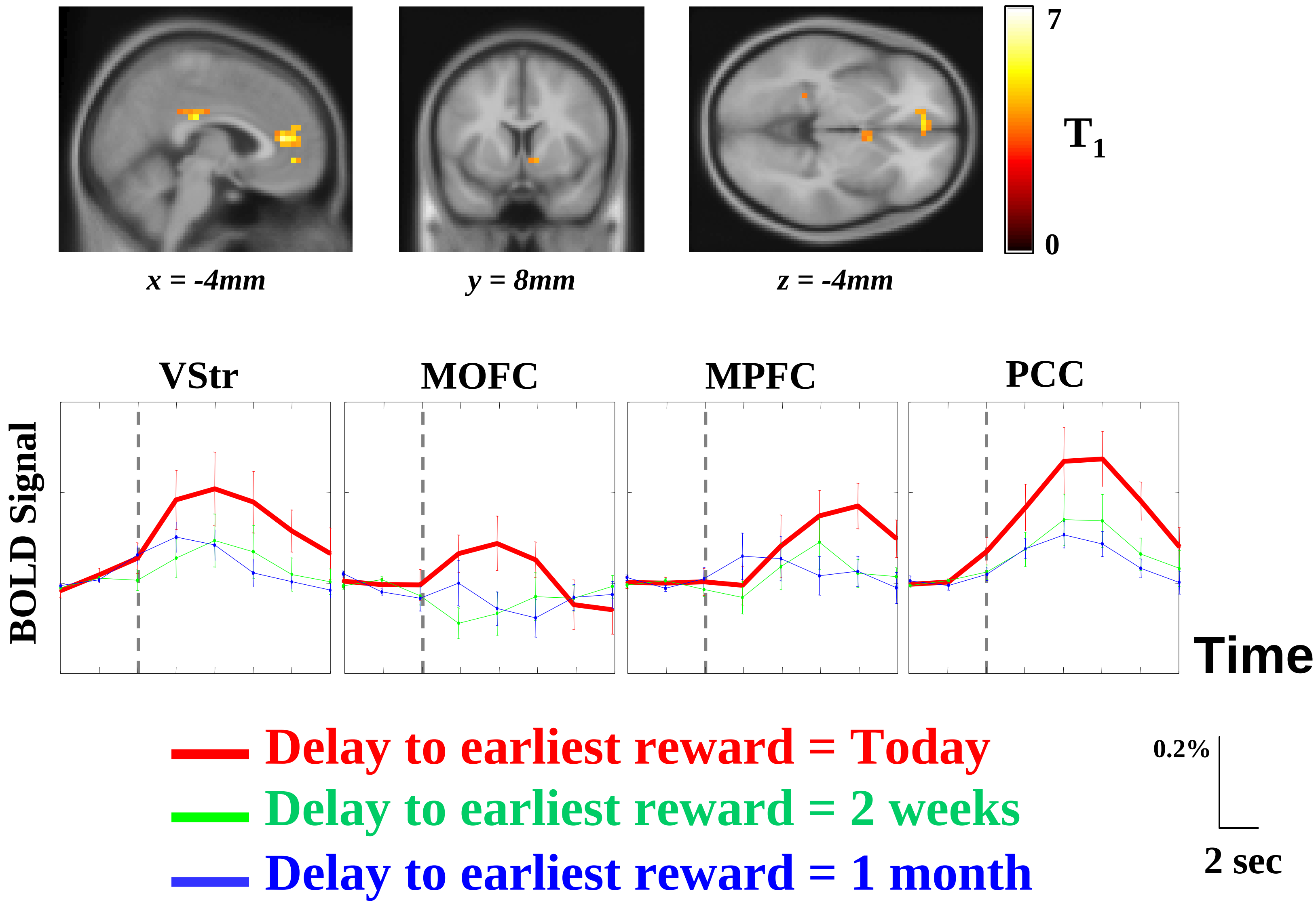
McClure, Laibson, Loewenstein, Cohen (2004)

- Intertemporal choice with time-dated Amazon gift certificates.
- Subjects make binary choices:

$20 now	**or**	**$30 in two weeks**
$20 in two weeks	**or**	**$30 in four weeks**
$20 in four weeks	**or**	**$30 in six weeks**


Meso-limbic dopamine
Fronto-parietal cortex
$20
$30
Fronto-parietal cortex
$20
$30
Fronto-parietal cortex
$20
$30

β areas respond “only” to immediate rewards

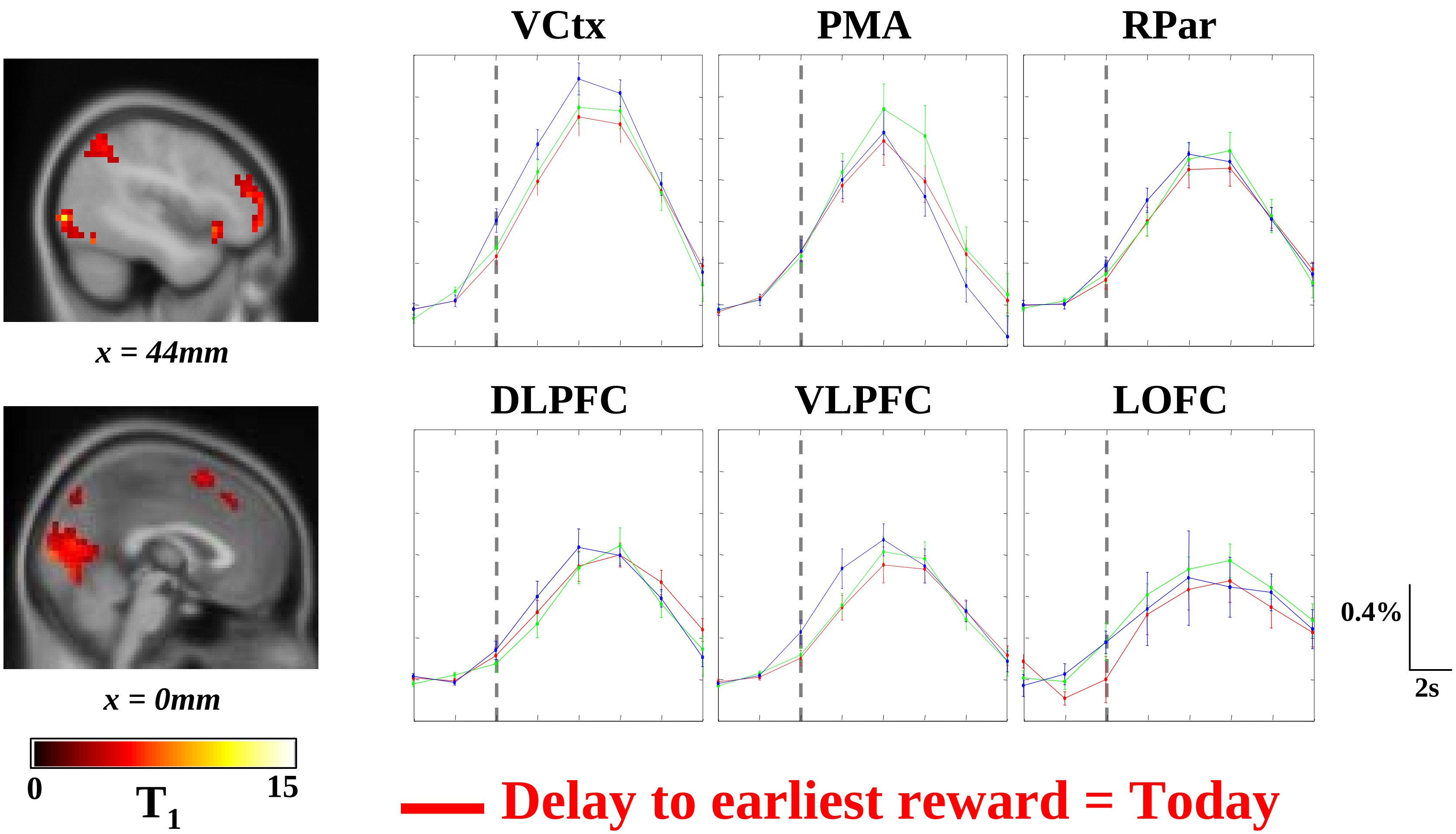


β Analysis
Summary of Significant Voxels

	X	Y	Z	Max T	n
Medial OFC	-8	48	-4	5.034	16
Ventral striatum	6	8	-4	4.369	5
L Posterior hippocampus	-26	-38	-8	4.582	7
Medial PFC	0	44	12	6.793	74
Posterior cingulate cortex	-8	-28	32	5.354	21

All voxels significant at $p < 0.001$

δ Areas respond equally to all rewards



Delay to earliest reward = Today

Delay to earliest reward = 2 weeks

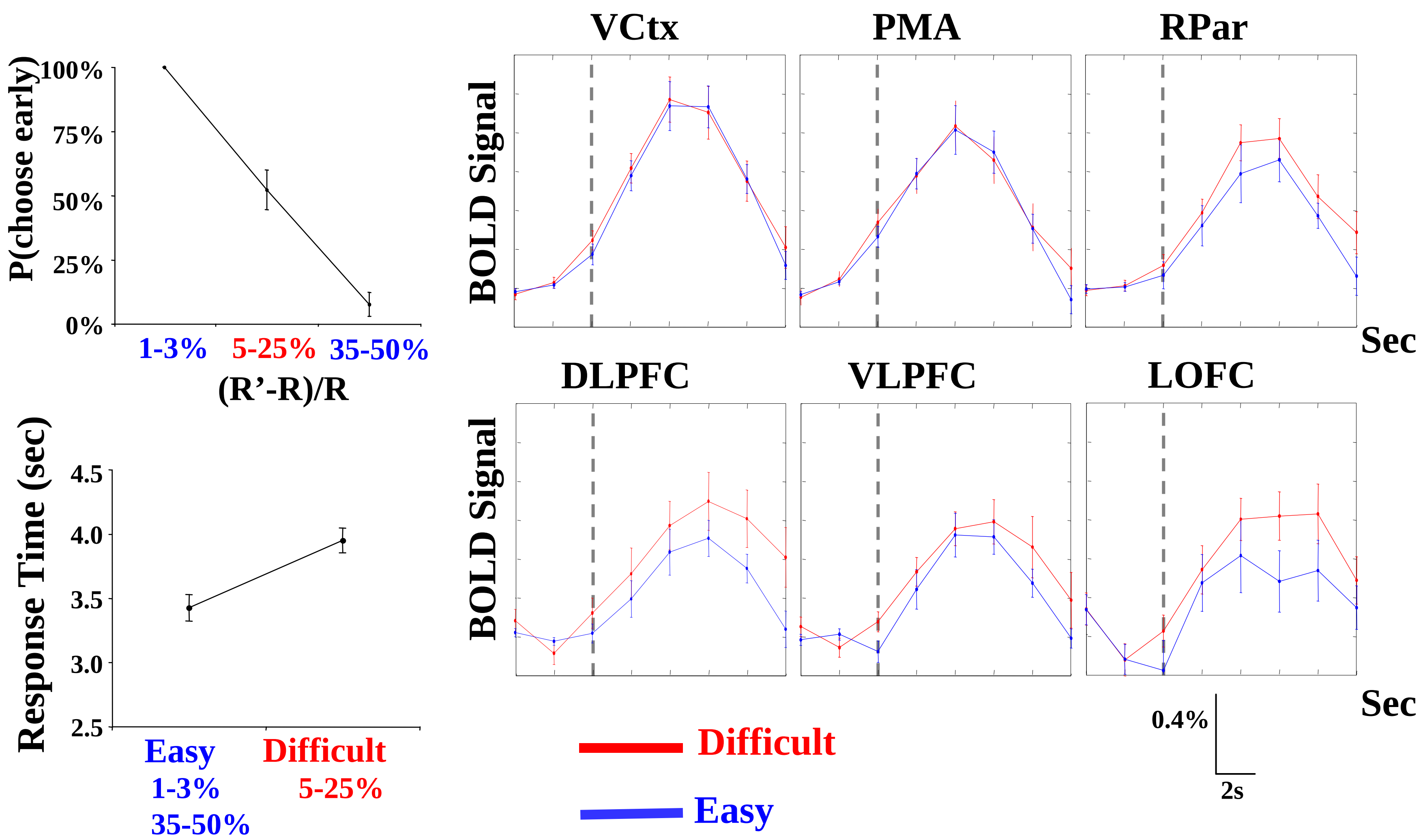
Delay to earliest reward = 1 month

δ Analysis:
Summary of Significant Voxels

	X	Y	Z	Max T	n
Visual cortex	-4	-80	4	14.787	1093
PMA	0	12	56	5.277	15
SMA	4	30	40	4.987	10
R Posterior parietal cortex	40	-60	44	7.364	191
L Posterior parietal cortex	-32	-60	52	9.568	73
R DLPFC	44	44	16	7.422	59
R VLPFC	40	20	-8	7.596	39
R Lateral OFC	24	50	-12	5.509	5

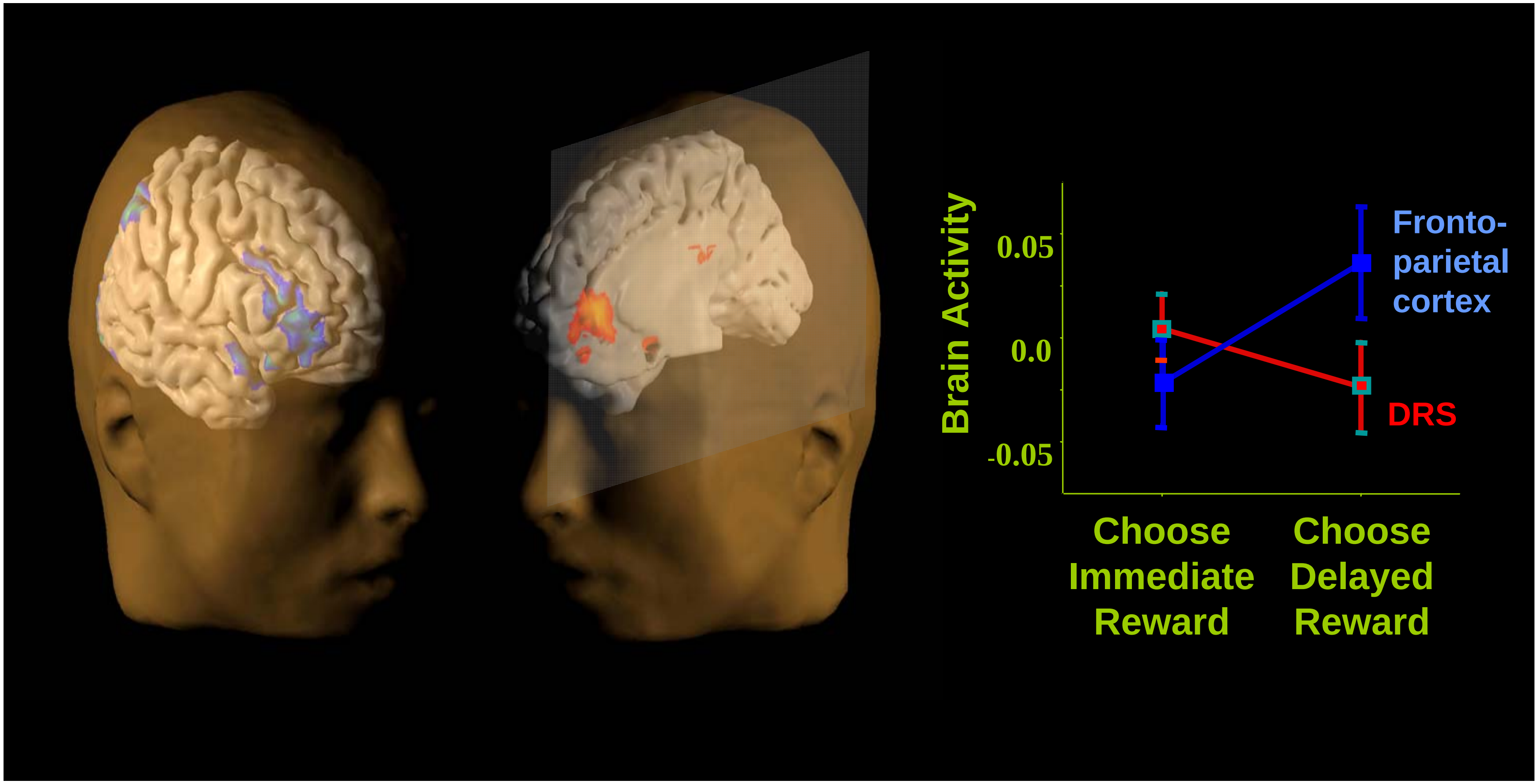
All voxels significant at $p < 0.001$

Effect of Difficulty



Brain activity in the frontoparietal system and mesolimbic dopamine reward system predict behavior

(Data for choices with an immediate option.)



McClure, Ericson, Laibson, Loewenstein, Cohen (2007)

Subjects water deprived for 3hr prior to experiment

From:
Subject: **I hate you**
Date: December 14, 2004 3:57:34 PM EST
To: dardenne@Princeton.EDU
Cc: smcclure@Princeton.EDU

I'm already thirsty! It's 4:00!

(a subject scheduled for 6:00)

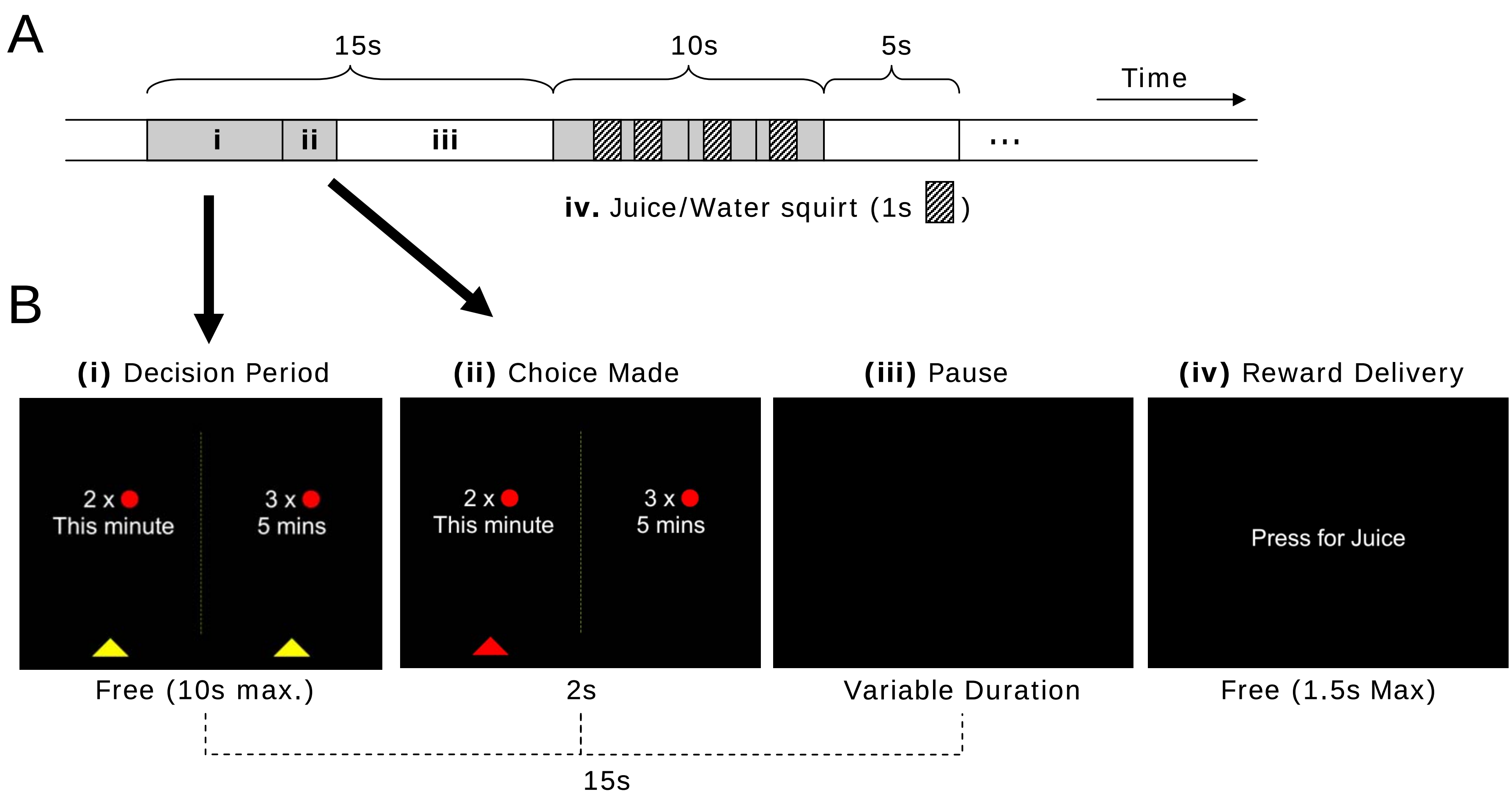
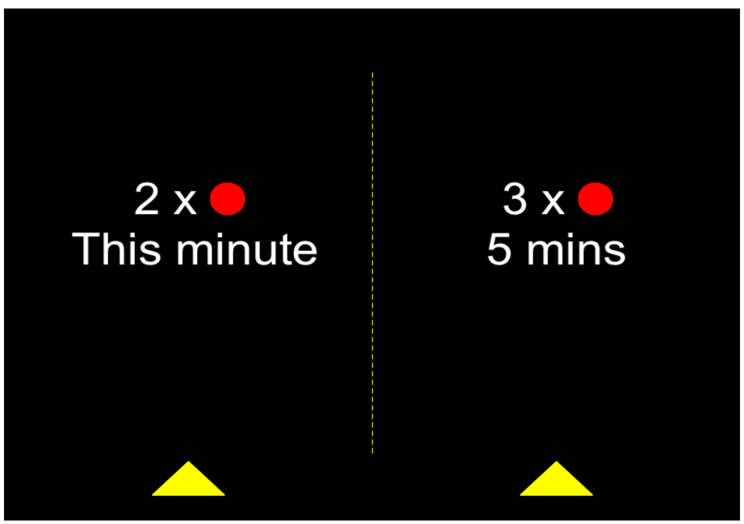

A
15s
10s
5s
Time
i
ii
iii
...
iv. Juice/Water squirt (1s)
B
(i) Decision Period
(ii) Choice Made
(iii) Pause
(iv) Reward Delivery
2 x
This minute
3 x
5 mins
2 x
This minute
3 x
5 mins
Press for Juice
Free (10s max.)
2s
Variable Duration
Free (1.5s Max)
15s

Figure 1

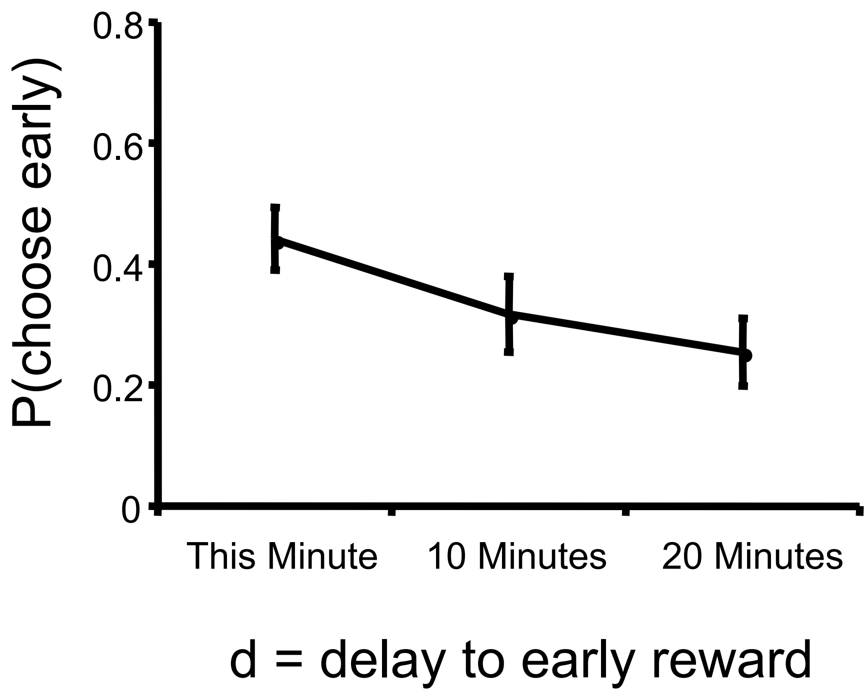
Experiment Design

d ∈ { This minute, 10 minutes, 20 minutes }
d'-d ∈ { 1 minute, 5 minutes }
(R, R') ∈ {(1,2), (1,3), (2,3)}

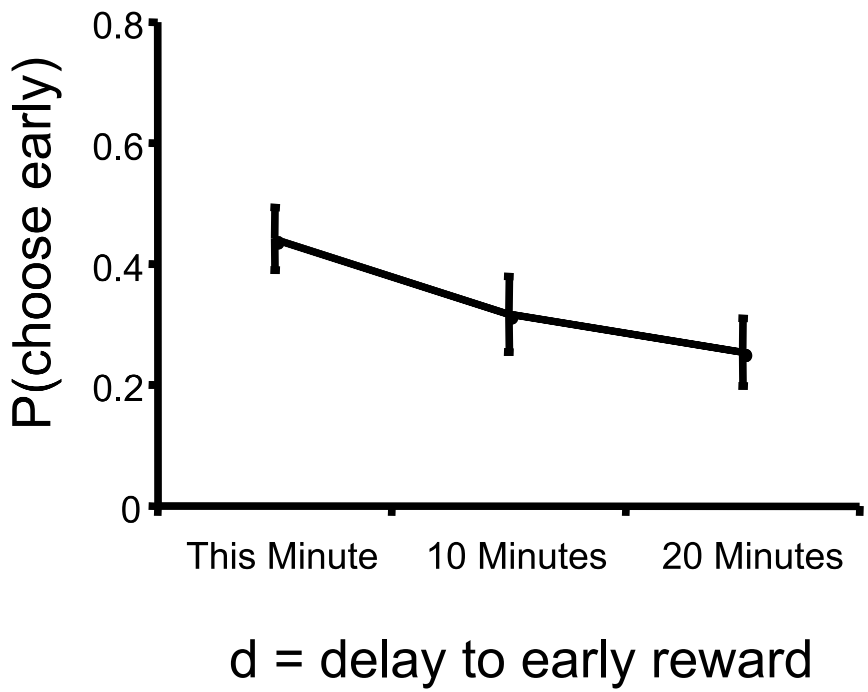


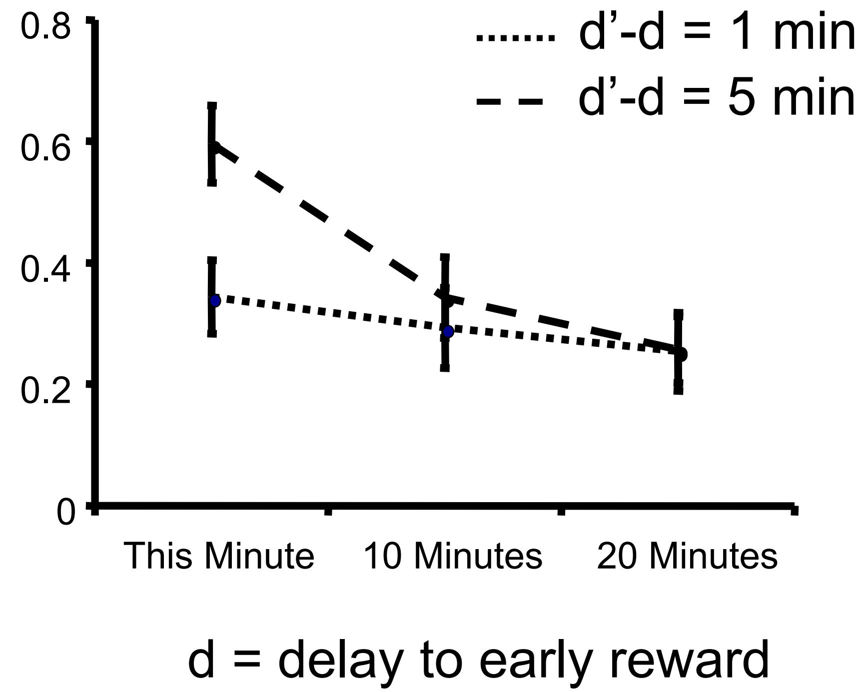
d = This minute
d'-d = 5 minutes
(R, R') = (2,3)

Behavioral evidence



Behavioral evidence





Discount functions fit to behavioral data

$$V(t) = u(c_t) + \beta \sum_{\tau=1}^{\infty} \delta^{\tau} u(c_{t+\tau})$$

$$V(t) = \left(\frac{1}{\beta} - 1\right) u(c_t) + \sum_{\tau=0}^{\infty} \delta^{\tau} u(c_{t+\tau})$$

Mesolimbic | Cortical

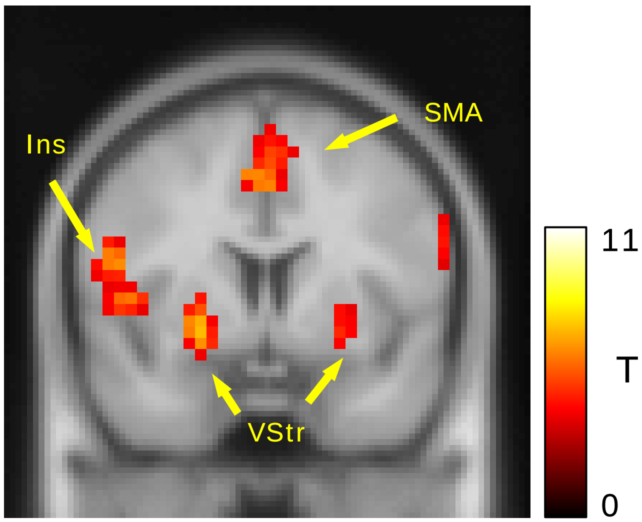
β = 0.53 (se = 0.041)
δ = 0.98 (se = 0.014)

$$V(t) = \omega \sum_{\tau=0}^{\infty} \beta^{\tau} u(c_{t+\tau}) + (1-\omega) \sum_{\tau=0}^{\infty} \delta^{\tau} u(c_{t+\tau})$$

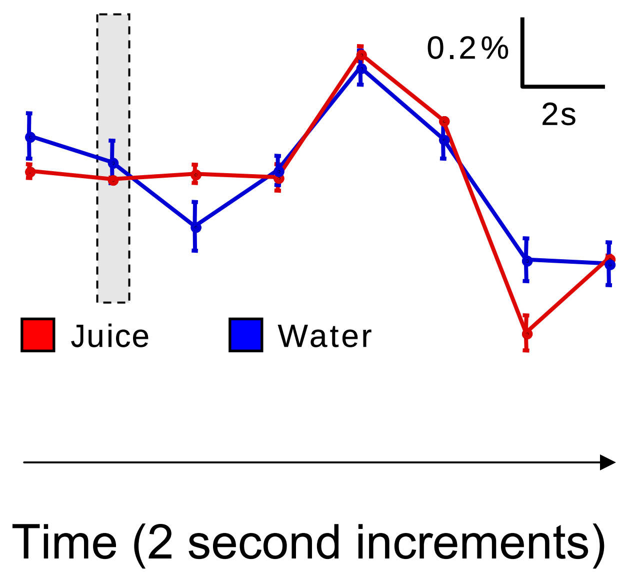
β= 0.47 (se = 0.101)
δ = 1.02 (se = 0.018)

- Evidence for two-system model
- Can reject exponential restriction with t-stat > 5
- Double exponential generalization fits data best

Juice and Water treated equally
(both behavioral and neurally)

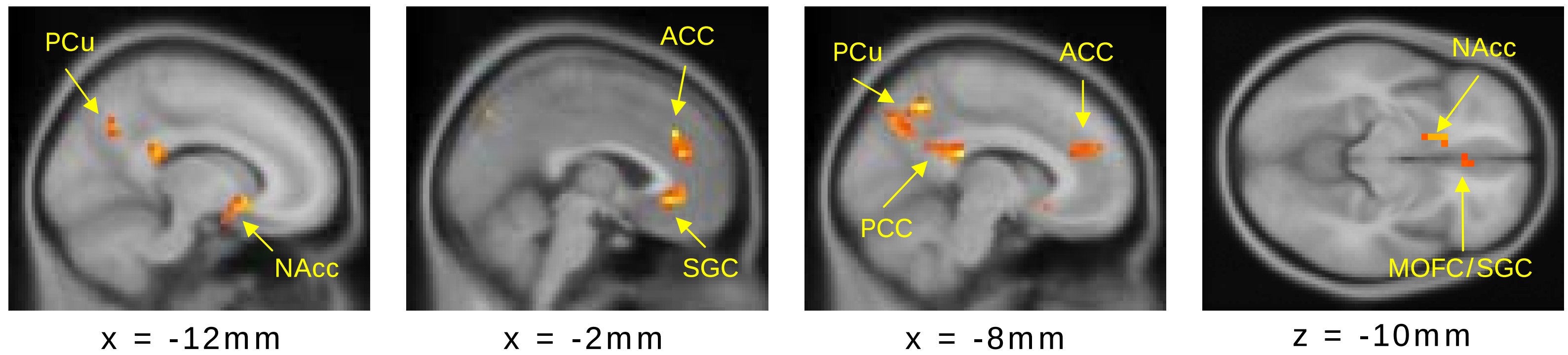


y = 12mm



Neuroimaging data estimated with general linear model.

A β areas: respond only to immediate rewards



B δ areas: respond to all rewards

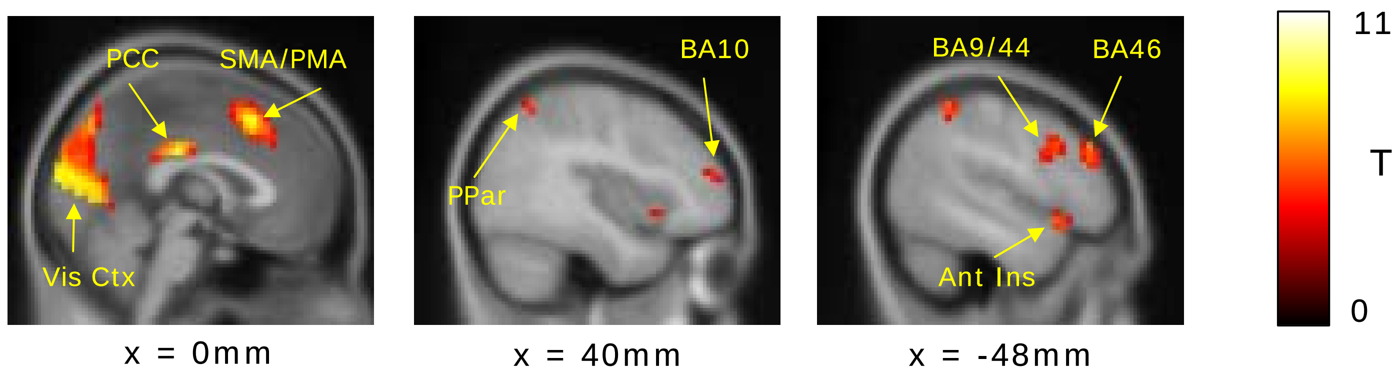
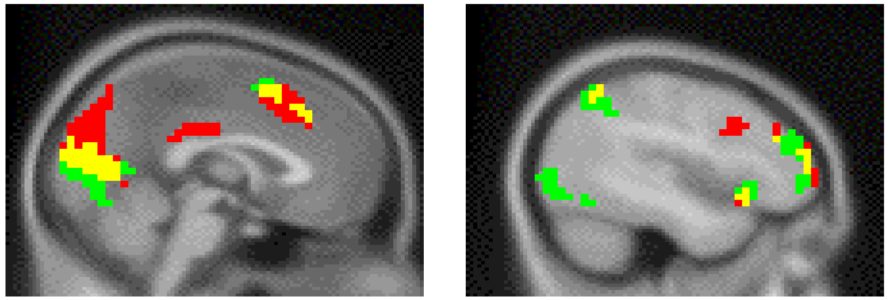


Figure 4

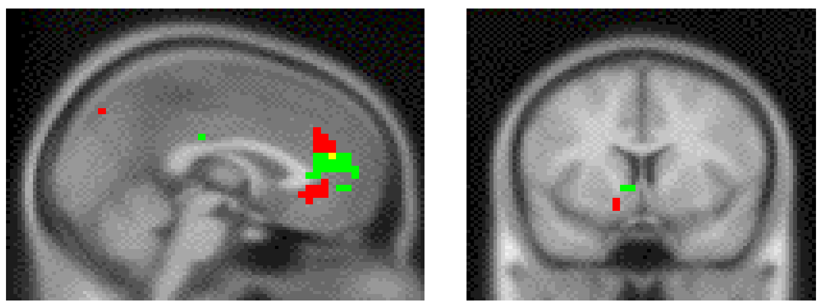
Relationship to Amazon experiment:

δ areas ($p<0.001$)



x = 0mm x = -48mm

β areas ($p<0.001$)



x = 0mm y = 8mm

Juice only Amazon only Both

Figure 5

Relationship to Amazon experiment:

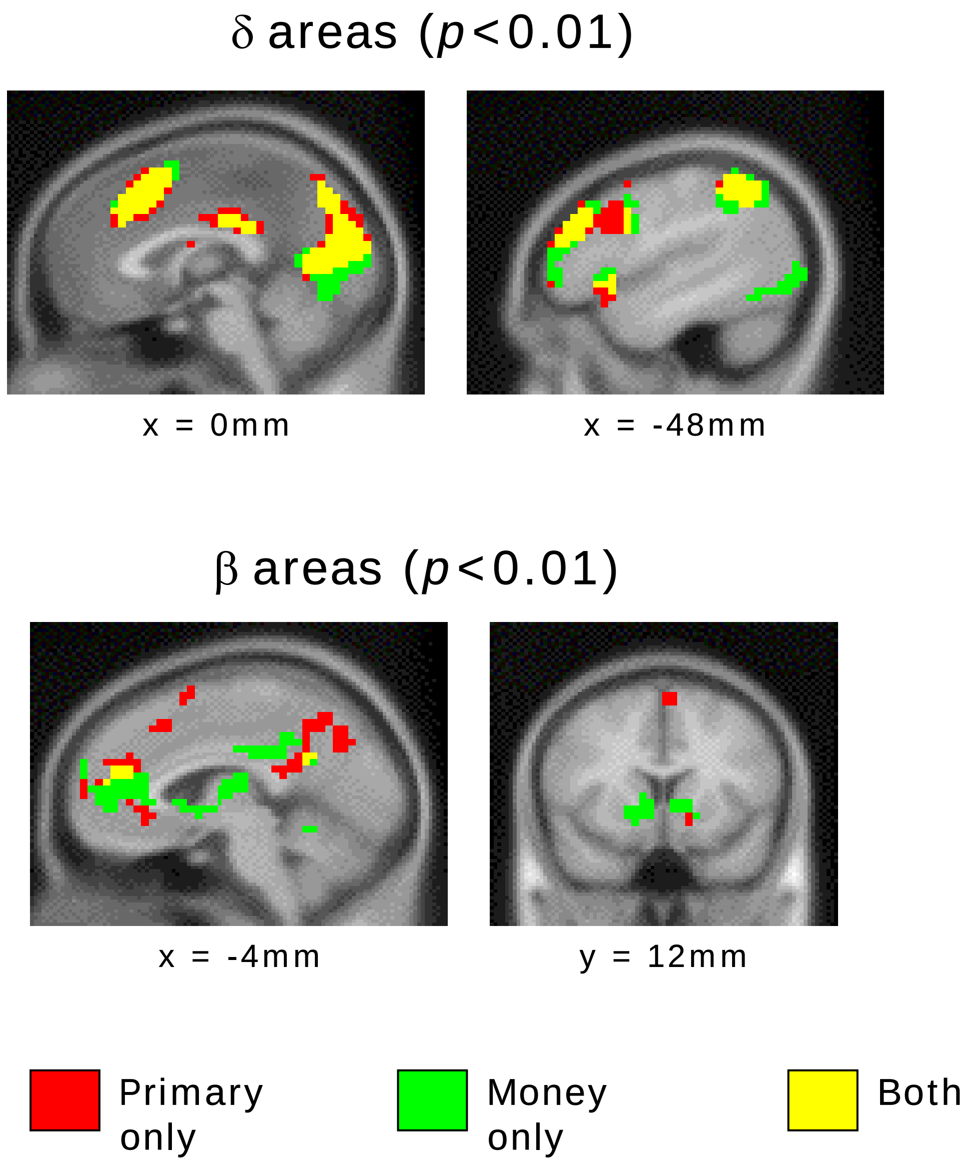


Figure 5

Fitting discount functions using neuroimaging data.

β systems

$$N(t) = \beta^{d}R + \beta^{d'}R' + X(t)\cdot\theta + \varepsilon(t)$$

$$\beta = 0.963\ (0.004)$$

$$0.963^{25} = 0.39$$

	Discount factor	*StdErr*
ACC	**0.1099**	***(0.132)***
MPFC	**0.0000**	***NA***
NAc	**0.9592**	***(0.014)***
PCC	**0.9437**	***(0.014)***
PCu	**0.9547**	***(0.011)***

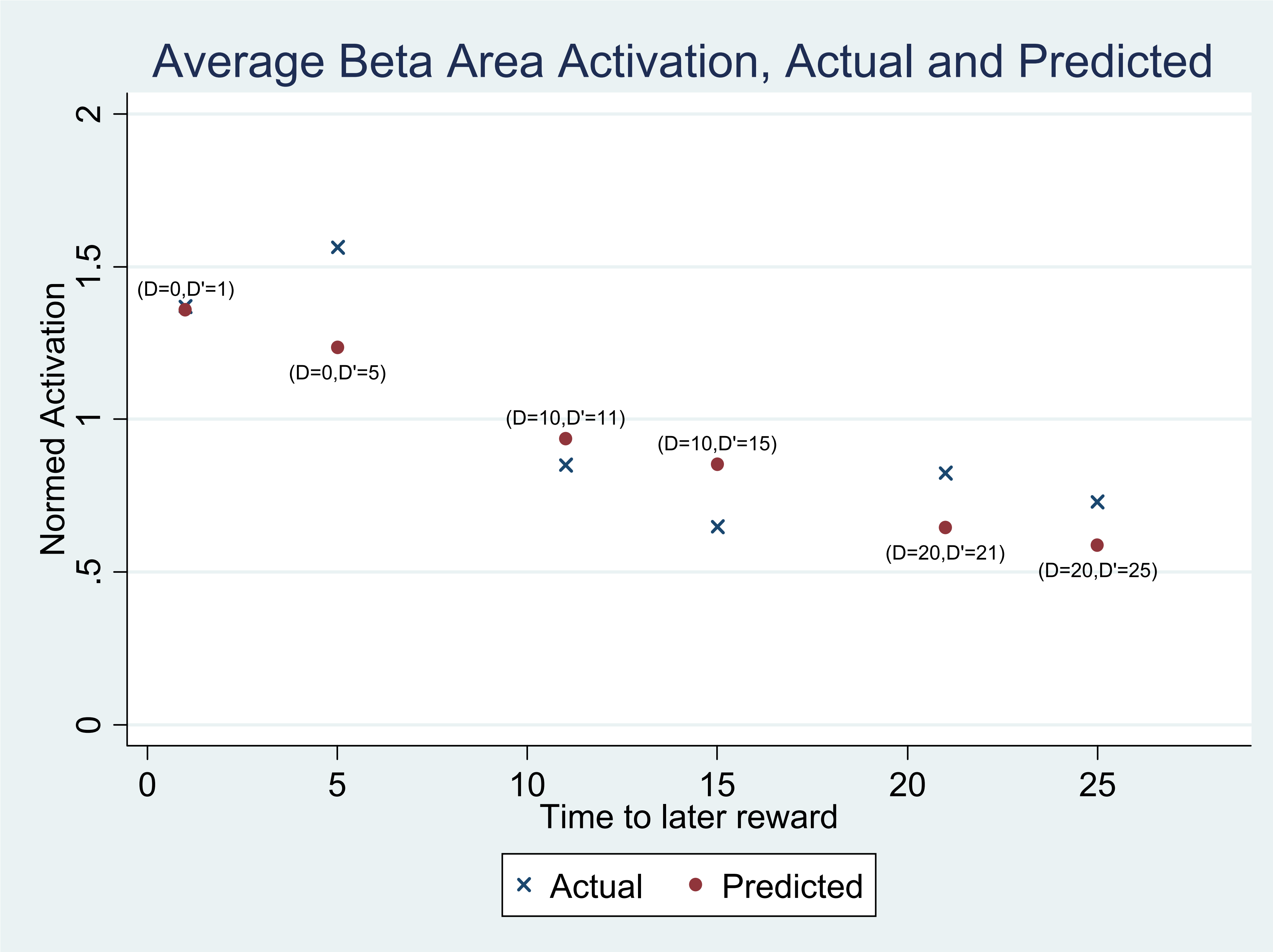
δ systems

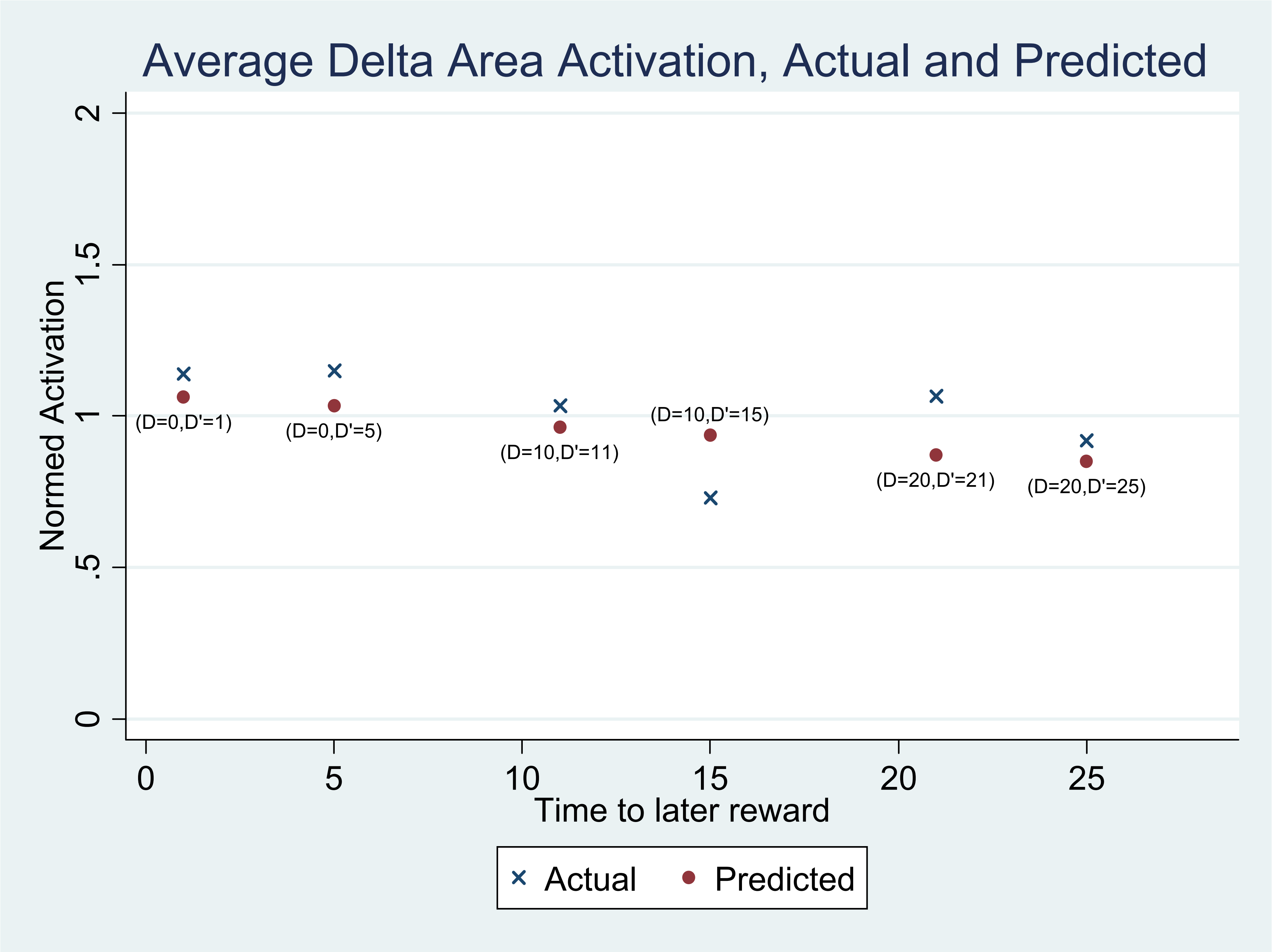
$$N(t) = \delta^{d}R + \delta^{d'}R' + X(t)\cdot\theta + \varepsilon(t)$$

$$\delta = 0.990\ (0.003)$$

$$0.990^{25} = 0.78$$

AntInsL	**0.9908**	***(0.007)***
AntInsR	**0.9871**	***(0.006)***
BA10L	**1.0047**	***(0.011)***
BA10R	**0.9953**	***(0.010)***
BA46	**0.9870**	***(0.008)***
BA944	**0.9913**	***(0.006)***
PCC	**0.9926**	***(0.006)***
PParL	**0.9970**	***(0.005)***
PParR	**0.9959**	***(0.005)***
SMAPMA	**0.9870**	***(0.006)***
VisCtx	**0.9939**	***(0.005)***


Average Beta Area Activation, Actual and Predicted
Normed Activation
0
.5
1
1.5
2
0
5
10
15
20
25
Time to later reward
(D=0,D'=1)
(D=0,D'=5)
(D=10,D'=11)
(D=10,D'=15)
(D=20,D'=21)
(D=20,D'=25)
Actual
Predicted



Average Delta Area Activation, Actual and Predicted
Normed Activation
0
.5
1
1.5
2
0
5
10
15
20
25
Time to later reward
(D=0,D'=1)
(D=0,D'=5)
(D=10,D'=11)
(D=10,D'=15)
(D=20,D'=21)
(D=20,D'=25)
Actual
Predicted

What determines immediacy?

Is mesolimbic DRS activation associated with relatively “early” (or earliest) options?

or

Do juice and money have different discount functions?

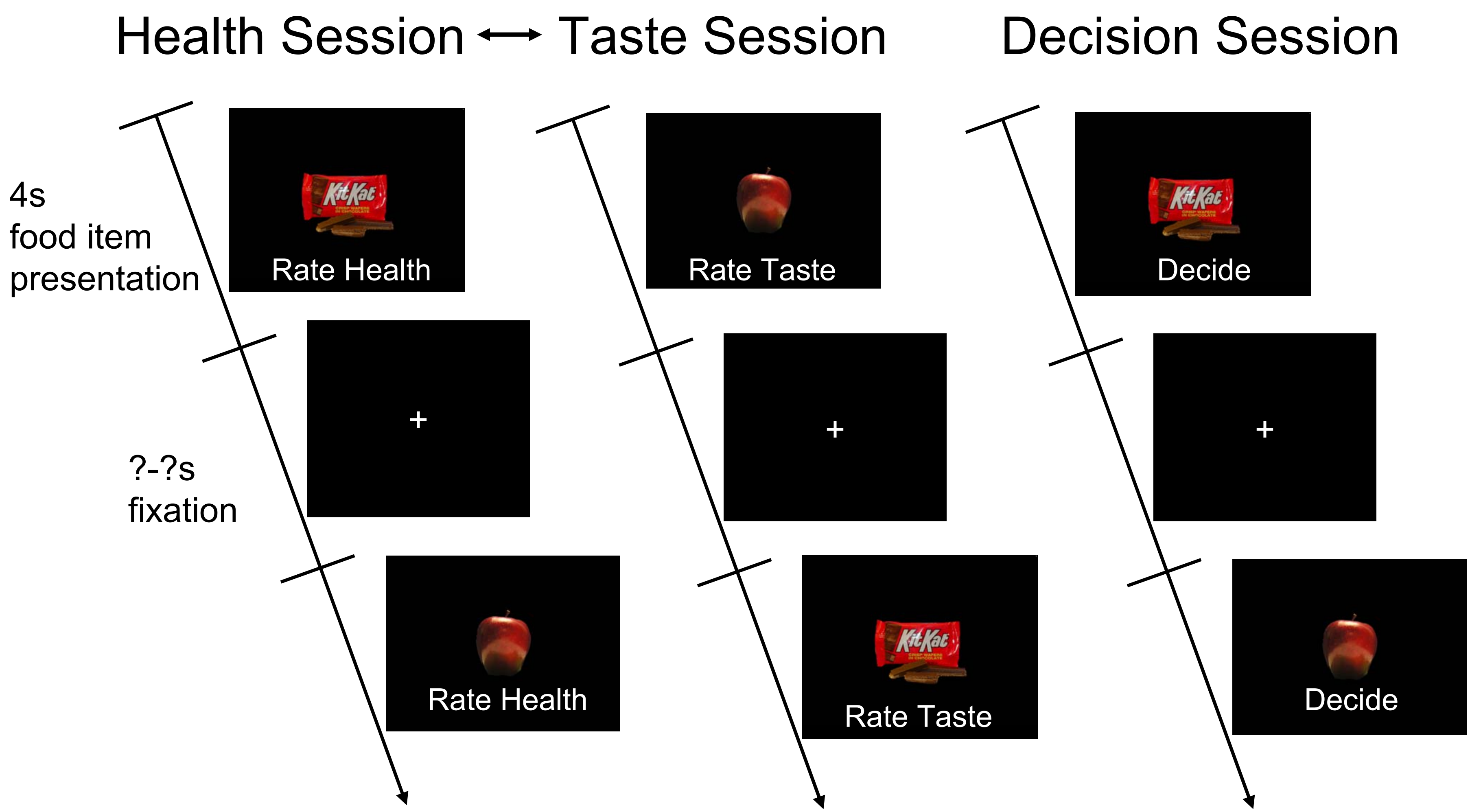
or

Does thirst invoke more intense discounting?

Our working hypotheses.

- One system associated with midbrain dopamine neurons (mesolimbic DRS) shows high sensitivity to time delay.
- Second system associated with lateral prefrontal and posterior parietal cortex shows less sensitivity to time delay.
- Combined function of these two systems explains decision making across choice domains.

Hare, Camerer, and Rangel (2009)



Details

- Taste and health ratings made on five point scale:

  -2,-1,0,1,2

- Decisions also reported on a five point scale: SN,N,0,Y,SY

  "strong no" to "strong yes"

- Subject choices sometimes reflect self control – rejection of an unhealthy, good tasting food, OR, acceptance of a healthy, bad tasting food

Subjects

- SC (self-control) group = 19 dieting subjects who showed self-control during the decision phase
- NSC (no self-control) group = 18 comparison subjects who did not exhibit self-control during the decision phase

Who is classified as a self-controller: SC? (must meet all criteria below)

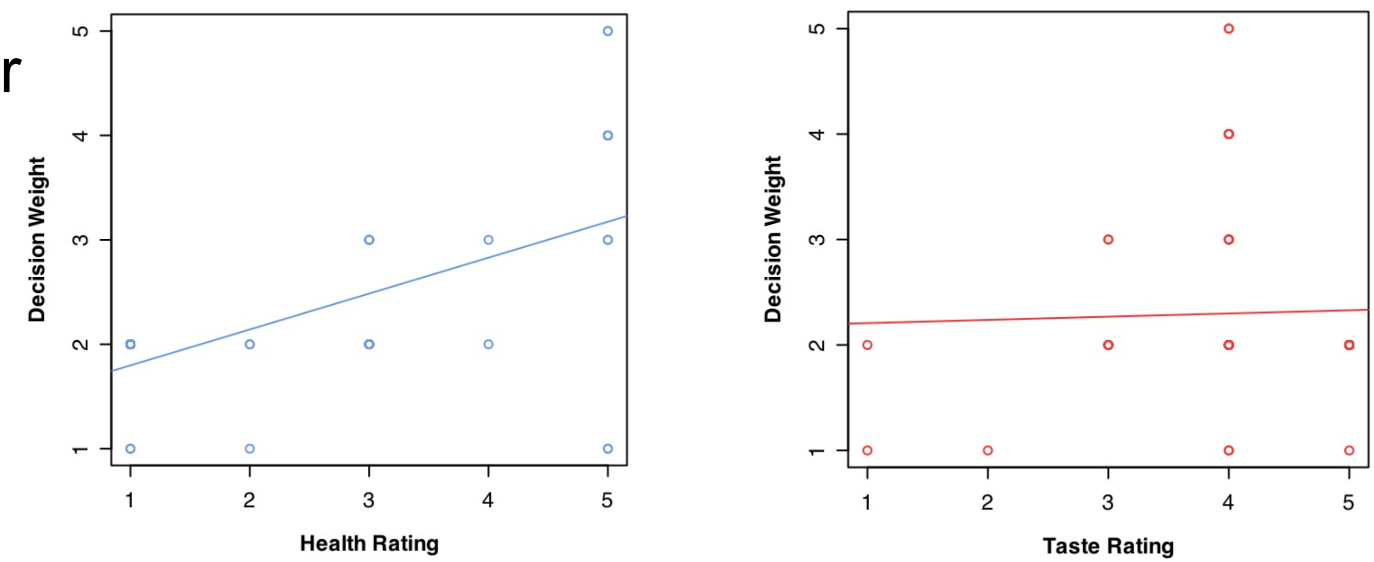
1) Use self-control on > %50 of trials in which self-control is required (decline Liked-Unhealthy items or choose Disliked-Healthy ones)

2) Decision = β_1HR + β_2LR + ε
   $\beta_1 > \beta_2$

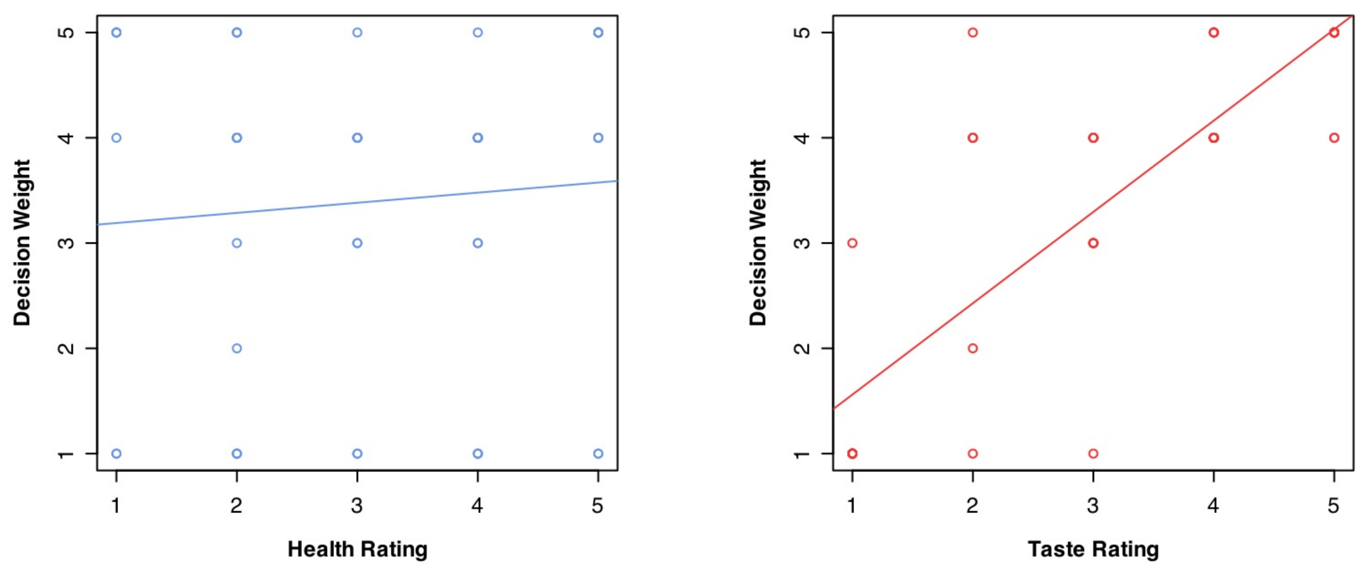
3) R^2 for HR > R^2 for LR

Examples of individual behavioral fits

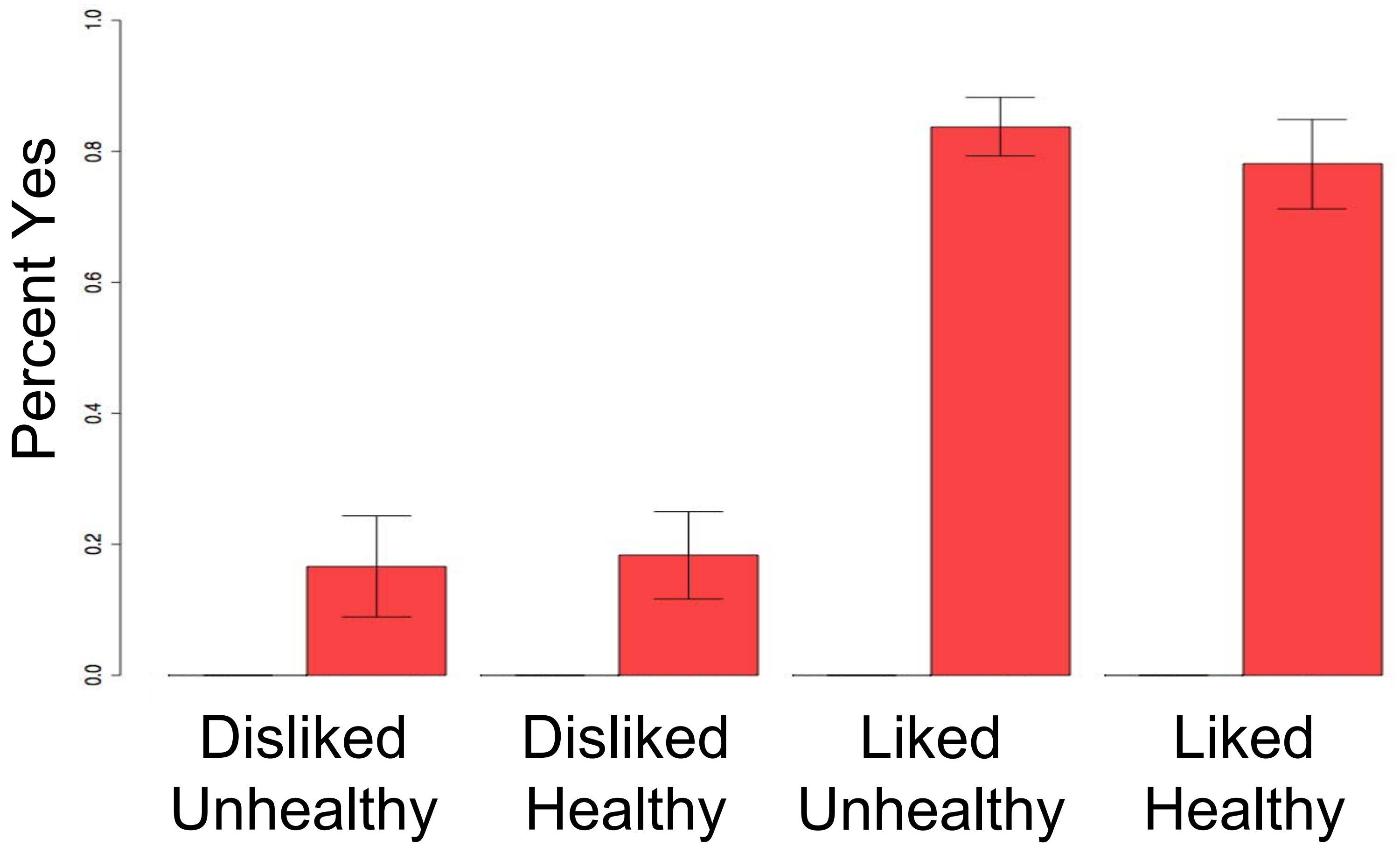
Self-controller



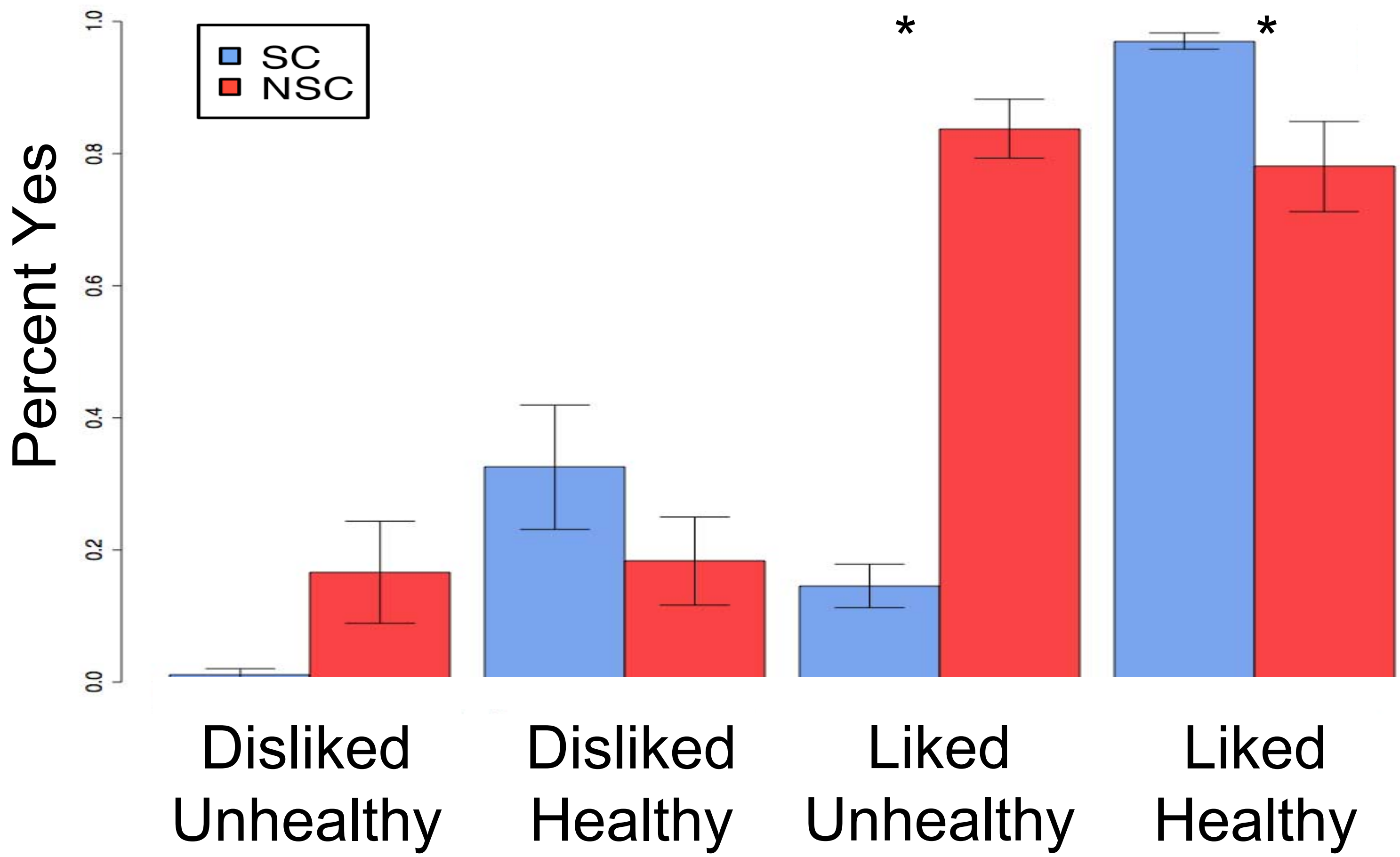
Non-
self-controller



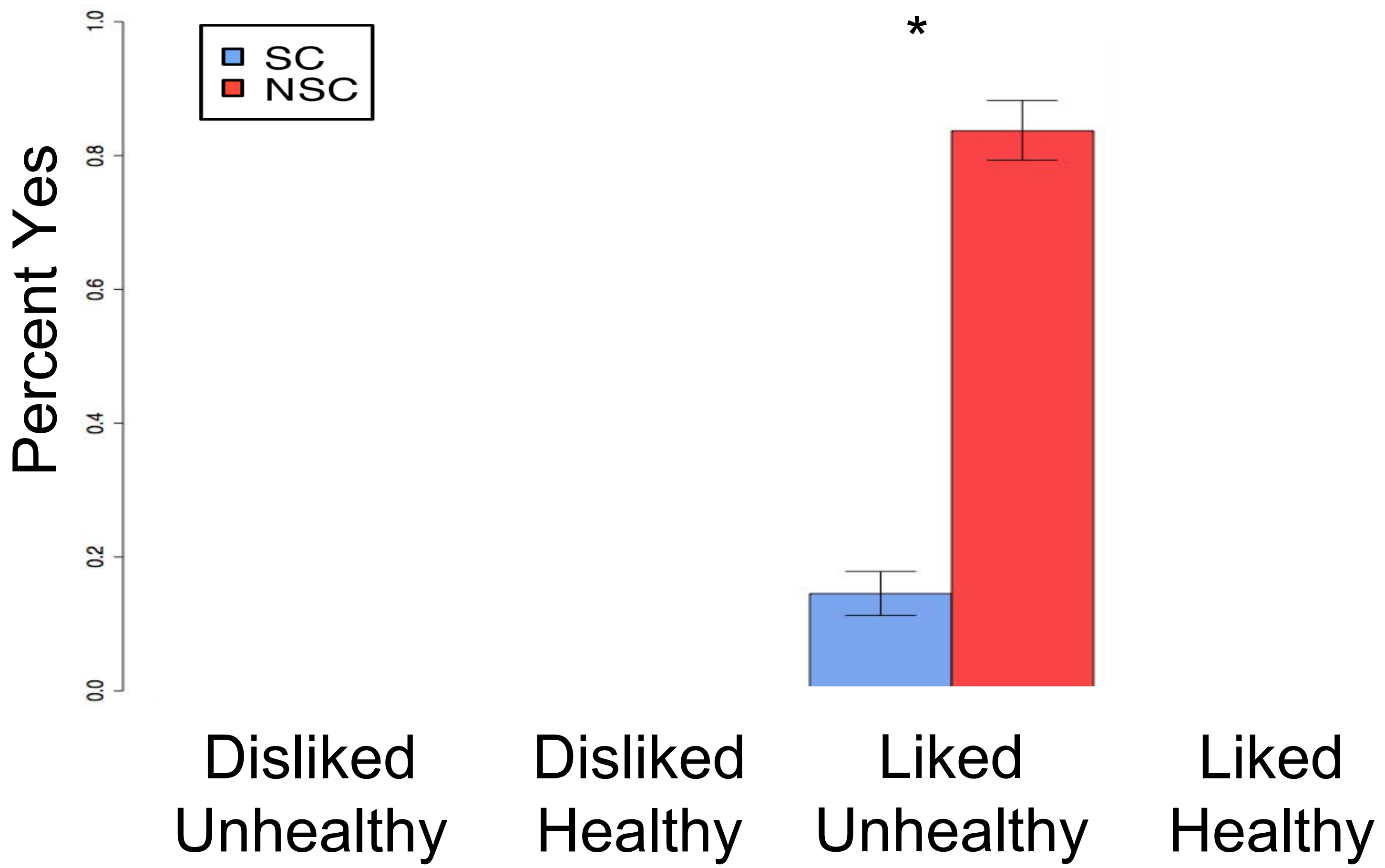
Result: NSC group chose based on taste



Result: SC group chose based on taste and health



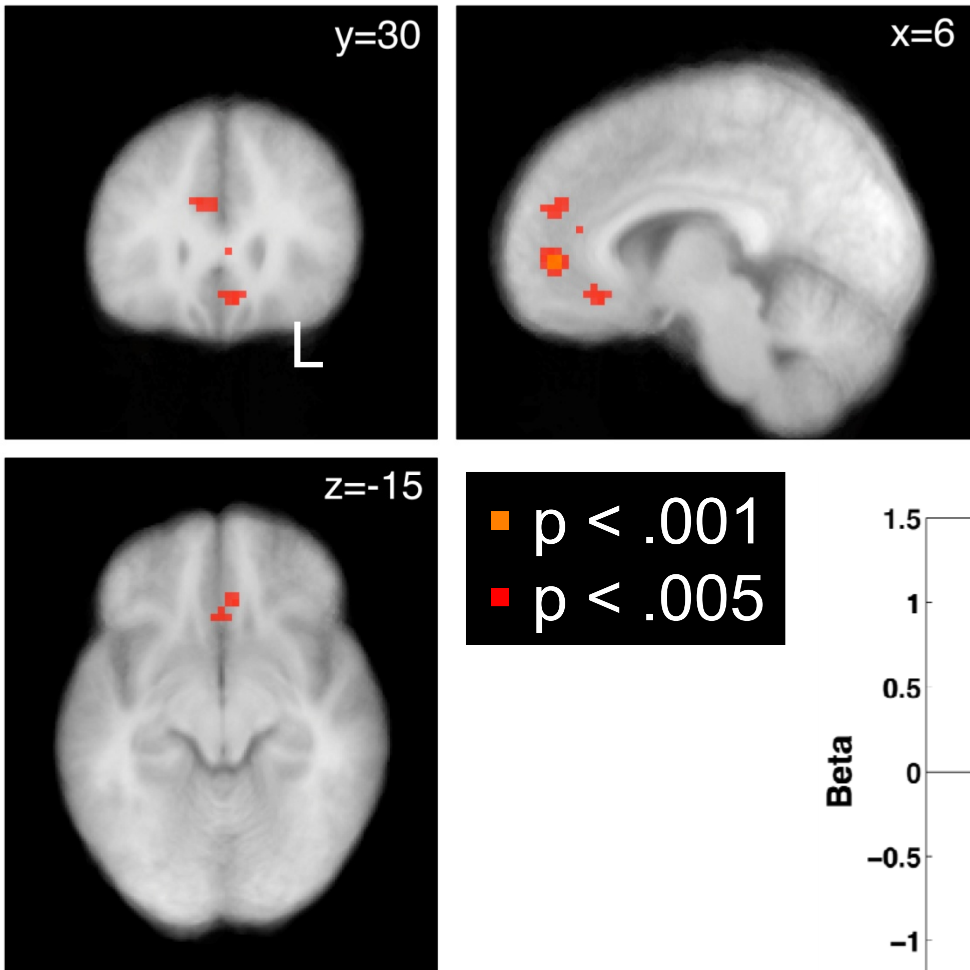
SC group versus NSC group

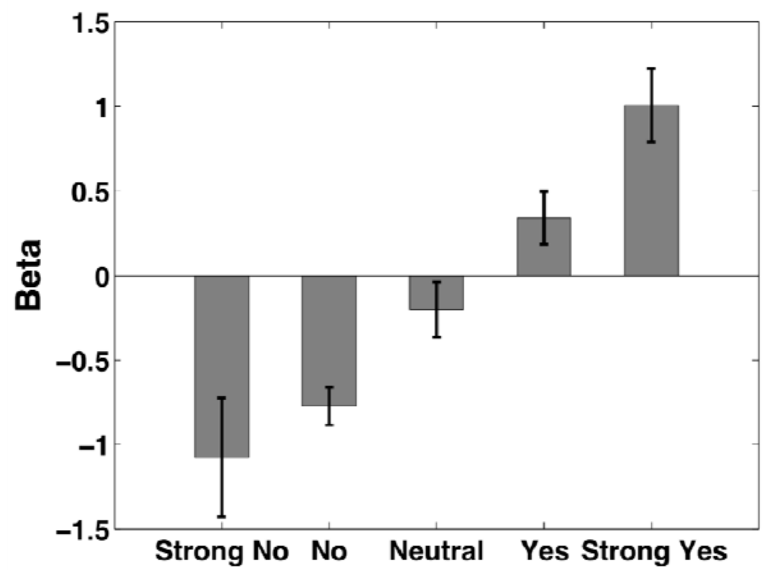


Neuroimaging Results

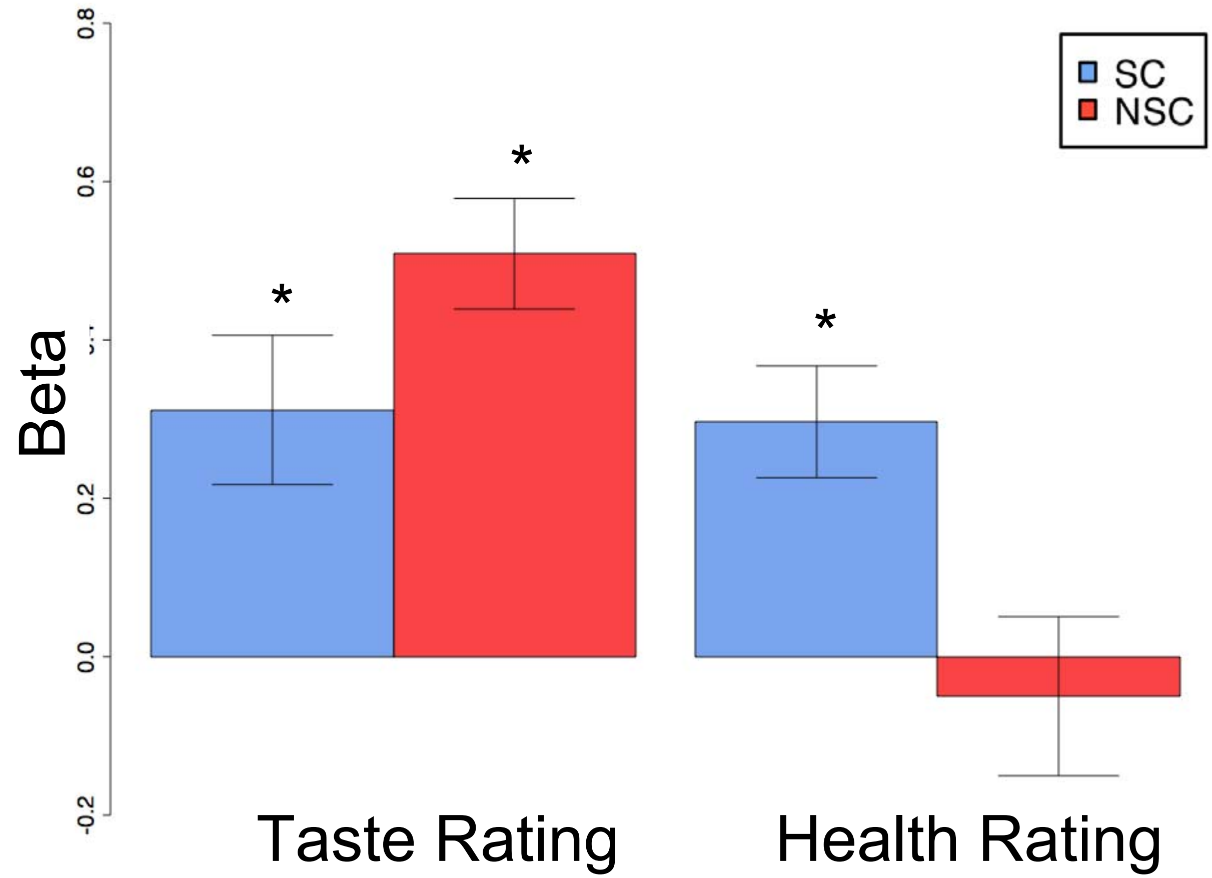
Question: Is there evidence for a single valuation system?

Activity in vmPFC is correlated with a behavioral measure of decision value (regardless of SC)

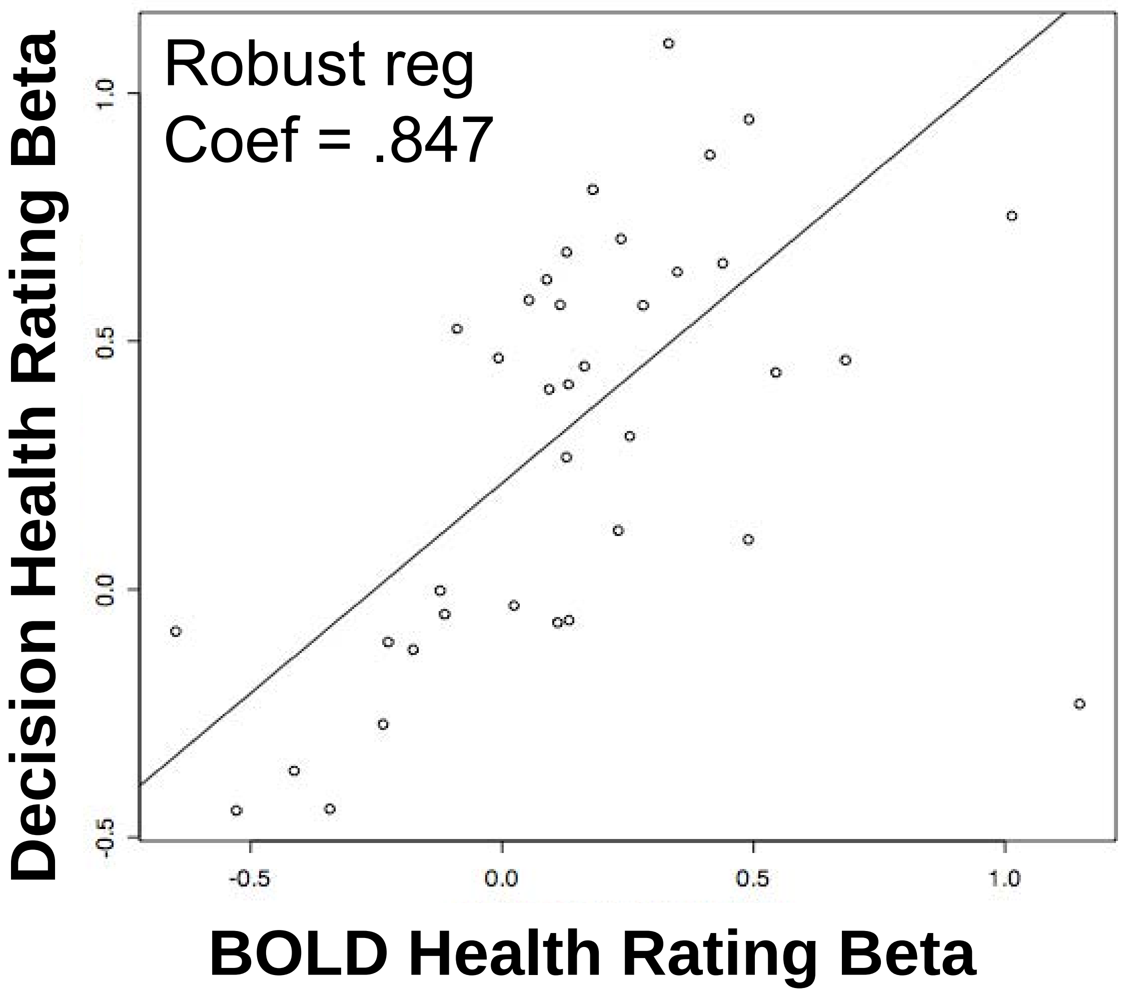




vmPFC BOLD signal reflects both taste and health ratings



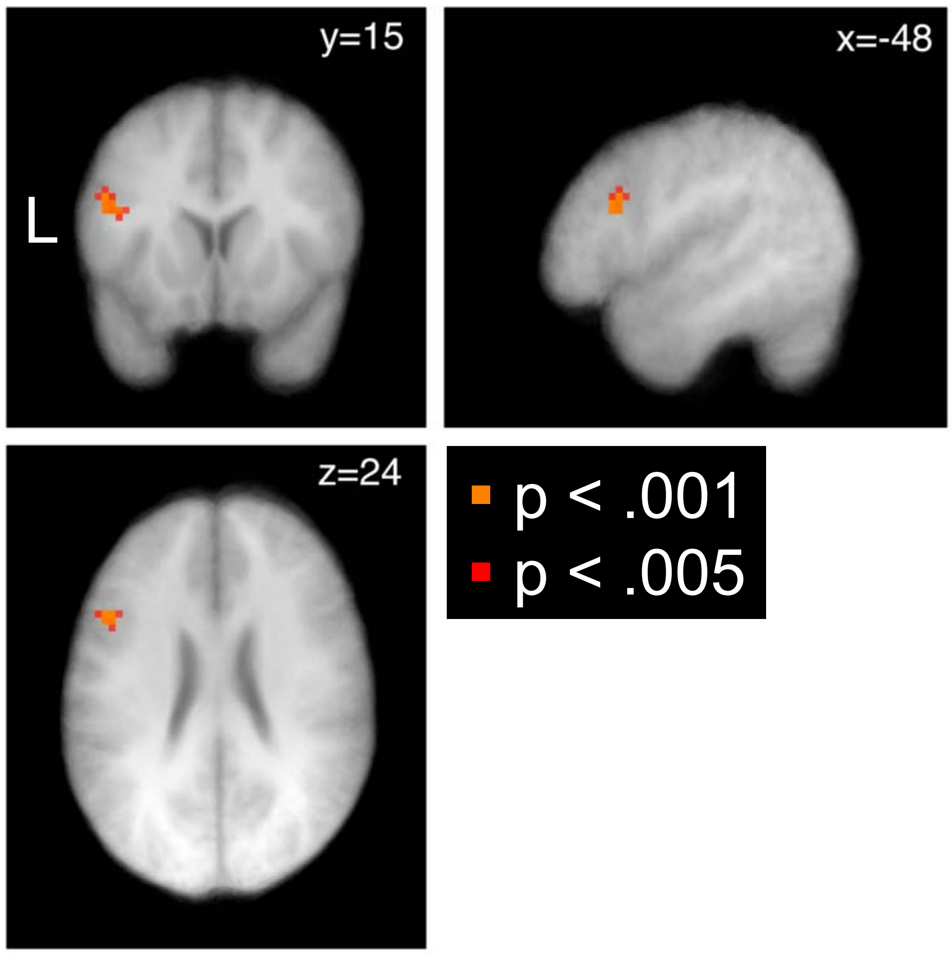
The effect of HR in the vmPFC is correlated with its effect on behavior



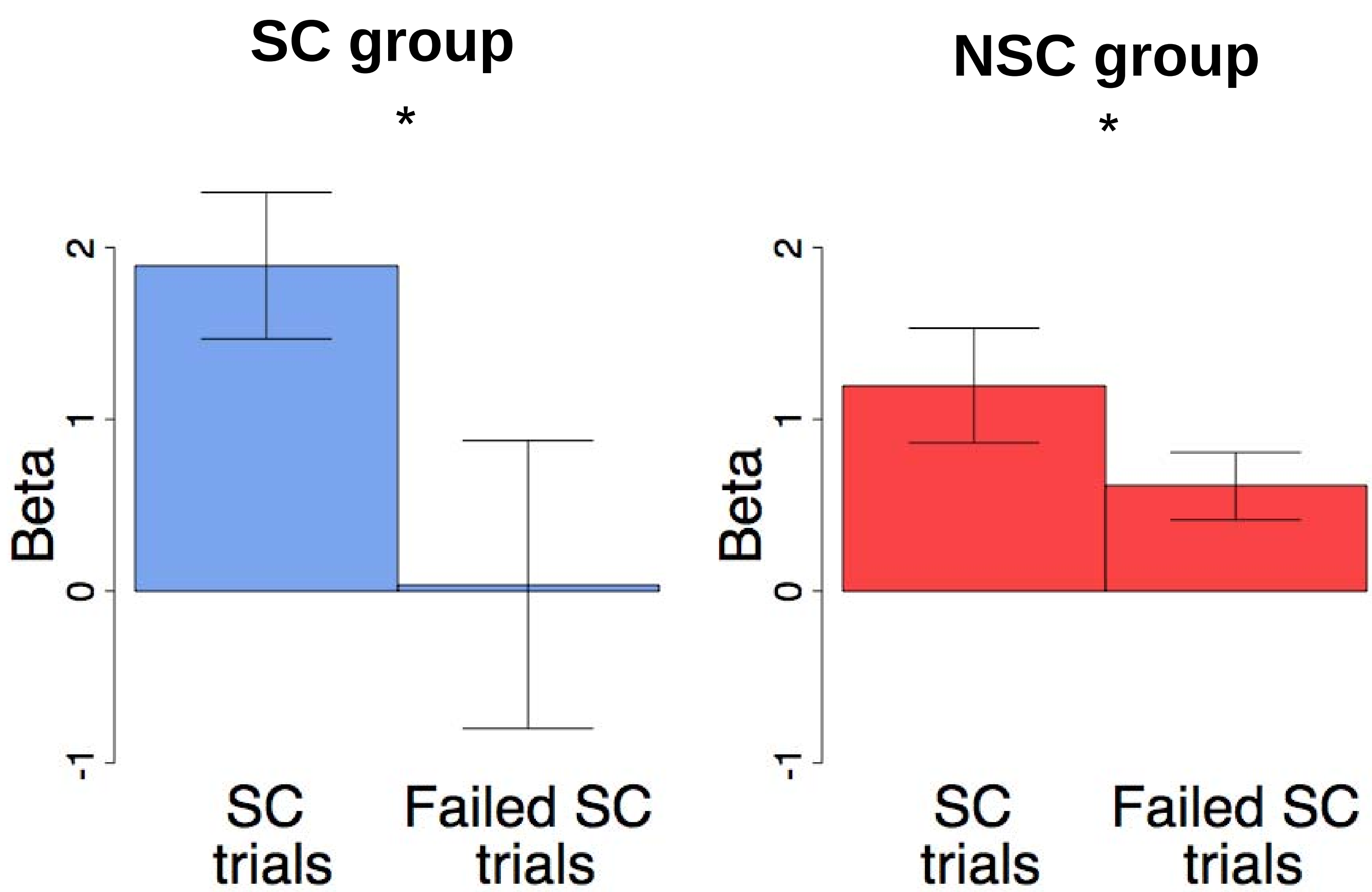
Neuroimaging Results

Question: Does self-control involve DLPFC modulation of the vmPFC valuation network?

More activity in DLPFC in successful SC trials than in failed SC trials



SC group has greater DLPFC than NSC when implementing self-control



Neuroimaging results

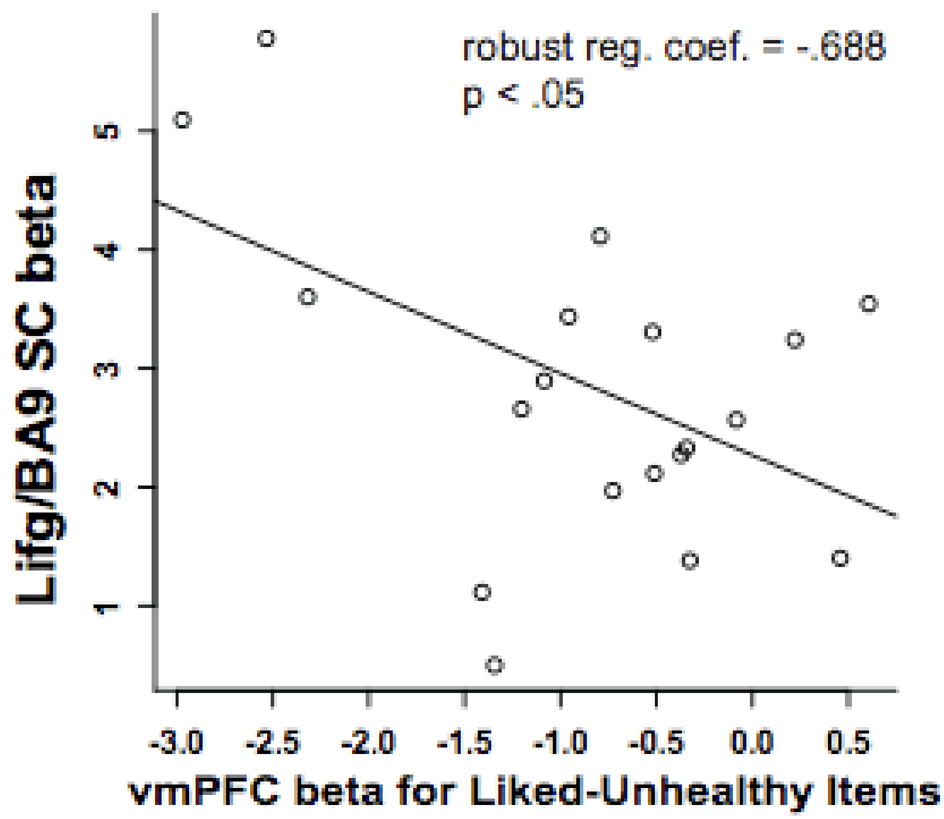
This result establishes that DLPFC activity increases during self-control.

Question: Is there more direct evidence that DLPFC modulates the vmPFC value networks

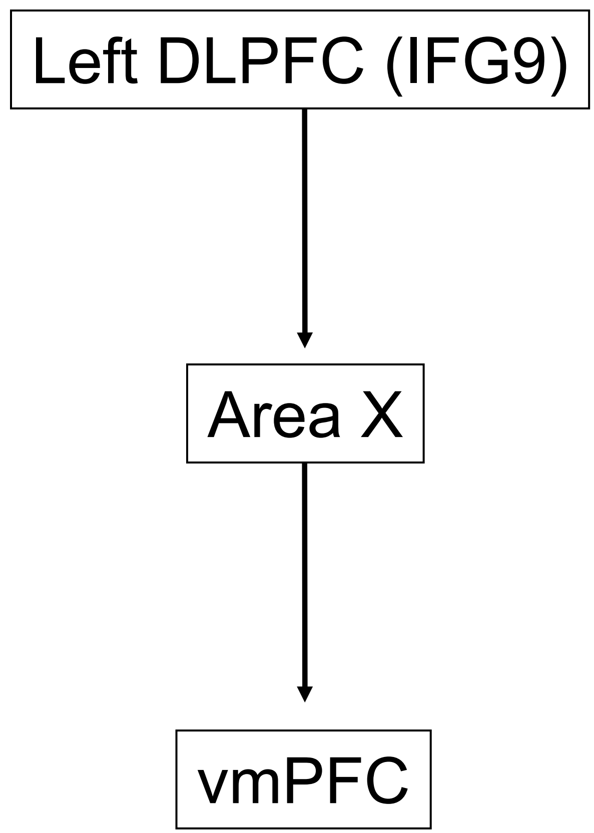
Functional Connectivity

- Initial psychophysiological interaction (PPI) results revealed no task related connectivity between left DLPFC and vmPFC value computation region

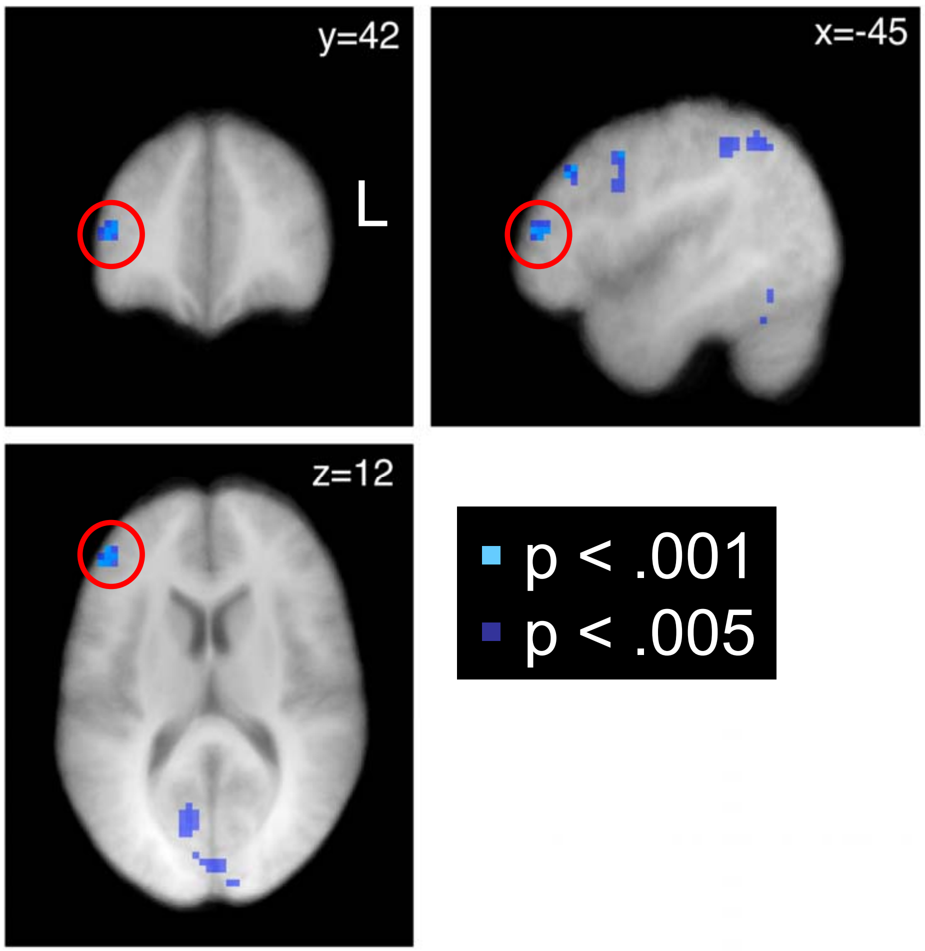
Some evidence for connectivity

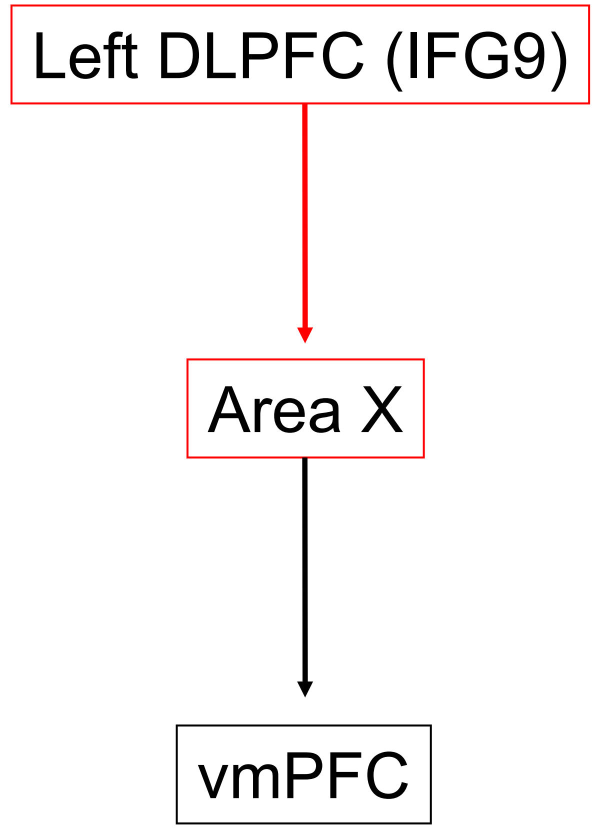


An indirect route?

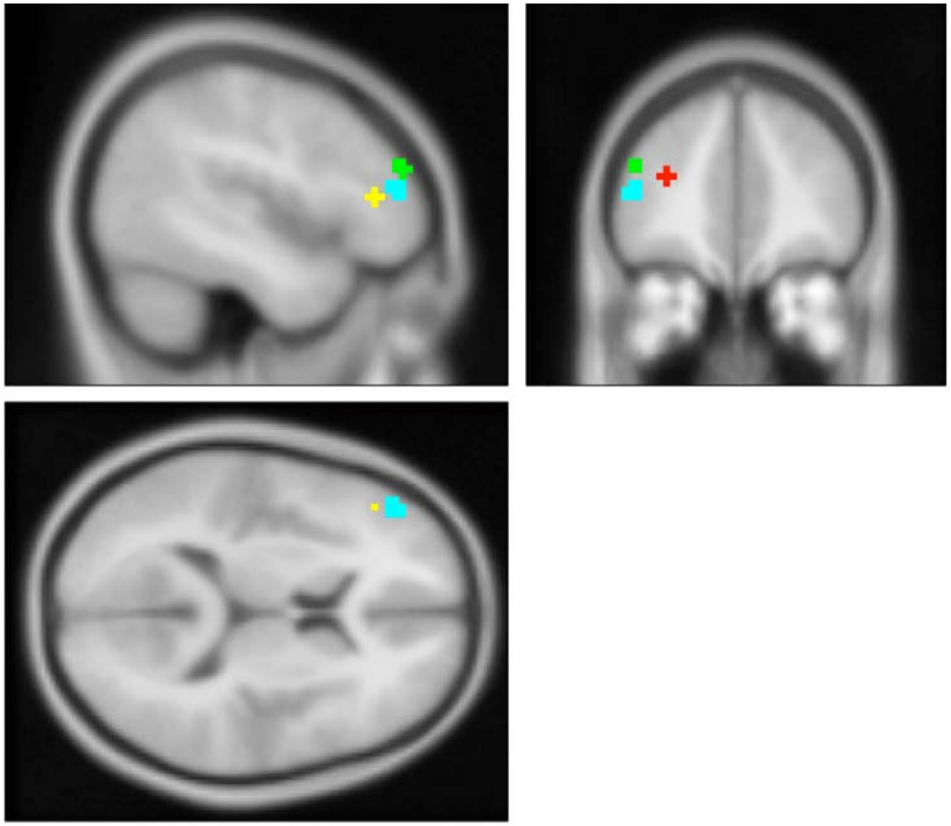


Node 1: Functional Connectivity with adjacent area of left DLPFC (negative correlation)

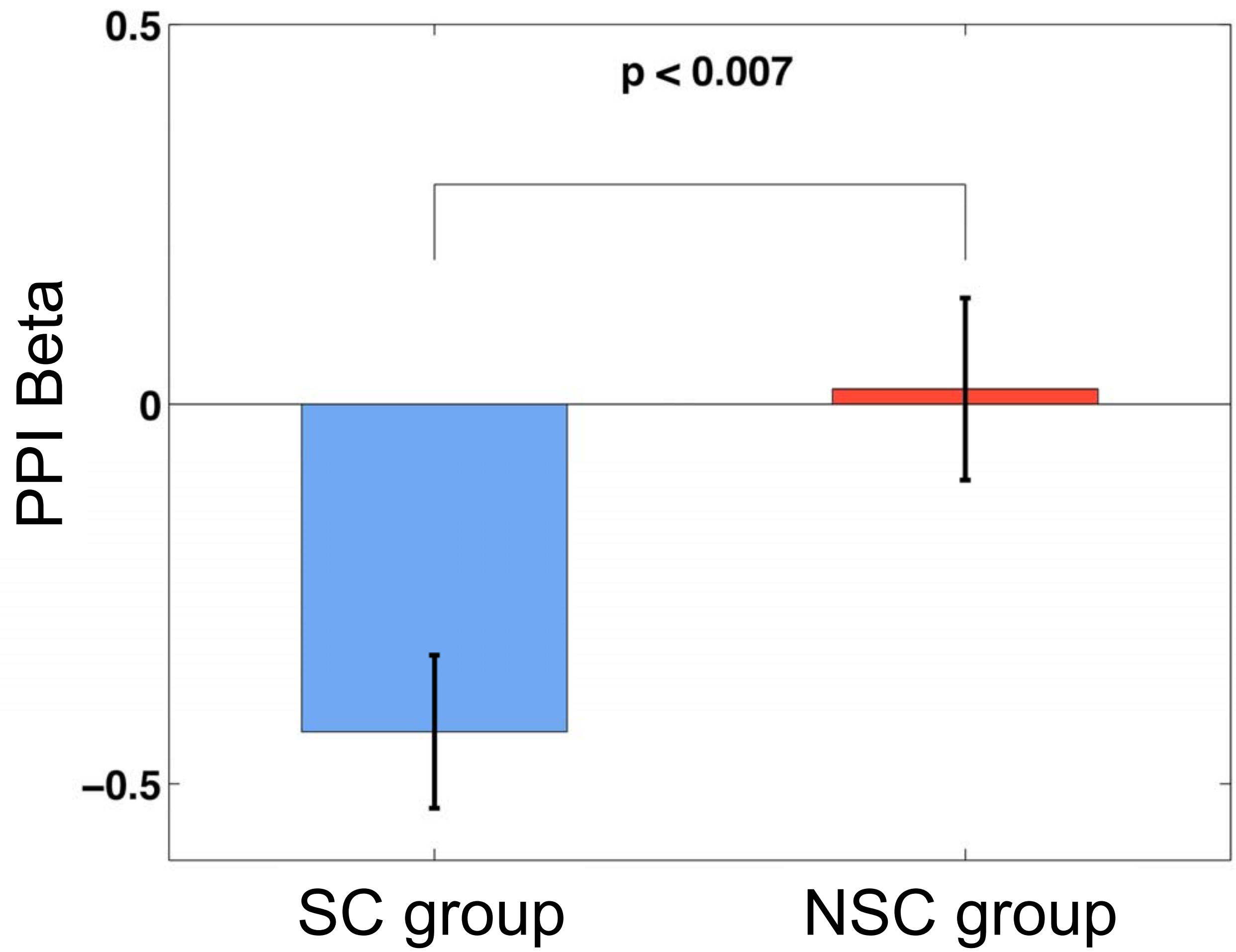




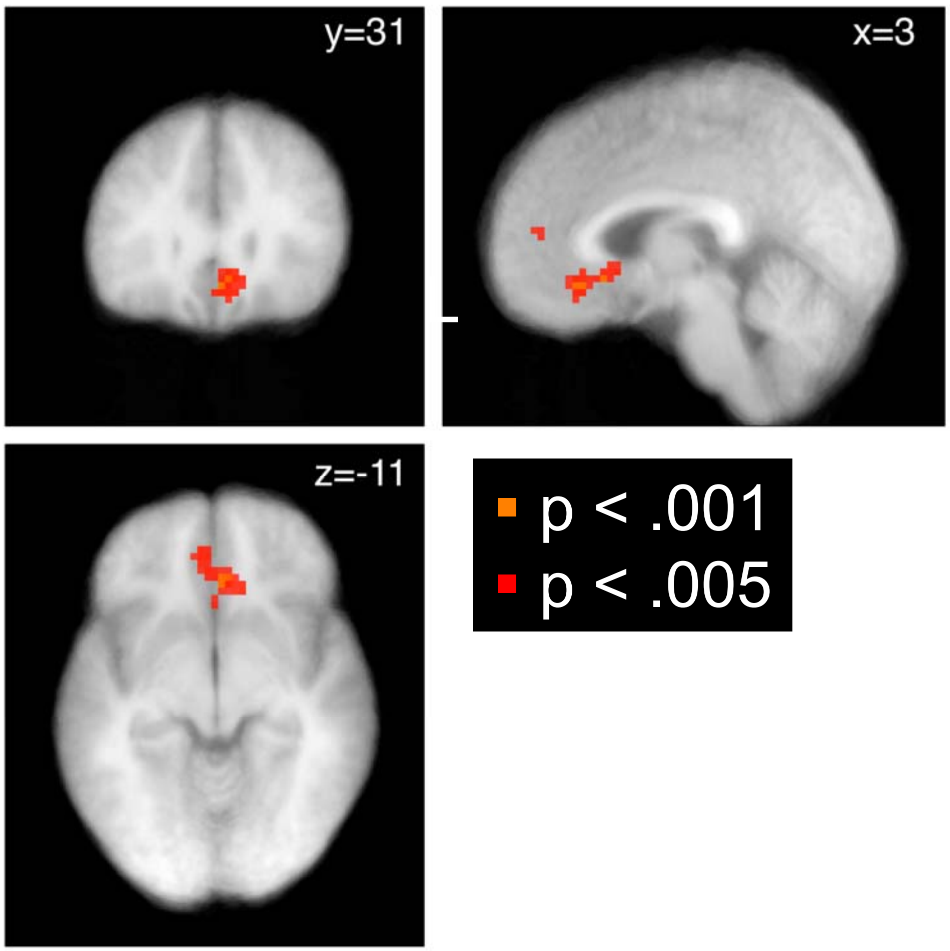
Justification for selection

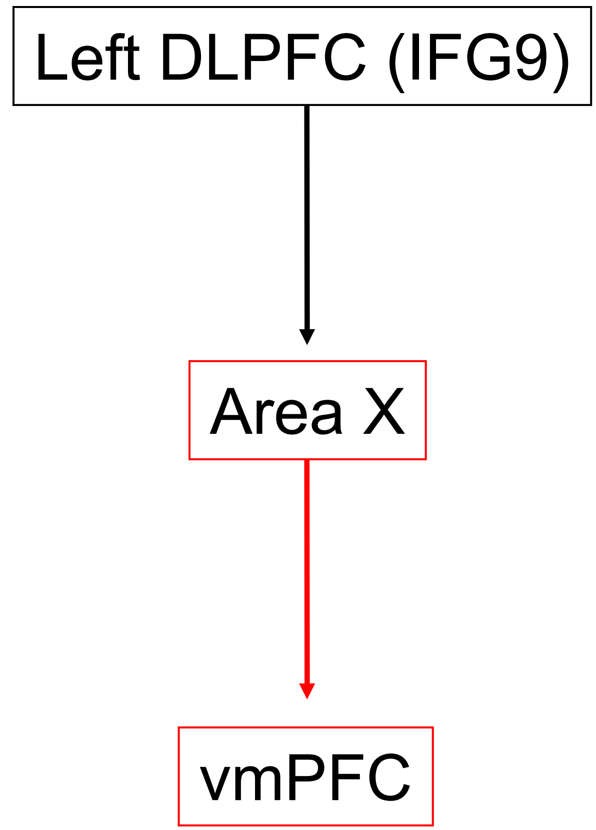


Group difference in PPI

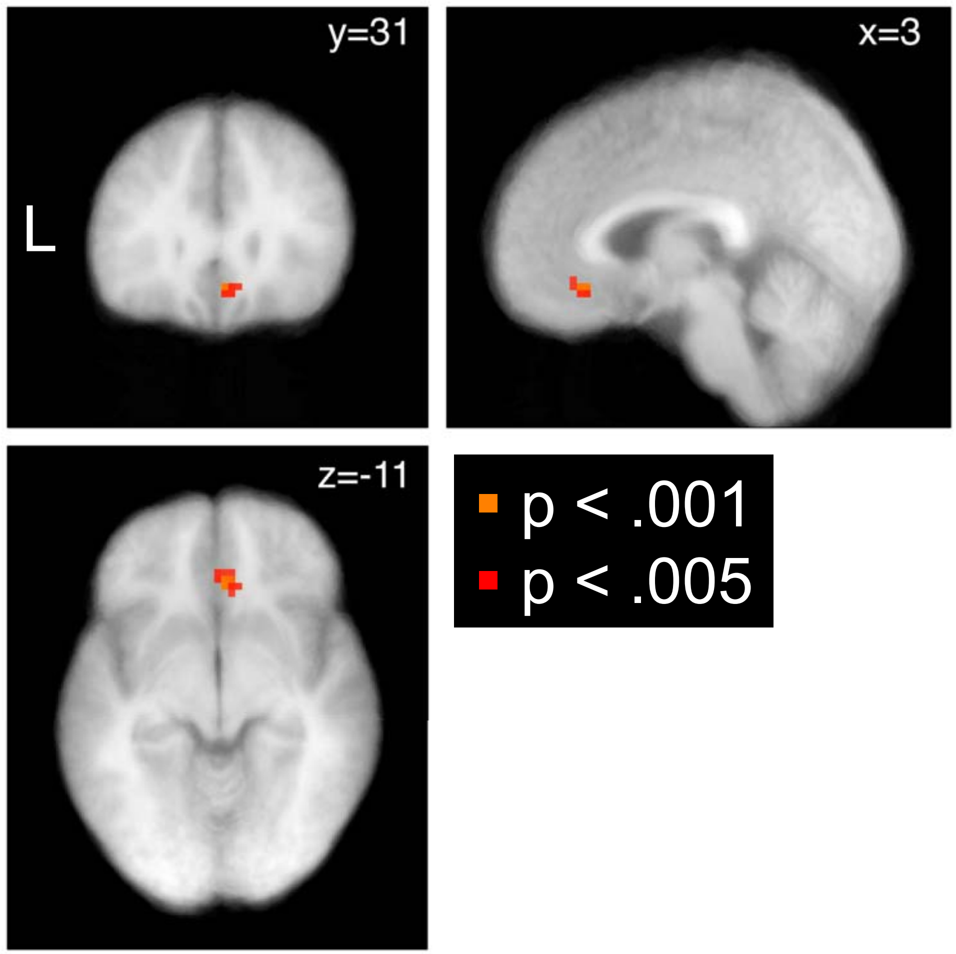


Node 2: Functional connectivity with anterior left DLPFC



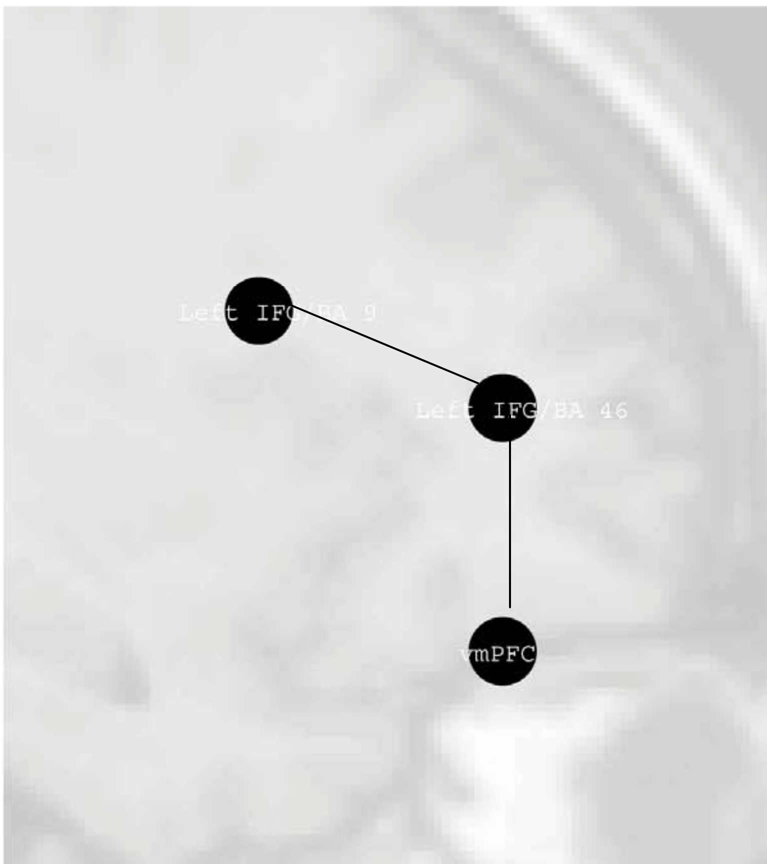


Result: Conjunction of regions showing connectivity and correlation with values

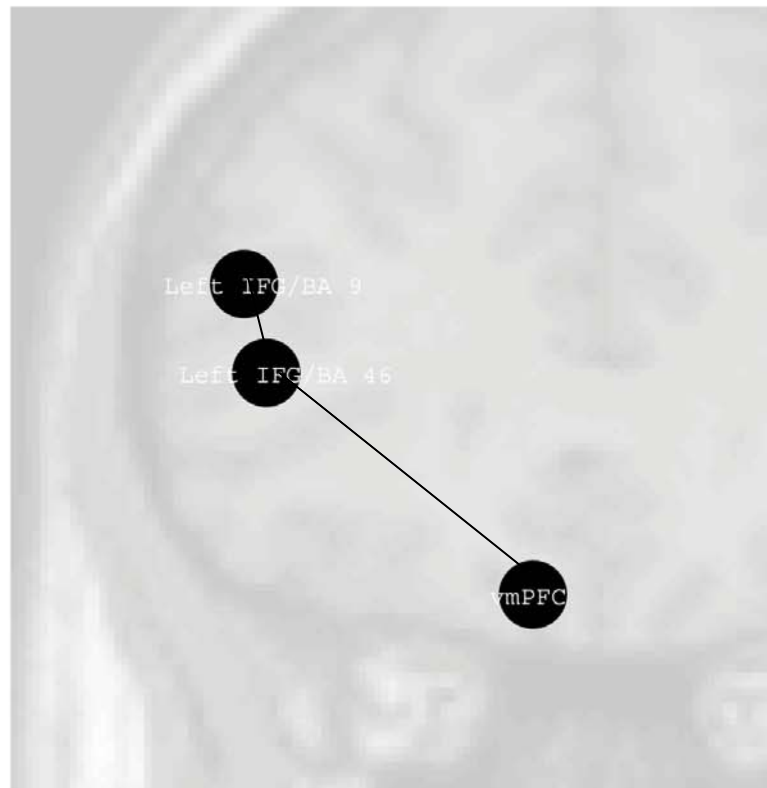


Self-control network

Sagittal



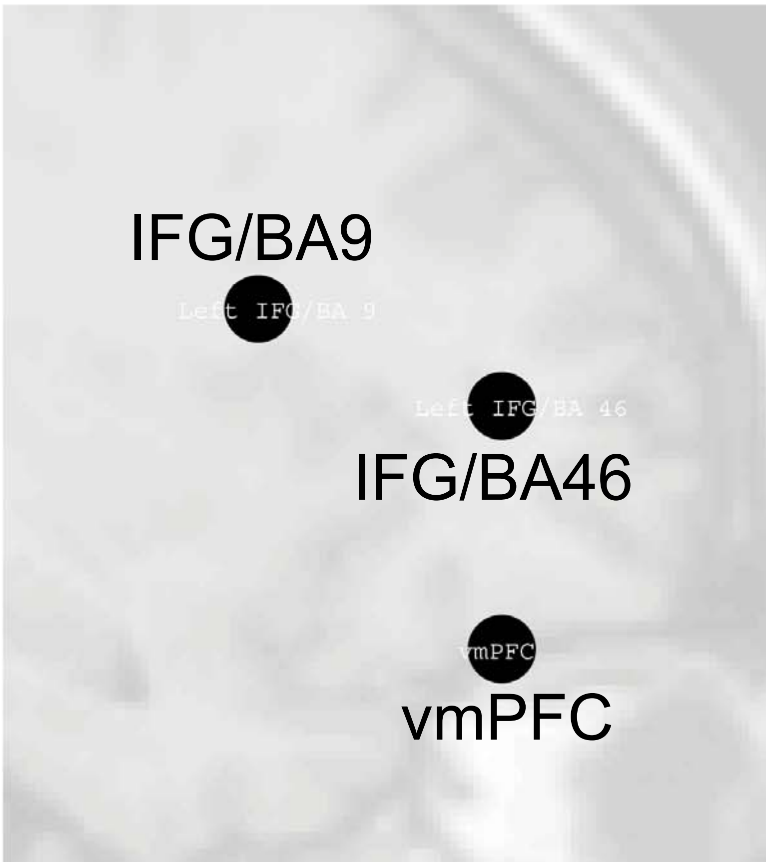
Coronal

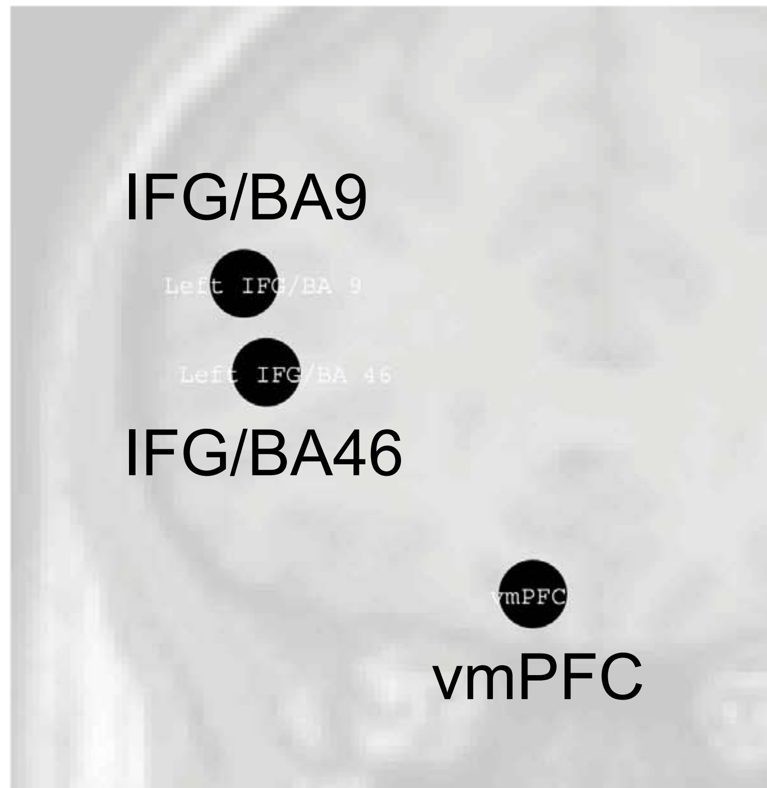


Dynamic Causal Modeling

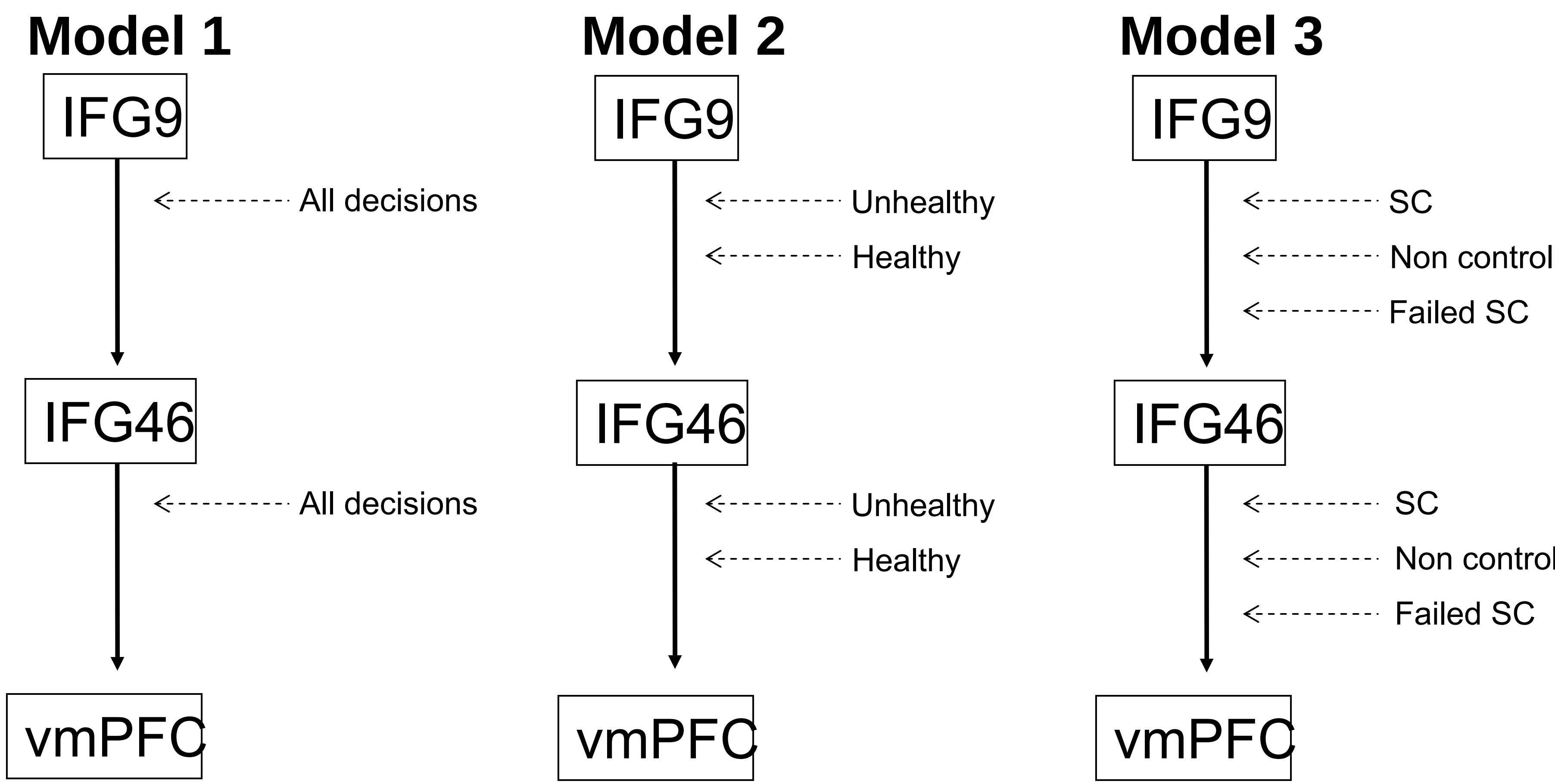
- PPI shows that activity in two brain regions is correlated in task dependent manner.
- Does not provide information about causation or directionality
- Dynamic Causal Modeling (DCM) can provide evidence for directionality
  Friston et al (2003) *NeuroImage*; Marreiros et al (2008) *NeuroImage*

Regions modeled

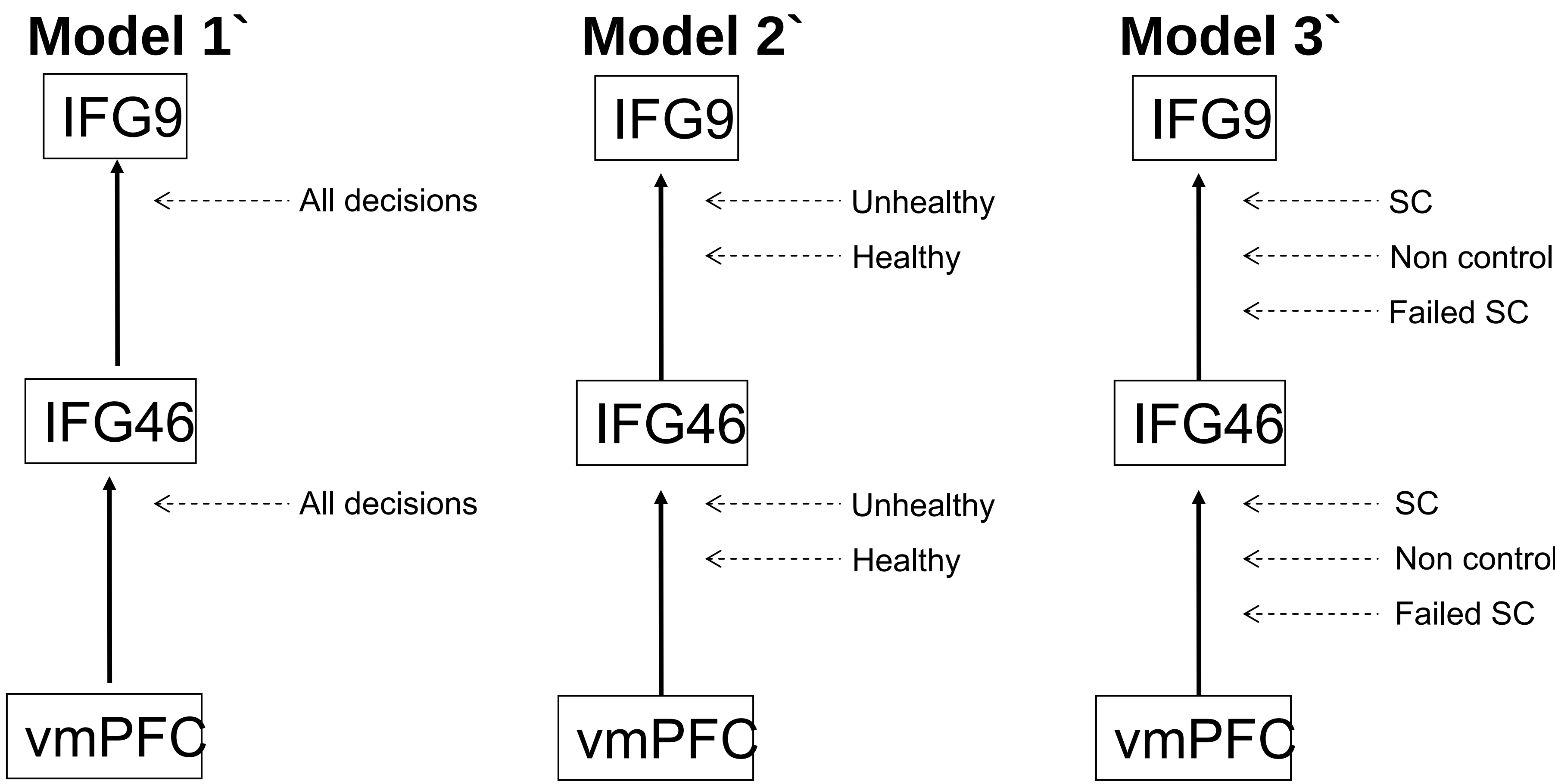




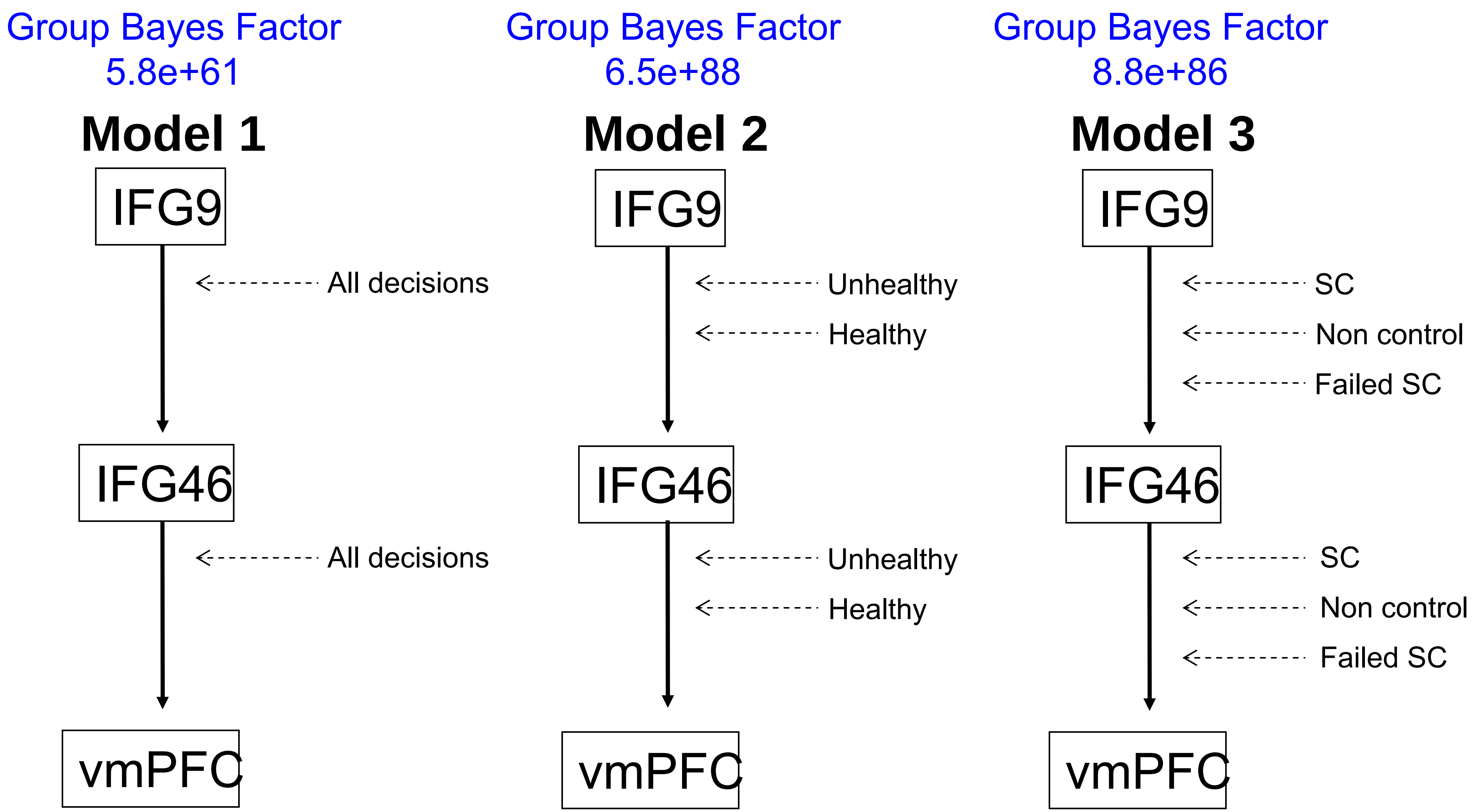
Multiple models tested



Multiple models tested

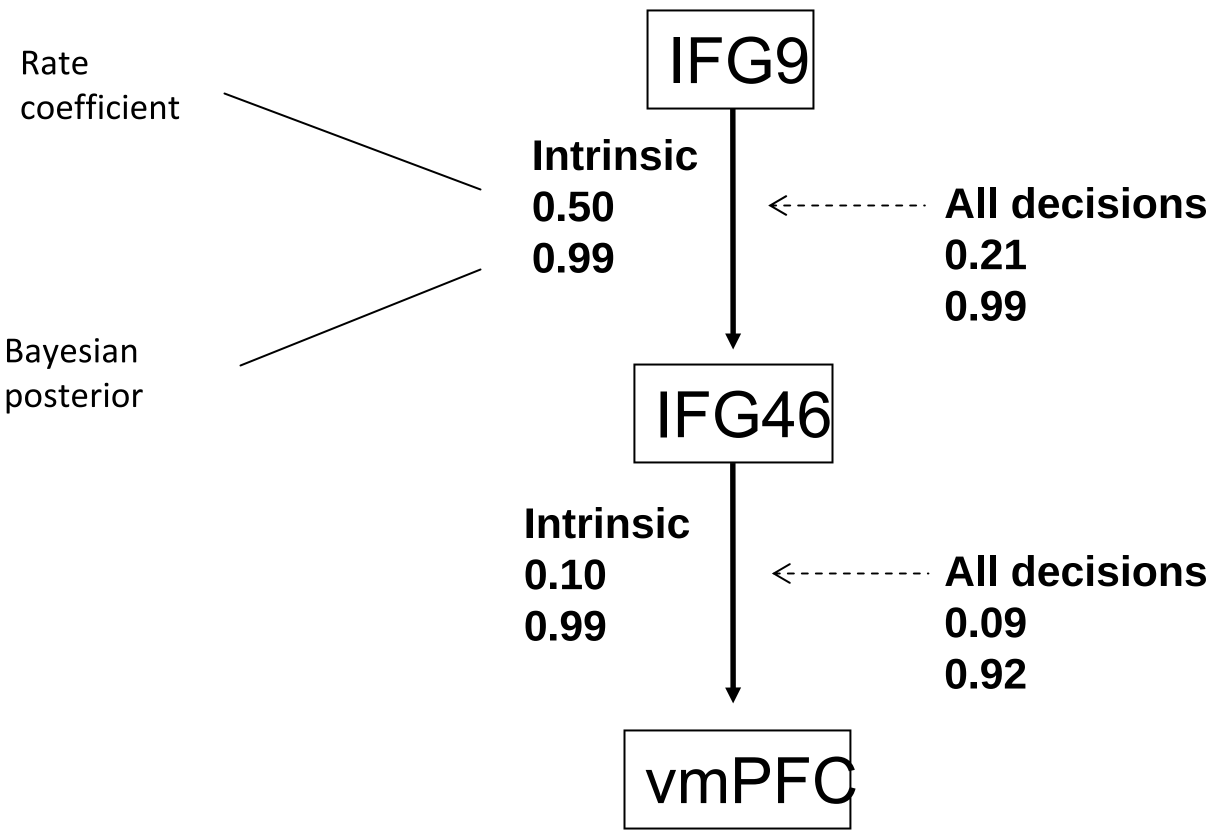


Multiple models tested



Best Model

Group Bayes Factor
>9e+100 for both comparisons



Summary of DCM results

Additional evidence for the hypothesis that that self-control involves DLPFC modulation of value computations in the vmPFC

What is Self-Control for the brain?

I: Goal directed decisions are guided by a common valuation system

II: Self-control involves DLPFC modulation of this valuation network

Evidence for Hypothesis I: Common value system

- Correlation between behavioral measures of value and neural activity in vmPFC across several studies
- vmPFC regions correlated with decision values reflect both taste and health ratings during self-control
  - Consistent with value computed based on two separate factors in the same region

Evidence for Hypothesis II: DLPFC modulation of value computation

- DLPFC is more active in the self-control group than the non-self-control group

- PPI & DCM results showing that activity in DLPFC influences activity in vmPFC value computation regions through a two node network

Making sense of previous fMRI literature on the neural basis of intertemporal discounting

- There seems to be a single value system
  - consistent with findings of Kable & Glimcher
- This value signal is modulated by DLPFC when self-control is required
  - consistent with McClure and colleagues showing increased DLPFC when choosing later larger reward

Bottom Line

- There seems to be a single system that computes a common value across multiple factors during goal directed decision-making.
- Self-control is exercised by DLPFC which modulates the input weights these factors have on the value signal computed in vmPFC networks.

Potential implications

- At their core, self-control deficits may be deficits of DLPFC function or connectivity between DLPFC and valuation areas
- This has implications for:
  - early detection of self-control deficits
  - co-morbidities between emotion regulation, cognitive control, and self-control failures

Kable and Glimcher (2007)

- Task: subjects make intertemporal monetary choices.
- Fit hyperbolic discount functions to subjects' behavior.
- Identifying assumption: behavioral preferences should match neural activation up to a linear transformation.
- Find brain regions that match the hyperbolic discount function estimated on behavioral data.
- This analysis identifies mPFC, PCC, VStr.
- Are these integration regions? (Especially PCC?).

Kable and Glimcher (2007)

Why don't Kable and Glimcher also identify PFC and Parietal Cortex?

1. These ROIs are *not* involved in the decision?
2. These ROIs *are* involved in the decision but they don't show sensitivity to time delay for this money task? (cf McClure et al 2004)
3. These ROIs *are* involved in the decision and they *do* show sensitivity to time delay but that sensitivity is a small part of total variance (at the subject level), so these regions are not accepted by the filter?

Hariri et al (2007)

- Measure individual differences in VStr responsivity using a task involving positive and negative feedback with monetary reward.
- Measure individual differences in impatience using choices among time-dated rewards.
- VStr responsivity correlates with impatience.
- “Increased preference for smaller immediate over larger delayed rewards reflects both a relatively indiscriminate and hyper-reactive VStr circuitry.”

Future work:

1. Are multiple system models a useful way of generating new hypotheses and models?
2. Are these systems localized? If so, where?
3. How do the systems communicate?
4. How are the inputs integrated?
5. When are the systems cooperative and when conflictual?
6. When they are in conflict, are they strategic?
7. What manipulations enhance or weaken the signals coming from these systems?
8. Can we influence individual systems in the lab?
9. Can we influence individual systems in the field?
10. Can we produce useful formalizations of their operation?

The Sweet Taste of Revenge

de Quervain et al
Science (2004)

- Two subjects, A and B, have the opportunity for a social exchange that benefits both parties.
- However, to complete the exchange, A has to move first while B moves second, i.e., A has to trust B, and B can cheat A.
- A is informed about B‘s action and A has an opportunity to punish B.
- A‘s brain is scanned when he is informed about B‘s action (i.e, whether B cheated him or not) and when A deliberates whether, and how much, he wants to punish B.

- Stage 1:
  - Both A and B receive an endowment of 10 MUs (1 MU = CHF 0.1).
  - A decides whether to keep or to transfer his endowment to B. In case of a transfer B receives 4 times A‘s endowment (40 MUs), in addition to his own endowment of 10 MUs.
- Stage 2:
  - B has the option to keep all the points he possesses or to give half of his MUs to A
    - If A trusts and B is trustworthy both earn 25.
    - If A trusts and B cheats, A has nothing and B has 50.
    - Both have 10 if A does not trust and B keeps everything.
- Stage 3:
  - Both players receive an additional endowment of 20 Mus
  - A has the opportunity to use this money to punish B or to keep the money.
  - Every MU invested into the punishment of B decreases B‘s payoff by 2 MUs.

Treatments

Player A sequentially faced 7 different B players.

- In 3 cases B cooperated (transferred half of his MUs back).
- In 3 cases B cheated (transferred no MUs back).
- In 1 case a random device forced B to „cheat“. A knew that in this case B‘s „decision“ was determined by rolling a die.

More treatments (when B cheats)

- Treatment where B decides himself and punishment is costly for both A and B (**Intentional & Costly, IC**). A is hypothesized to experience a desire to punish cheating and he can in fact punish.
- Treatment where B decides himself and punishment is free for A but costly for B (**Intentional & Free, IF**). A is hypothesized to experience a desire to punish cheating and he can in fact punish – even without a cost.
- Treatment where B decides himself and punishment is only symbolic, i.e., A and B have no costs of punishing (**Intentional & Symbolic; IS**). A is also hypothesized to experience a desire to punish cheating but he cannot punish.
- Treatment where B‘s decision is randomly imposed on B; punishment is costly for both A and B (**Non-intentional & Costly, NC**). A is to have no or a much less of a desire to punish B for „non-cooperation“ because B is not responsible for his „action“.

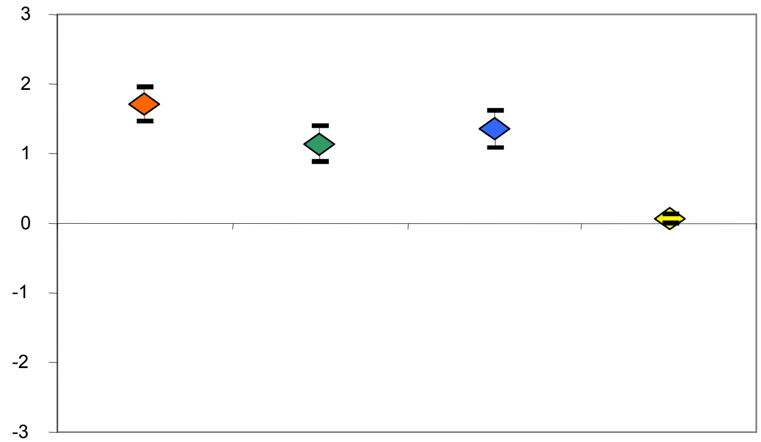
Hypotheses:

- The following contrasts are hypothesized to show activation of reward related (N Accumbens, N caudate) neural circuits:
  - (IC – IS) because no punishment possible in IS.
  - (IF – IS) because no punishment possible in IS.
  - (IC – NC) because no desire to punish in NC.
  - (IF – NC) because no desire to punish in NC.
  - (IC + IF) – (IS + NC) because in IS & NC there is either no possibility or no desire to punish so that there cannot be any satisfaction from punishing.
- The following contrast is hypothesized to show activation of cognition related (PFC, OFC) neural circuits.
  - (IC – IF) because in IC subjects must weigh the costs and benefits of punishing while in IF there are no costs.
  - Related to the Sanfey et al. (2003) hypothesis.

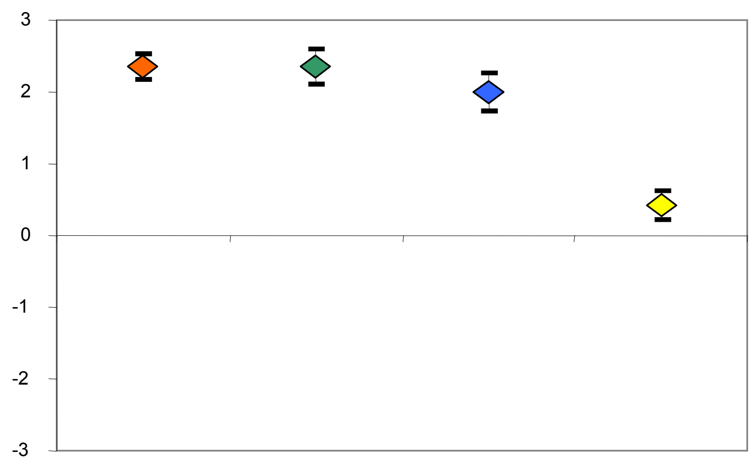
Behavioral and Questionnaire Results

intentional & costly IC
intentional & free IF
intentional & symbolic IS
non-intentional & costly NC

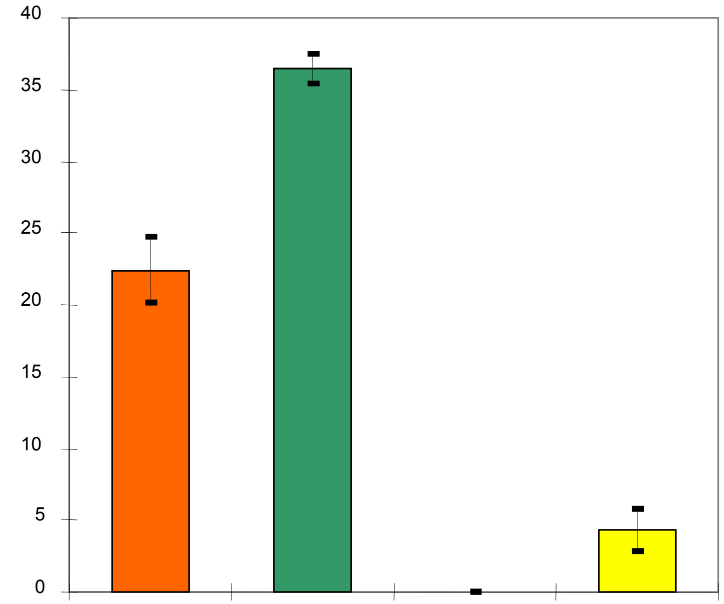
Perceived unfairness



Desire to punish

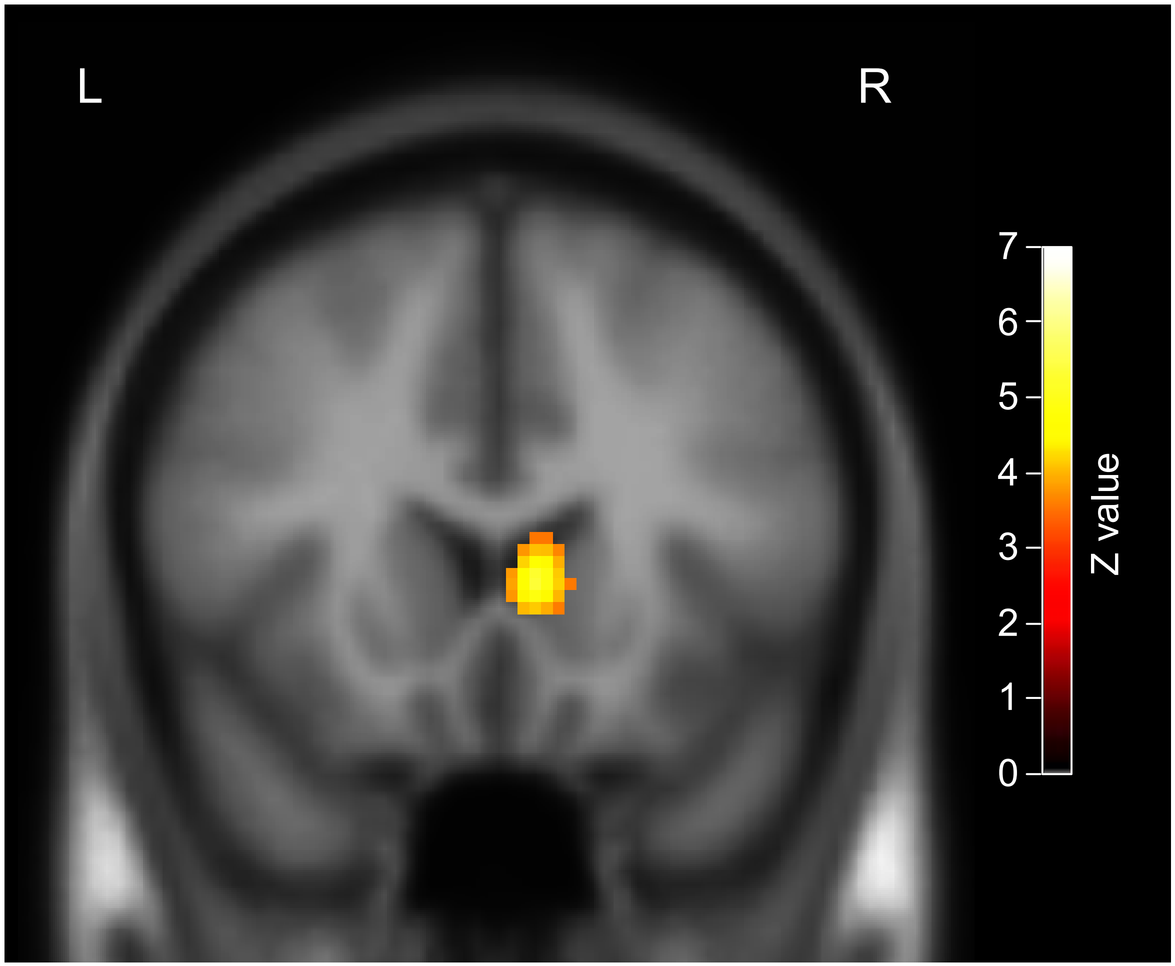


Actual payoff reduction imposed on B



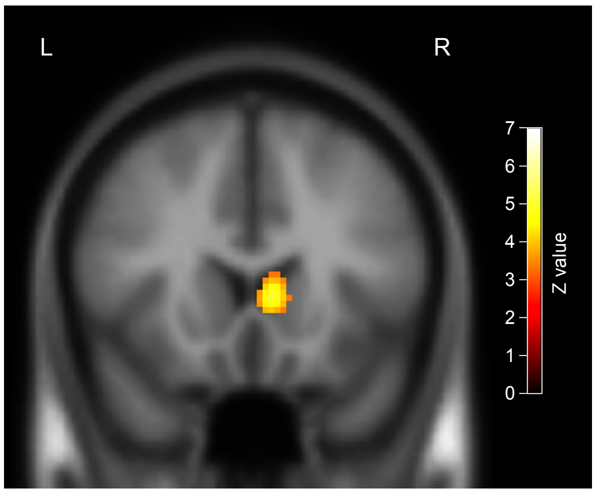
(IC + IF) - (IS + NC)

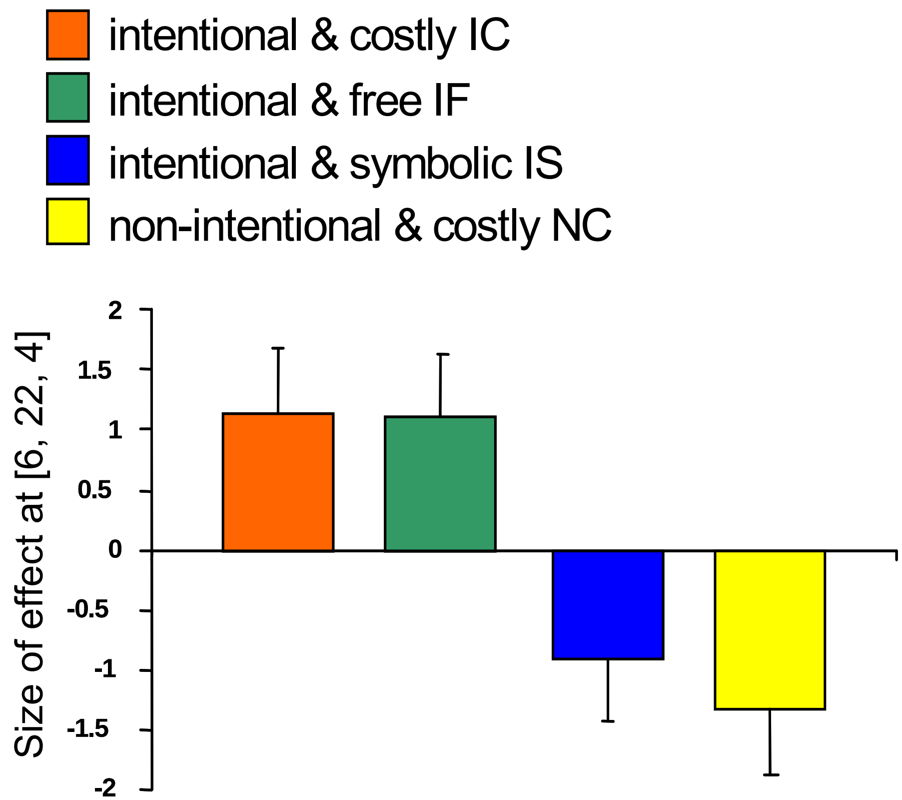
activates the nucleus caudate



Differential activation of Nucleus Caudate relative to mean brain activation

(IC + IF) - (IS + NC)





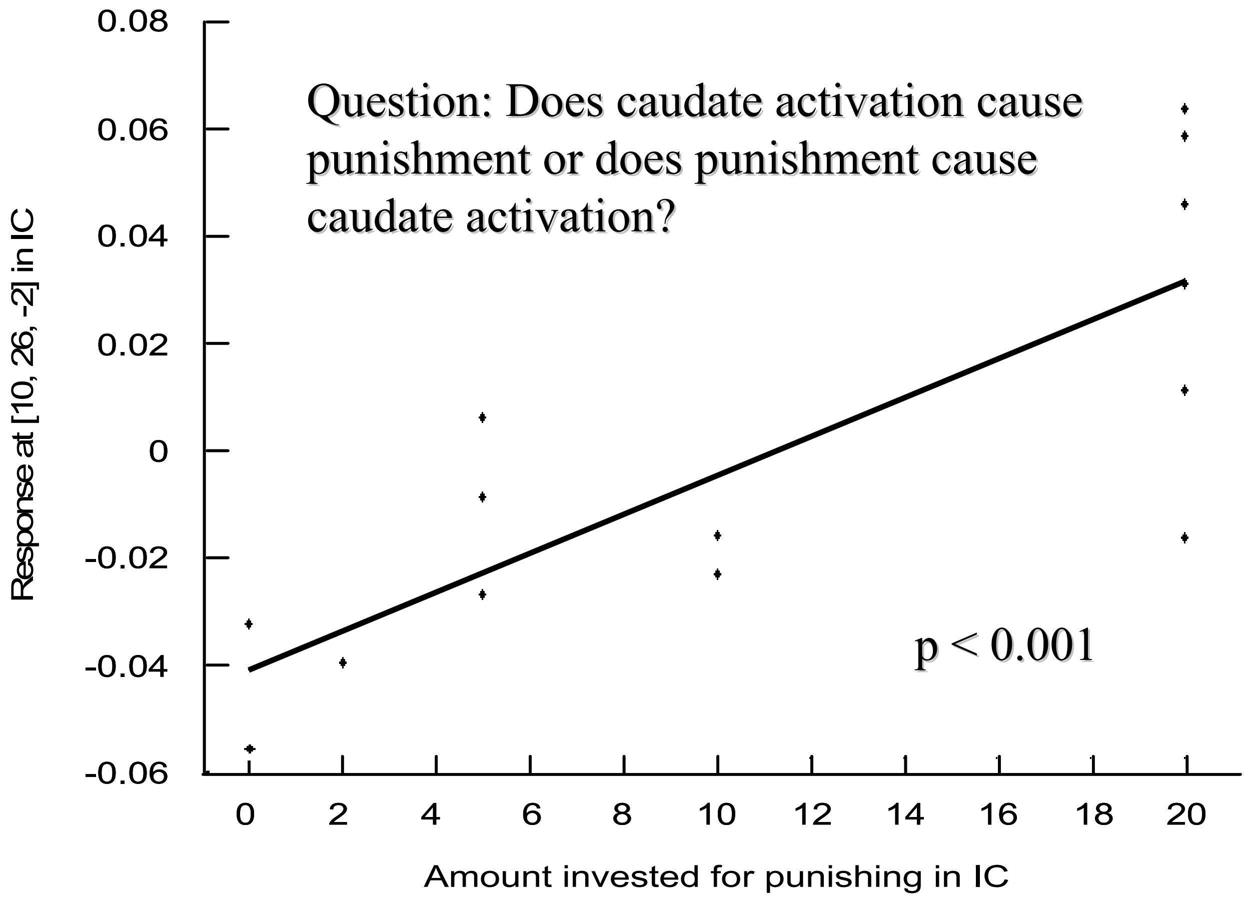
Differential activations across contrasts

Table 1. PET results

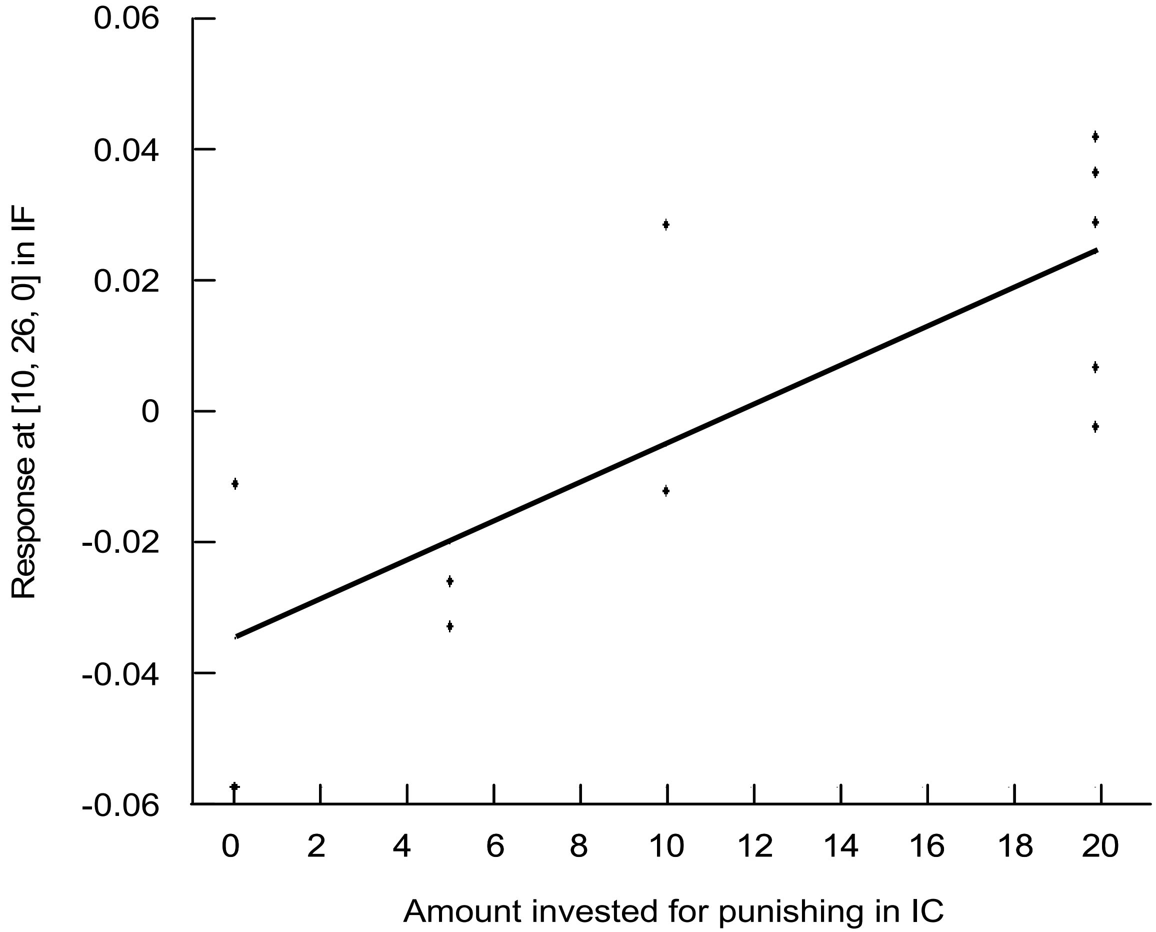
Contrast	Region (BA)	Coordinates			Z value
		x	**y**	**z**	
(IC+IF)-(IS+NC)	Caudate nucleus	6	22	4	5.11*
	Thalamus	22	-24	10	4.43*
IF-IS	Thalamus	22	-22	10	4.21
	Caudate nucleus	6	22	4	3.55
IC-IS	Thalamus	22	-22	10	4.15
	Caudate nucleus	6	24	2	3.70
IF-NC	Caudate nucleus	6	22	4	4.18
IC-NC	Caudate nucleus	6	22	4	4.23
IC-IF	Ventromedial prefrontal /	2	54	-4	4.59
	medial orbitofrontal cortex	-4	52	-16	3.35

The coordinates (x, y, z) locate the maxima of changes in blood flow. * indicate significant activations at a level < 0.05 corrected, otherwise p<0.001 uncorrected. BA, Brodmann area. I, intentional no-transfer; N, non-intentional no-transfer; C, costly punishment; F, free punishment; S, symbolic punishment.

If caudate activity reflects satisfaction from punishment subjects with stronger caudate activation should punish more

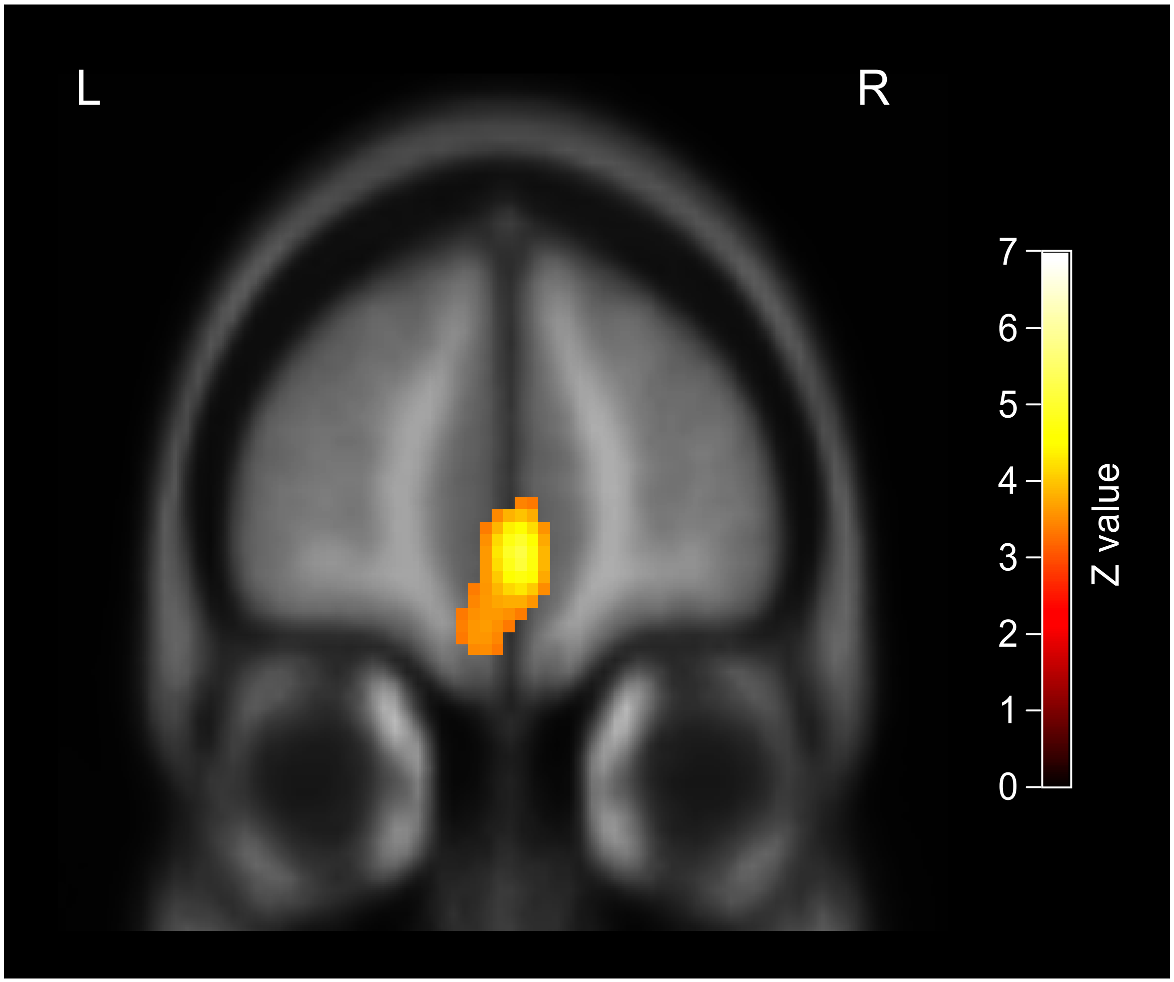


Are subjects with higher caudate activation in the IF condition willing to pay more for punishing in the IC condition?



If caudate activation reflects satisfaction from punishment subjects with stronger caudate activation in IF should be willing to pay more for punishment.

Does (IC – IF) activate the PFC/OFC?



Summary of revenge study

- Experimental evidence and social preference theories suggest that many people prefer to punish unfair acts.

- This study supports the view that satisfaction is associated with the punishment of unfair acts.

- Correlation between caudate activity and punishment in IC, and between caudate activity in IF and punishment in IC supports this interpretation.