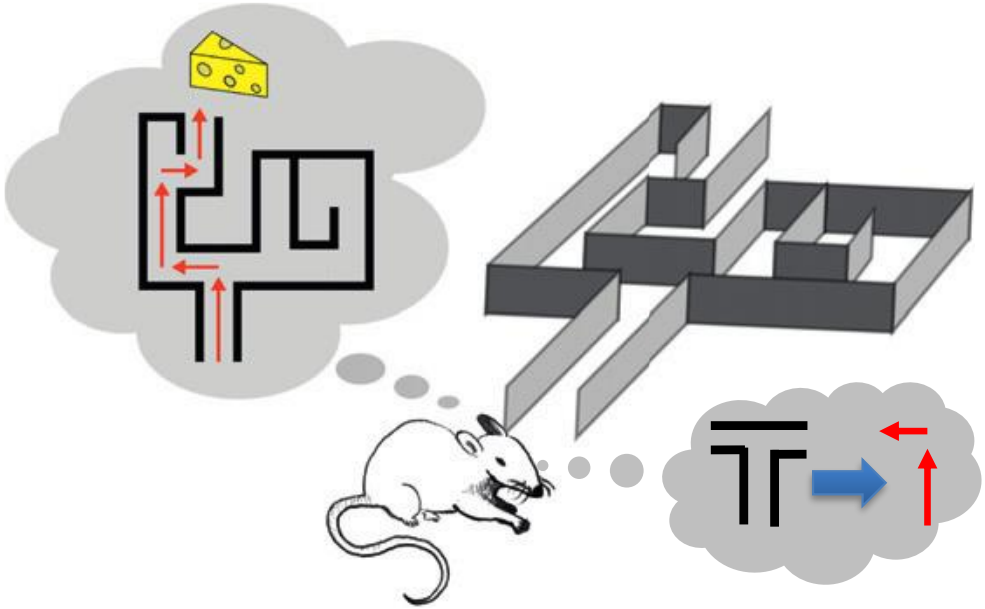
# computational mechanisms of decision making: individual differences

nathaniel daw
princeton neuroscience institute & department of psychology
princeton university

ringberg symposium

# learning for decisions



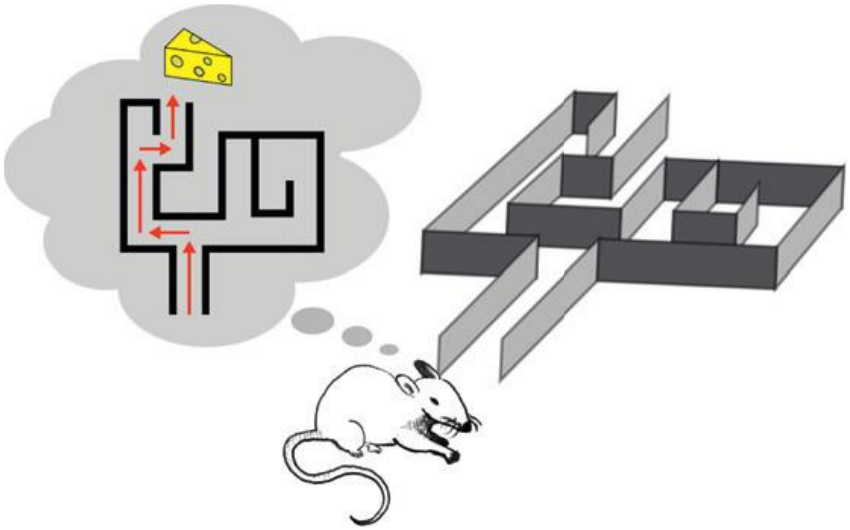
multiple systems for trial-and-error decision making:

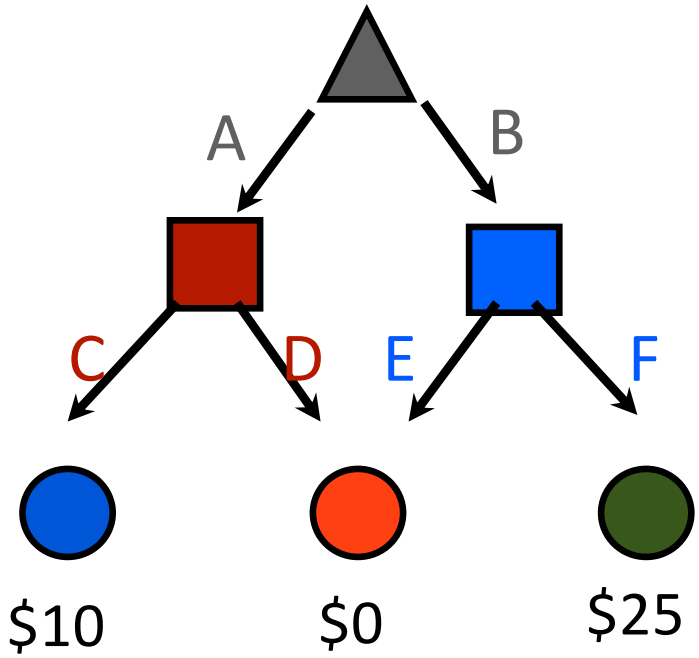
habits, slips of action, control, & compulsion

1. characterizing this distinction computationally via different learning strategies: model-based and model-free RL
2. individual variation in these mechanisms: eg. are they compromised in psychiatric disorders?

# model-based learning

$$Q(s_t, a_t) = r(s_t) + \sum_{s_{t+1}} P(s_{t+1}|s_t, a_t) \left[ r(s_{t+1}) + \sum_{s_{t+2}} \dots \right]$$



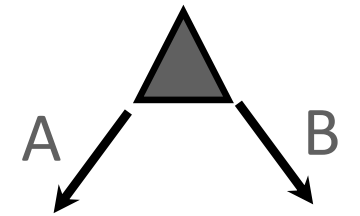



"model-based" learning:

- learn one-step rewards & transitions;
- iterative, tree-structured computation

# “model-free” learning

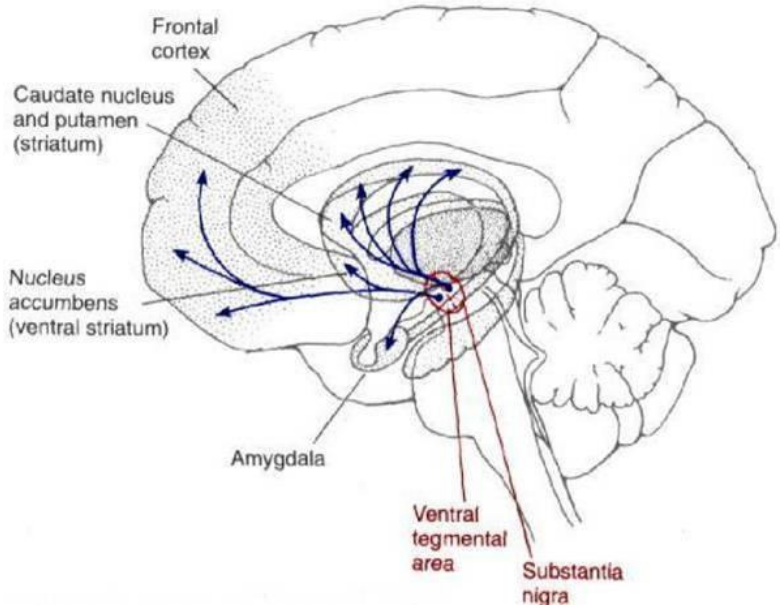
$$Q(s_t, a_t) = r(s_t) + \sum_{s_{t+1}} P(s_{t+1}|s_t, a_t)\left[r(s_{t+1}) + \sum_{s_{t+2}} \dots\right]$$

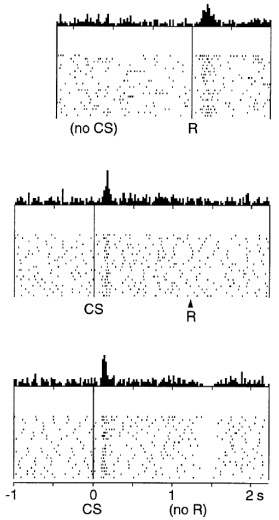



$10 $25

shortcut: cache endpoints of computation (long-run action values)

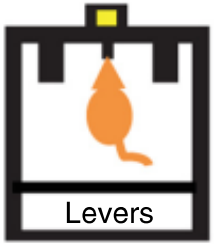
- simplifies choice-time computation (just retrieve)
- these can be learned directly (TD learning)
- standard theory of dopamine, reward prediction errors etc
- intermediates (e.g. partial evaluation)

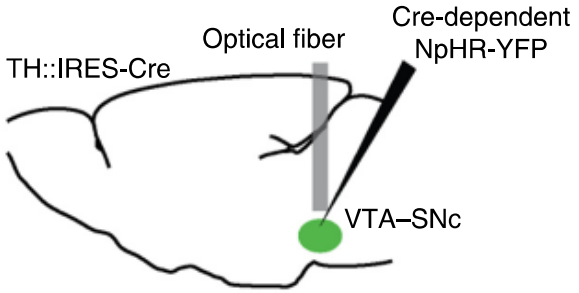



(Schultz et al 1997)

# model-free reinforcement



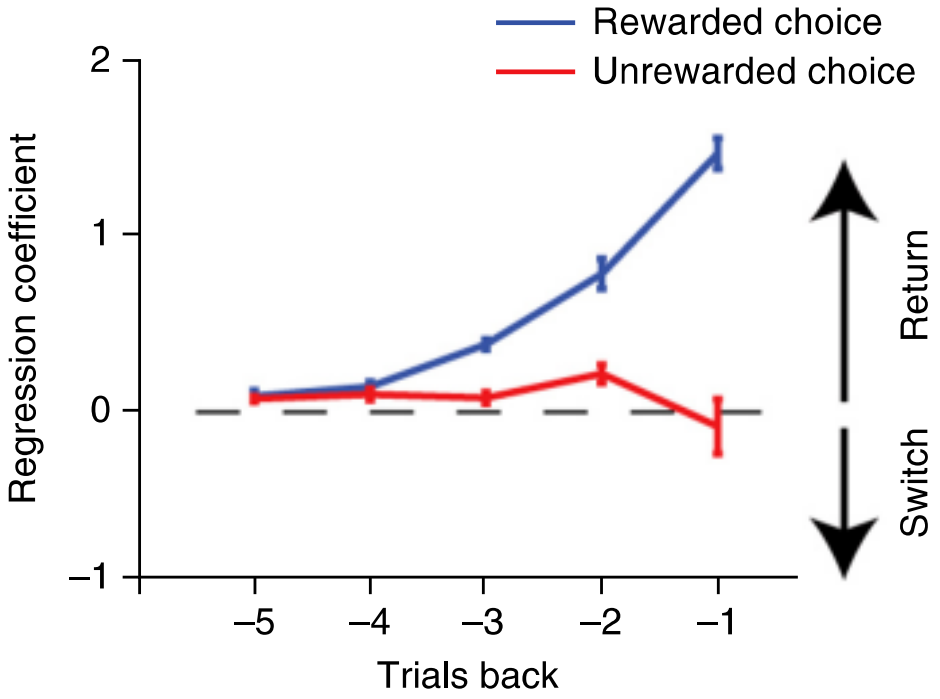

delta rule predicts exponential form

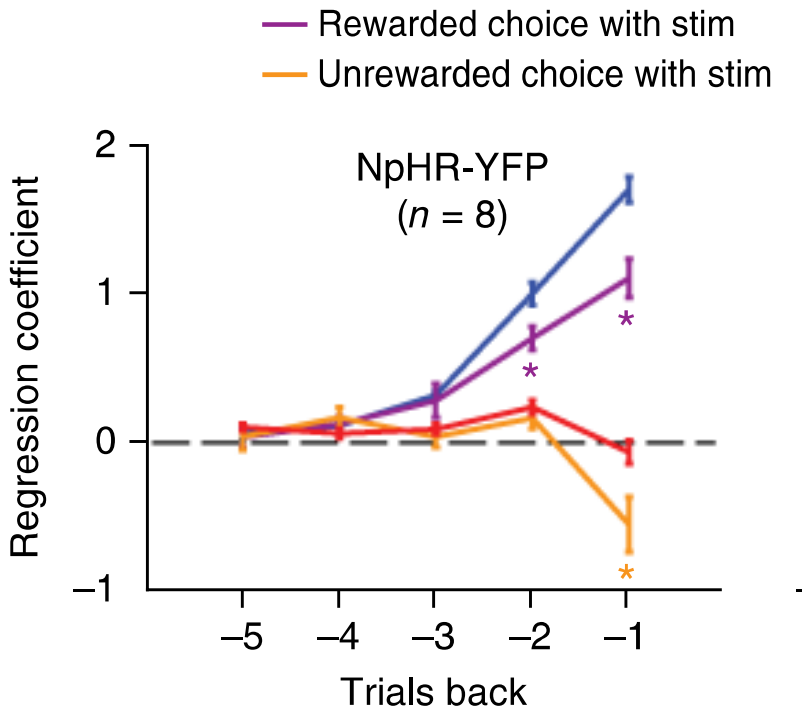
$$P(choice_t = c) \propto \exp(\beta V_t(c))$$

$$V_{t+1}(choice_t) = V_t(choice_t) + \alpha\delta_t$$

$$\delta_t = reward_t - V_t(choice_t)$$

timed suppression of dopamine neurons on 10% of trials



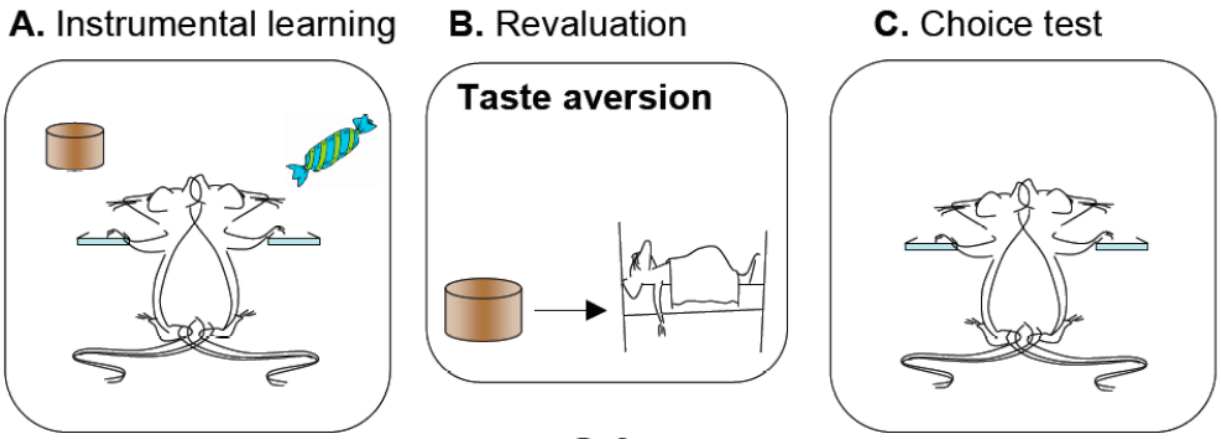




(Parker, …, Daw & Witten, Nature Neuroscience 2016)

# model-based learning

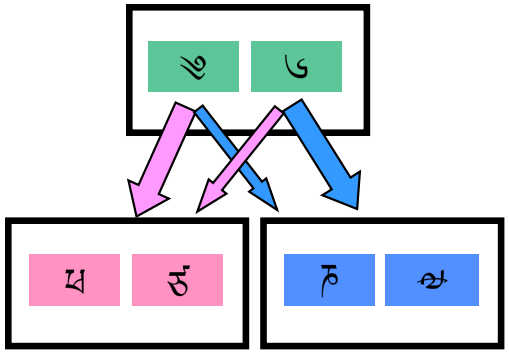
decision behavior and neural decision variables incorporate knowledge other than reinforcement history; integrate over separate experiences

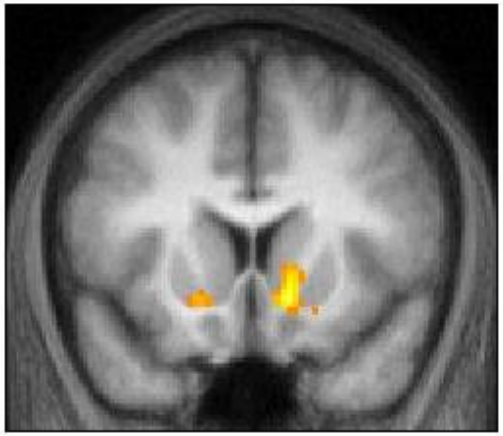
**revaluation**




(Dickinson, Balleine)

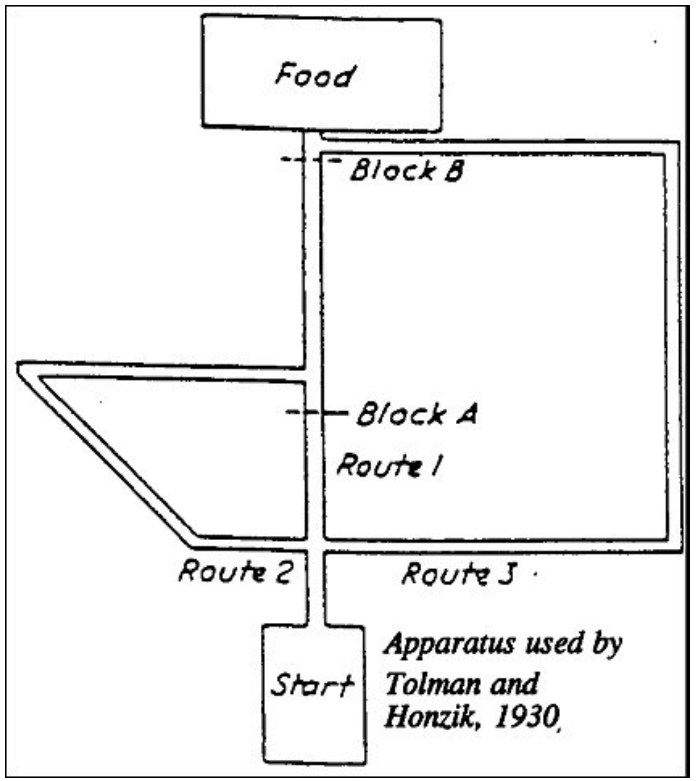
**sequential decision tasks**





(Daw et al. 2011)

**replanning**




$$Q(s,a) = r(s) + E[Q(s',a')]$$

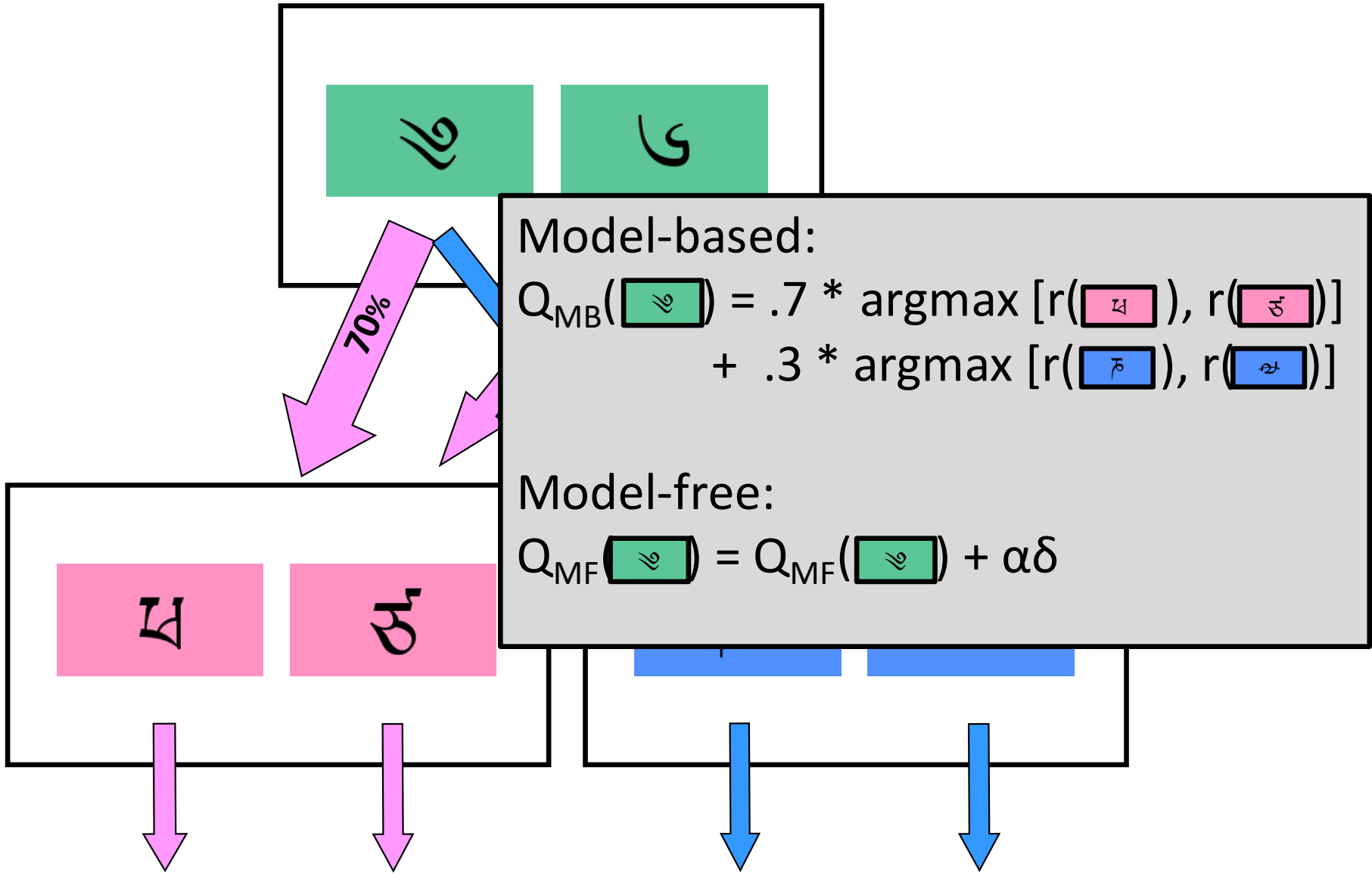
"model-free" "model-based"

(Doya 1999; Daw et al. 2005)

# issues

1) Can we study this tradeoff with more targeted experimental designs?
2) What is the mechanism of model-based evaluation?
3) How do these mechanism vary across individuals? Is there a relationship between compulsion and habits?

# sequential decision task




*with prob*: 26% 57% 41% 28%

(*all slowly changing*)

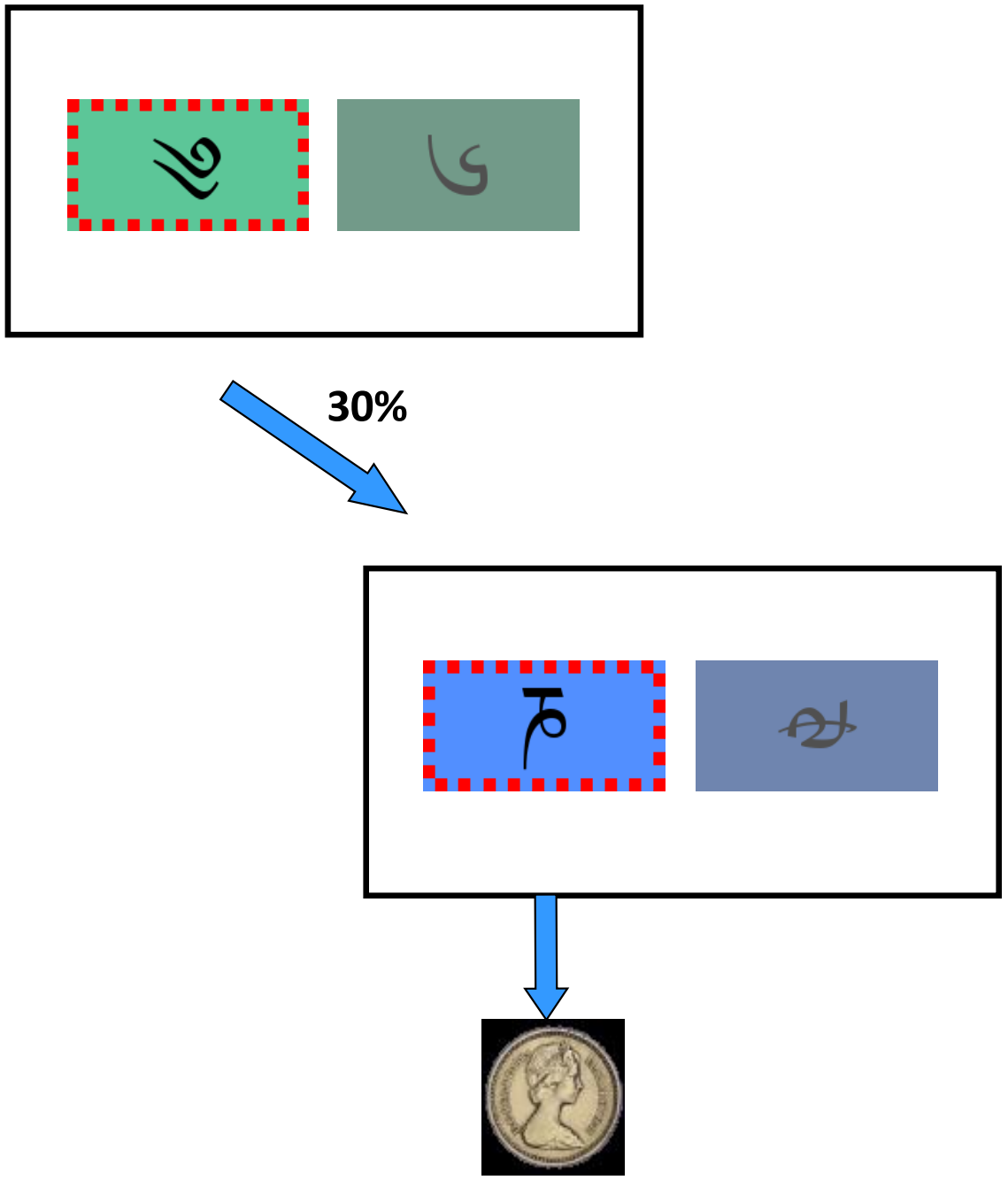
(Daw et al Neuron 2011)

# idea

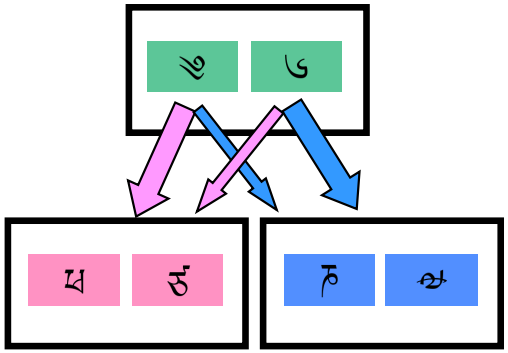
Are top-stage valuations mediated by bottom-stage states?

Example: rare transition at top level, followed by win

- Which top-stage action is now favored?

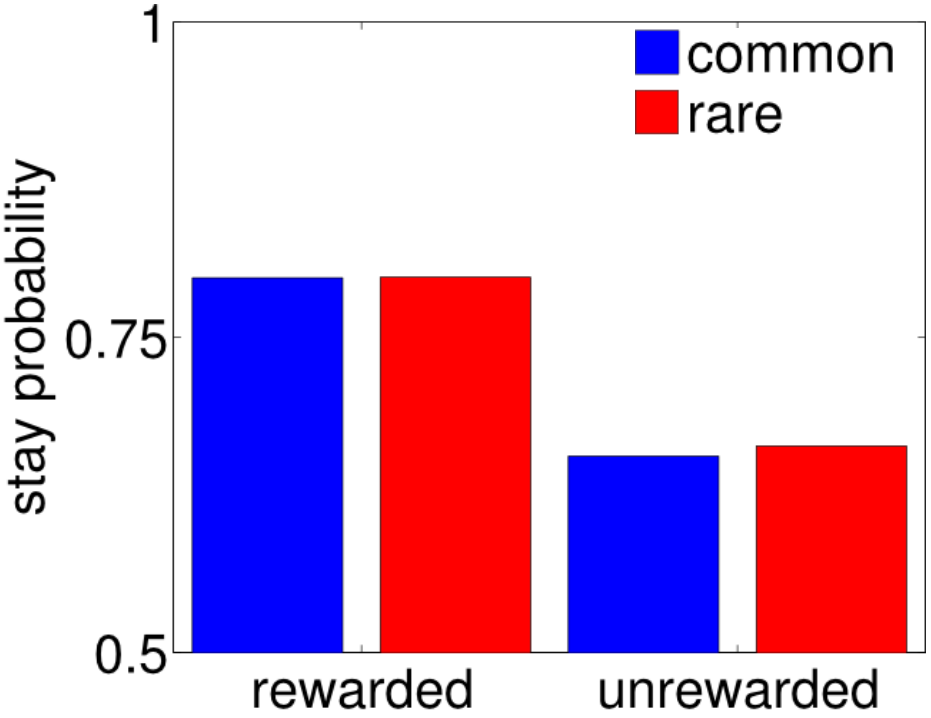


# predictions



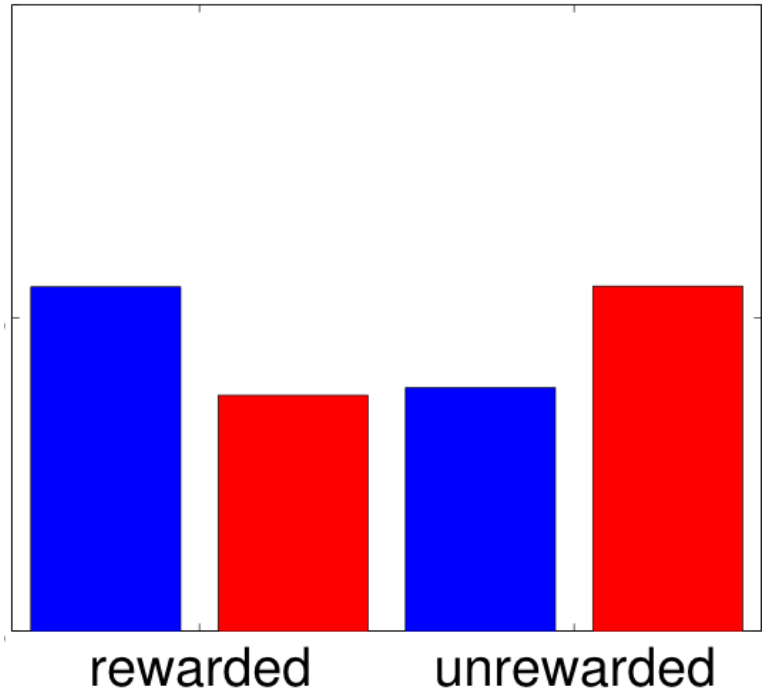
direct reinforcement

ignores transition structure

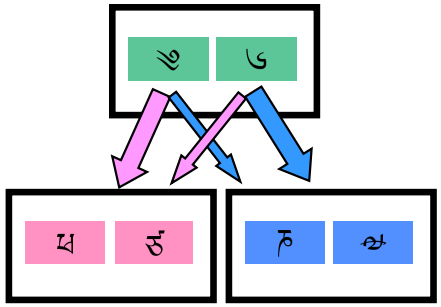



model-based planning

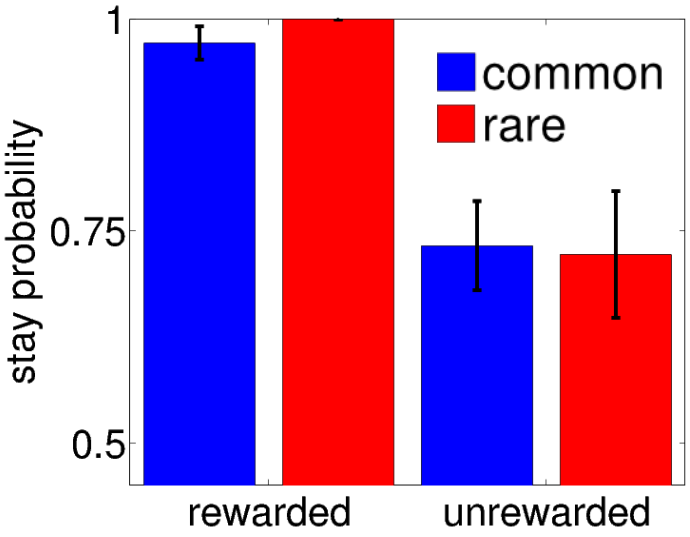
respects transition structure

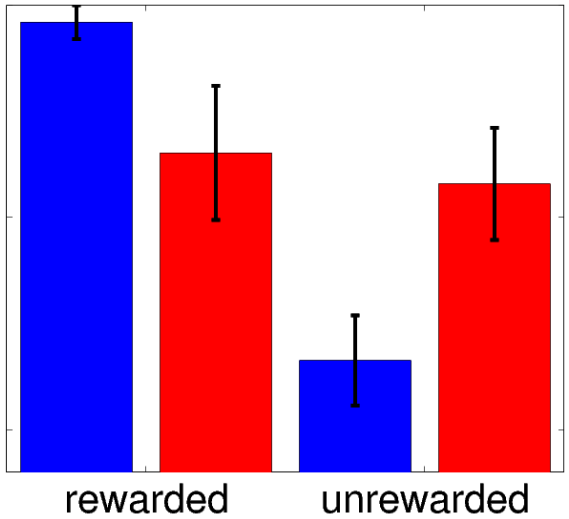


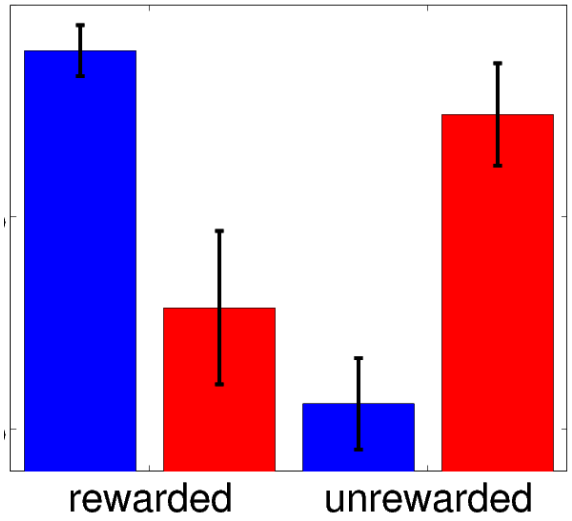
# data



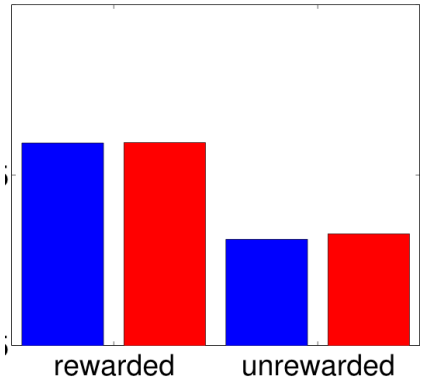
individual subs x 201 trials each



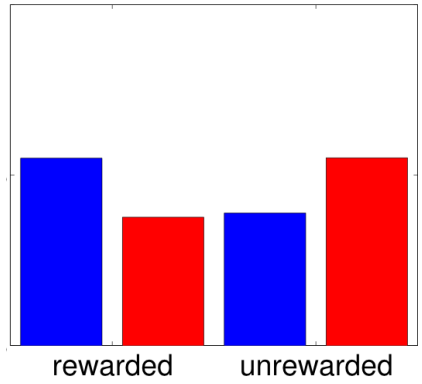




reinforcement

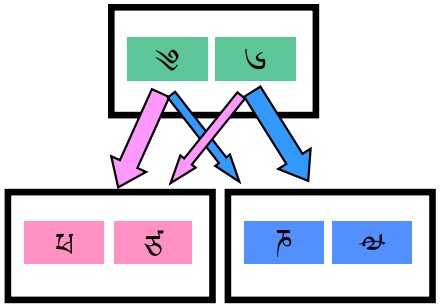



planning

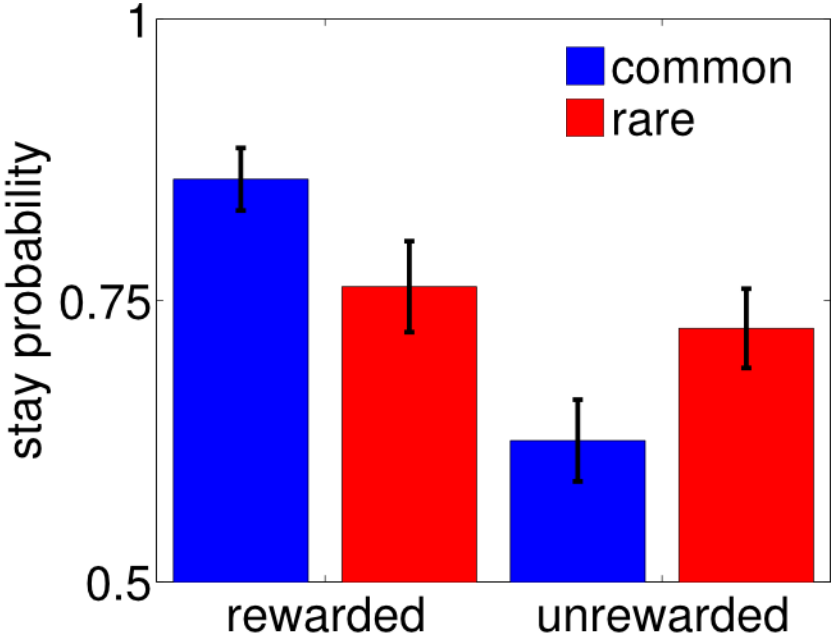



(Daw et al Neuron 2011)

# data

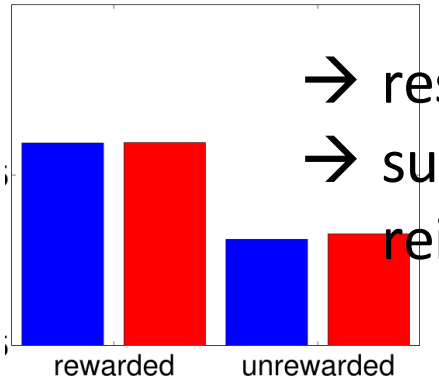


17 subs x 201 trials each



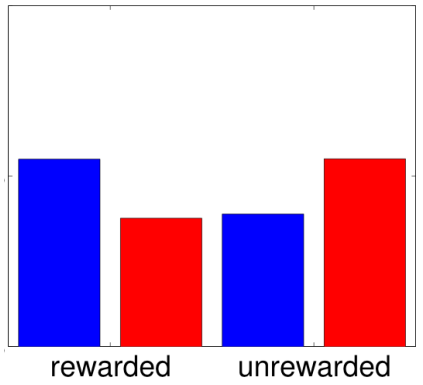

reward: p<1e-8
reward x rare: p<5e-5
(mixed effects logit)

reinforcement



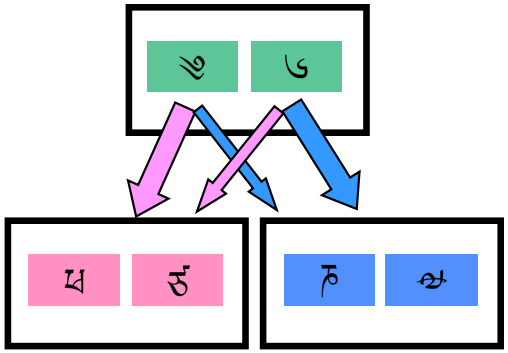

→ results reject pure reinforcement models
→ suggest mixture of planning and reinforcement processes

planning

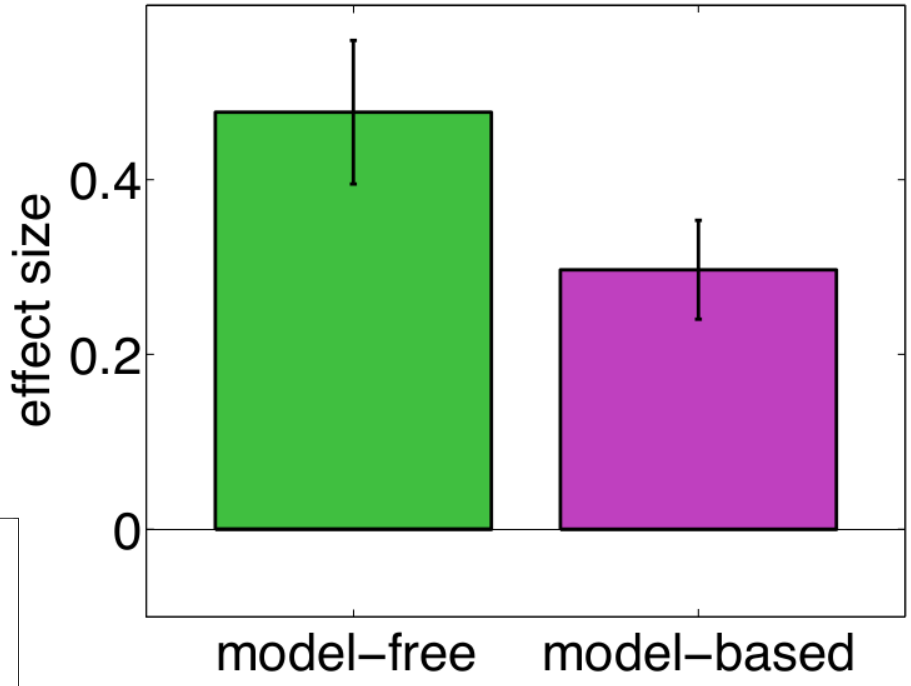



(Daw et al Neuron 2011)

# data

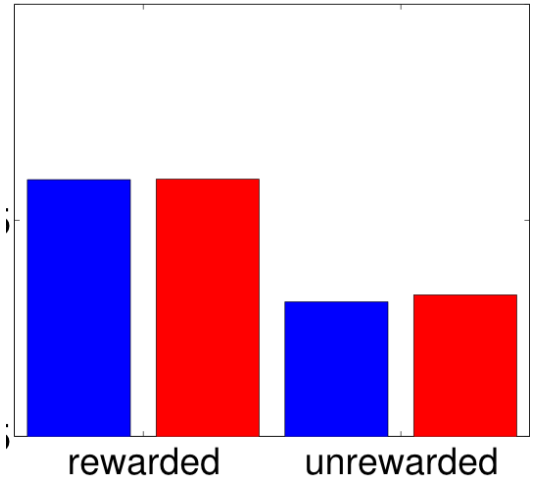


17 subs x 201 trials each

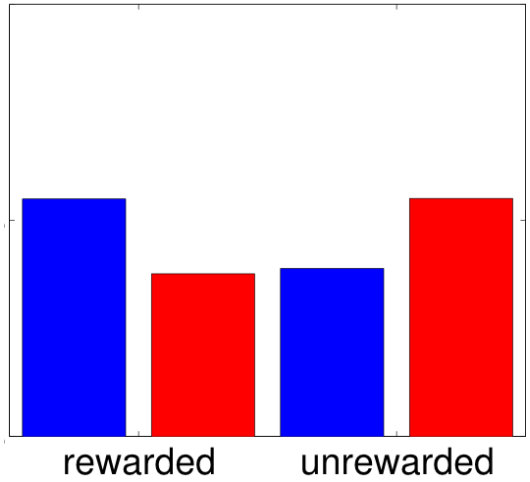



reward: p<1e-8
reward x rare: p<5e-5
(mixed effects logit)

reinforcement



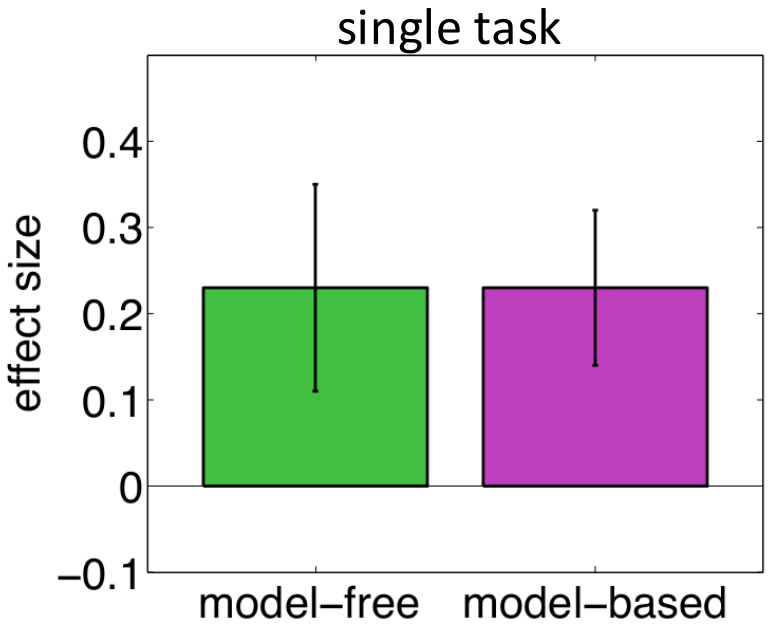

planning

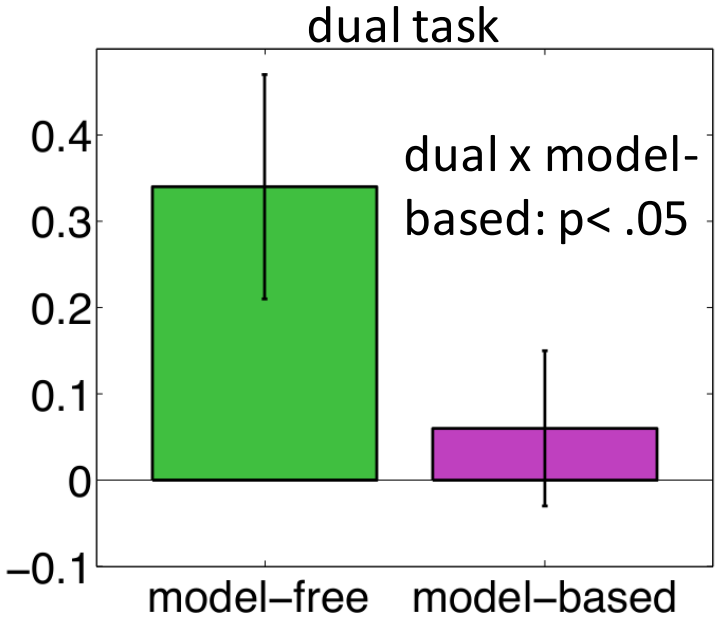



(Daw et al Neuron 2011)

What controls the tradeoff between these two sorts of learning?

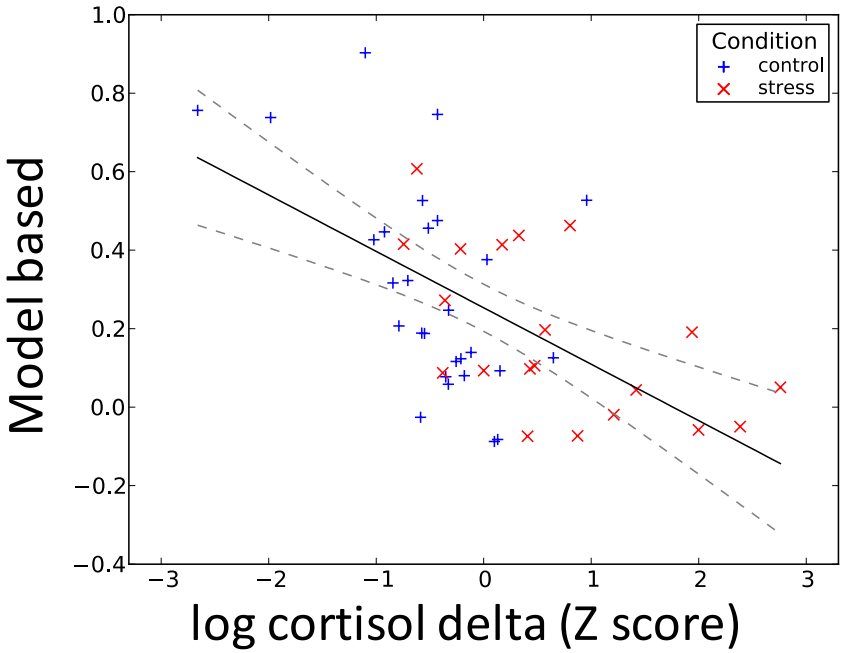
# interference




*(Otto et al Psych Science, 2013)*

# stress




*(Otto et al PNAS, 2013)*

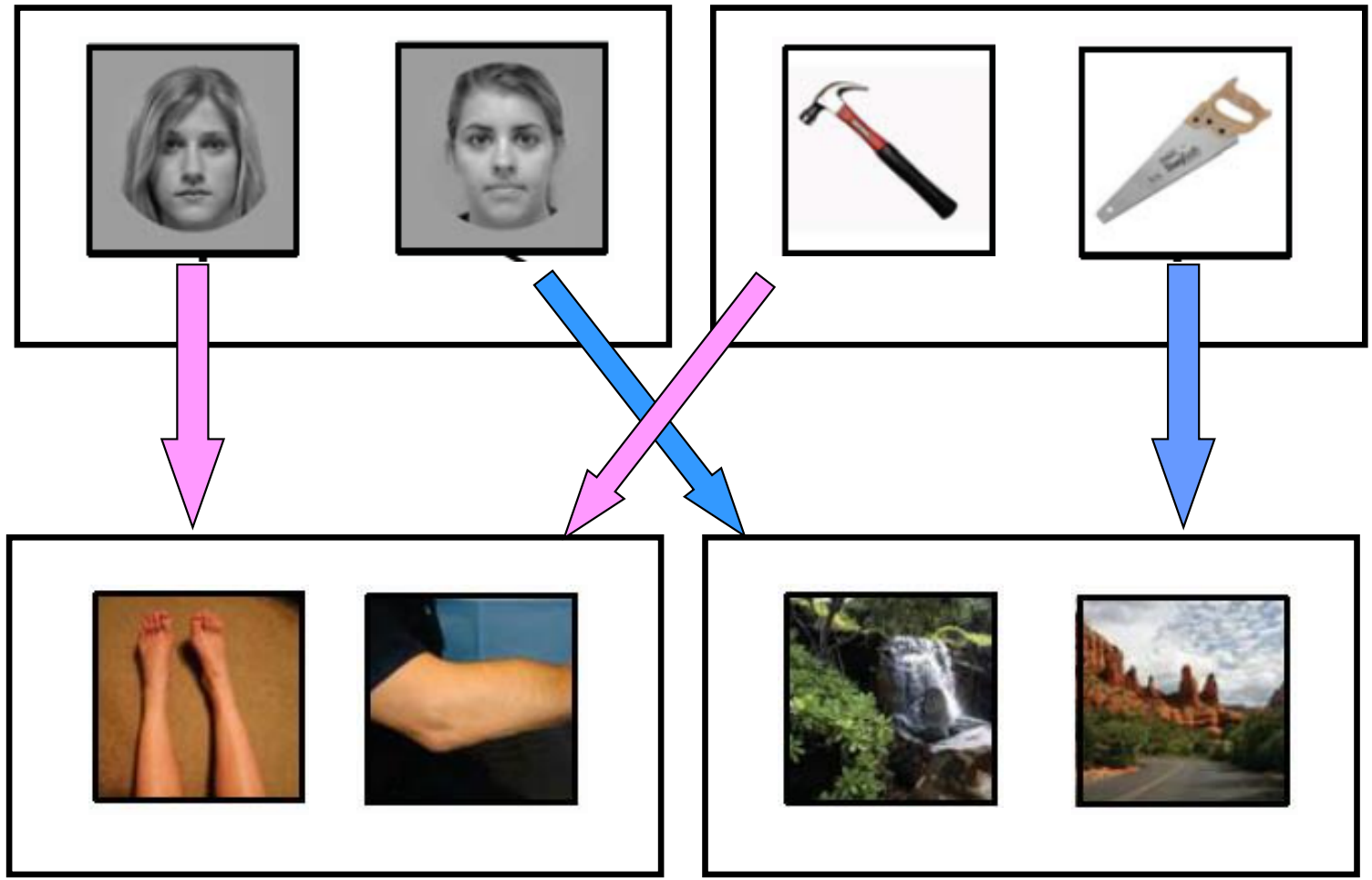
Also:

- Development & aging (Decker ea, 2016; Eppinger ea 2013)
- IQ (Schad ea 2014; Gillan ea 2016)
- cognitive control (Otto ea 2015)
- PFC TMS (Smittenaar ea 2013)
- COMT (PFC DA) genotype (Doll ea 2016)
- Parkinson's disease & meds (Sharp ea 2016; Wunderlich ea 2012)
- dopamine PET (Desserno ea 2015)
- psychopathology (Voon ea 2014, Gillan ea 2016)

what are the neural mechanisms underlying this evaluation?

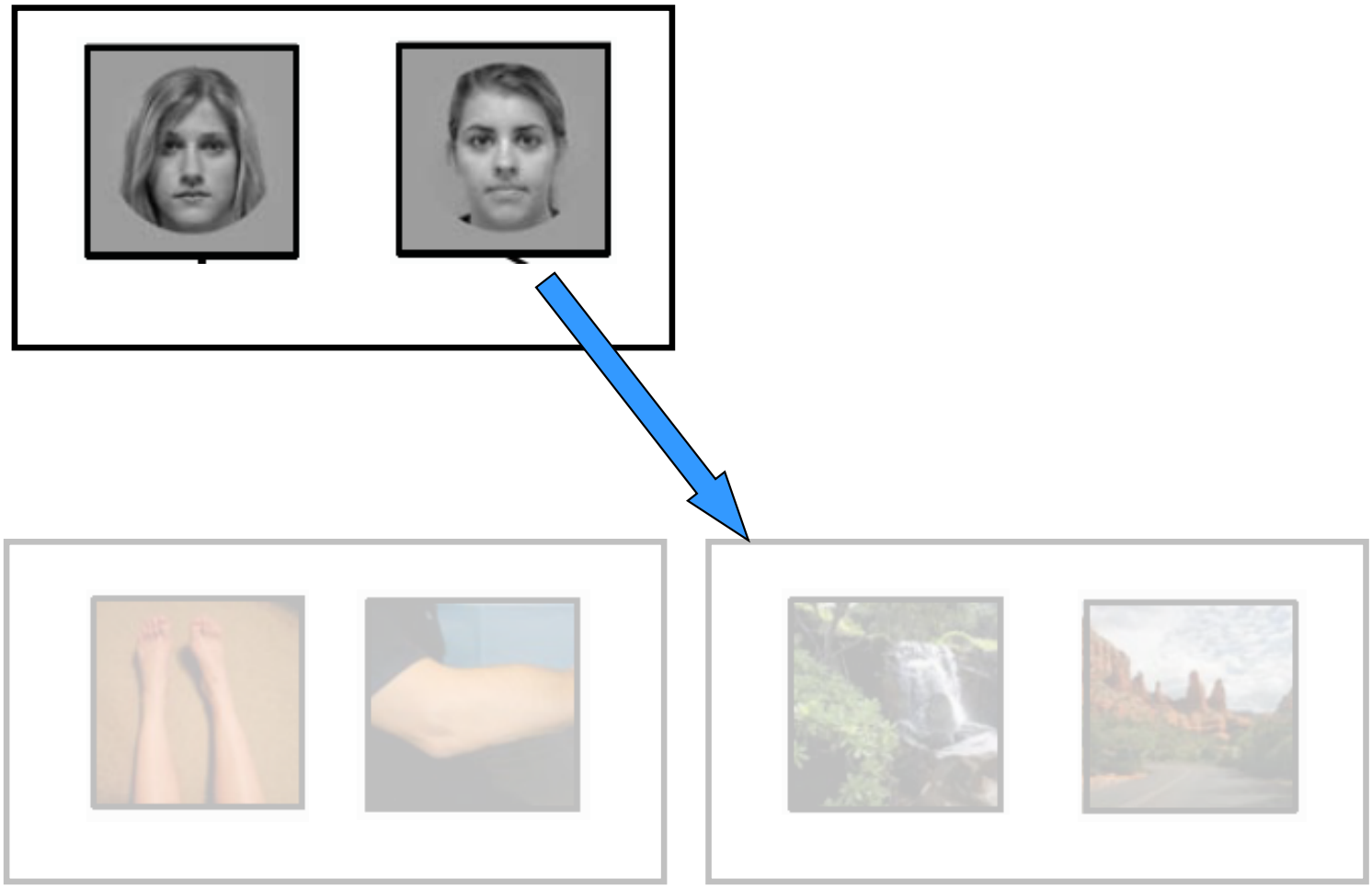
Is model-based learning really decision by simulation?

# decodable stimuli



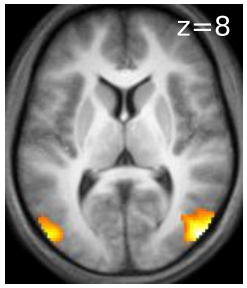
(Doll, Duncan, Simon, Shohamy & Daw *Nature Neuroscience* 2015)

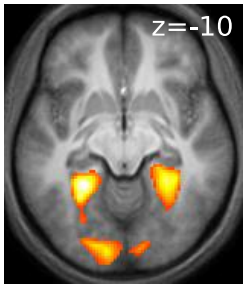
# catch trials



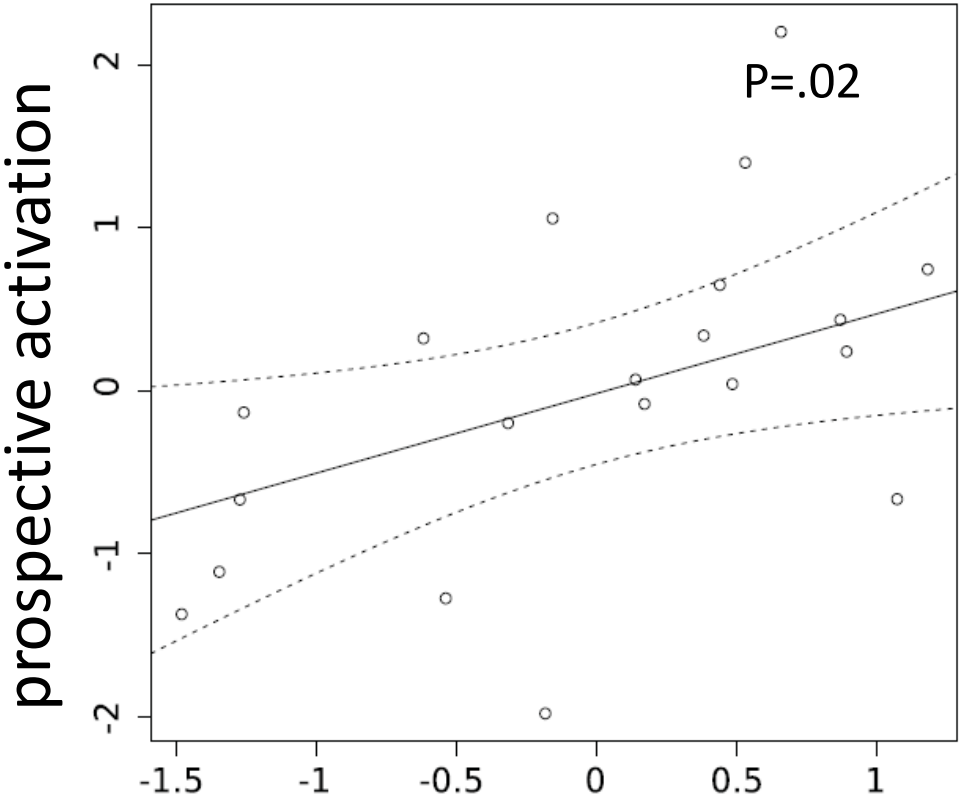
(Doll, Duncan, Simon, Shohamy & Daw *Nature Neuroscience* 2015)

## prospection (category selective ctx)



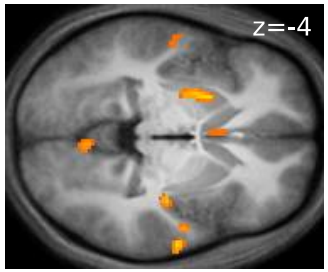



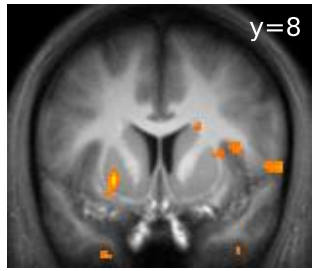

body parts scenes

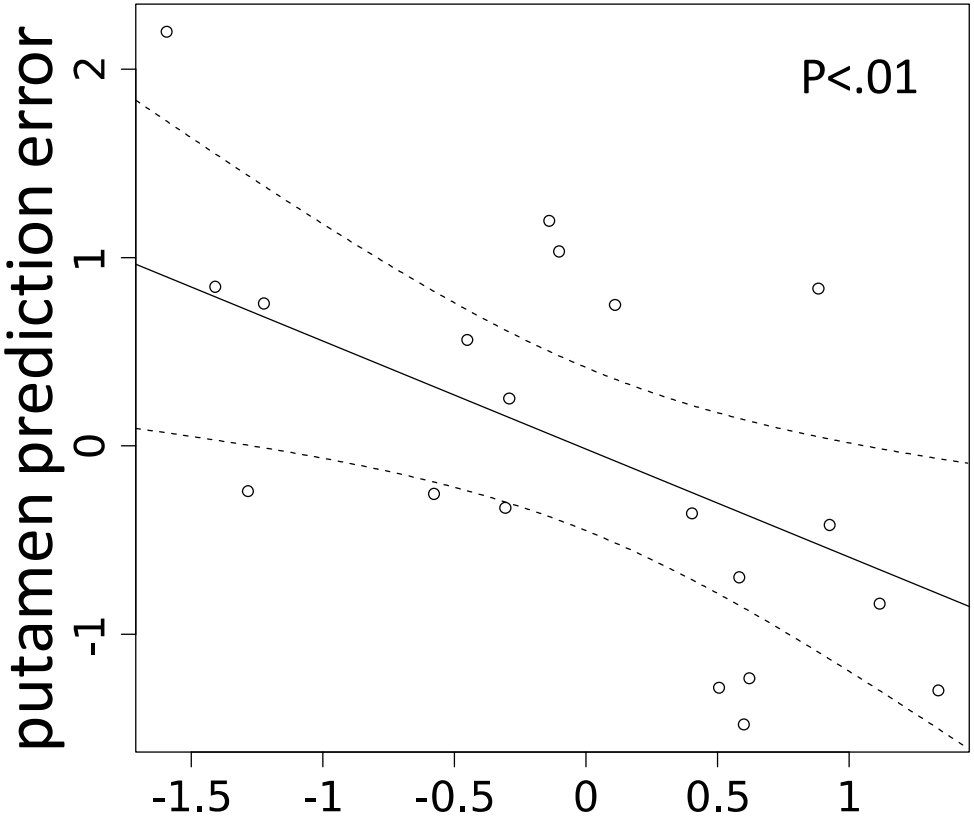



← MF behavior MB →

## RPE (ventral putamen)




← MF behavior MB →

(Doll, Duncan, Simon, Shohamy & Daw *Nature Neuroscience* 2015)

Signatures of two dissociable neural evaluation mechanisms

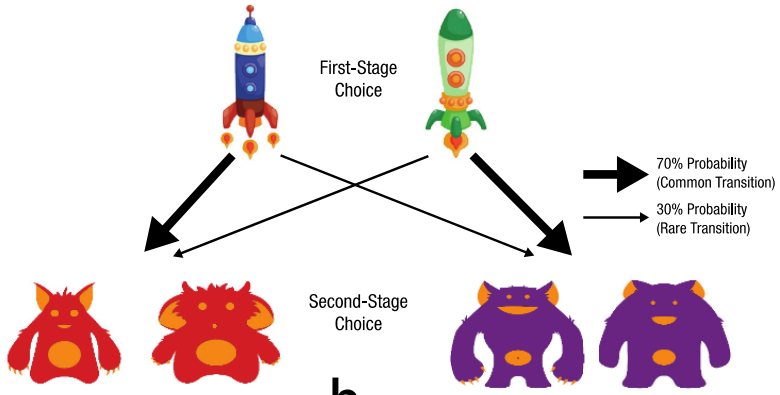
1. forward search
2. error-driven updating

which have the expected relationships to choice behavior

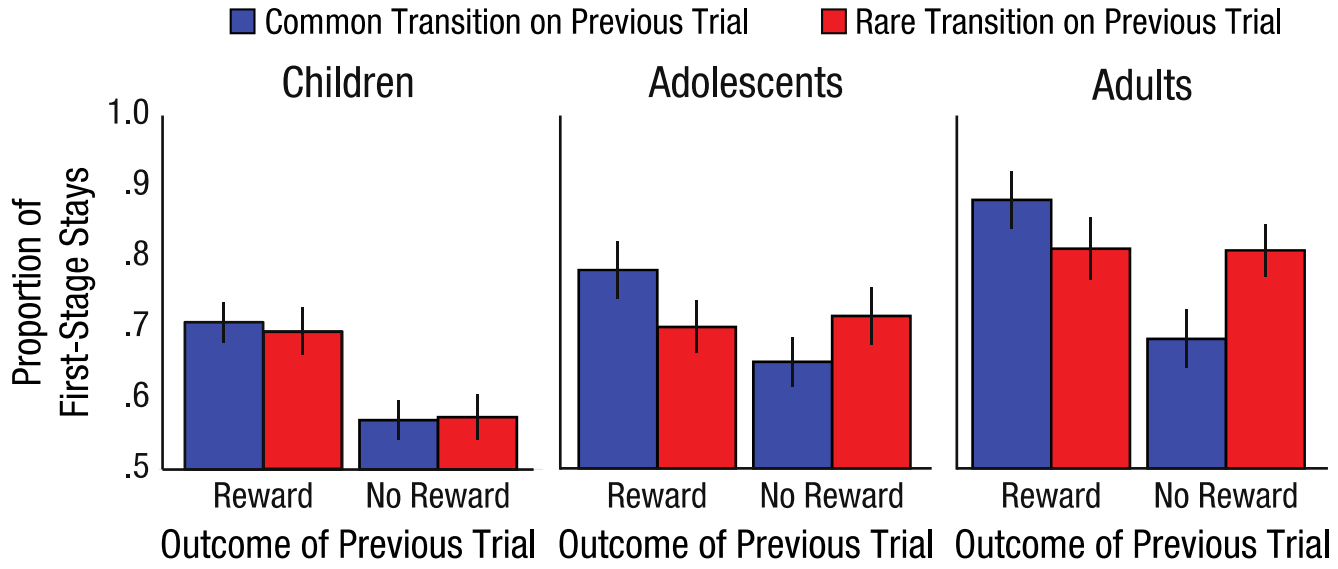
→ is this really related to compulsion?

# how do these change over development?

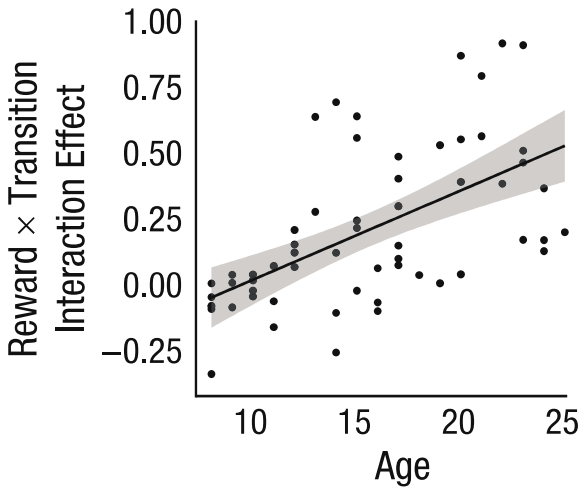
late maturation of PFC, hippocampus
impulsive behavior in youth




a




b

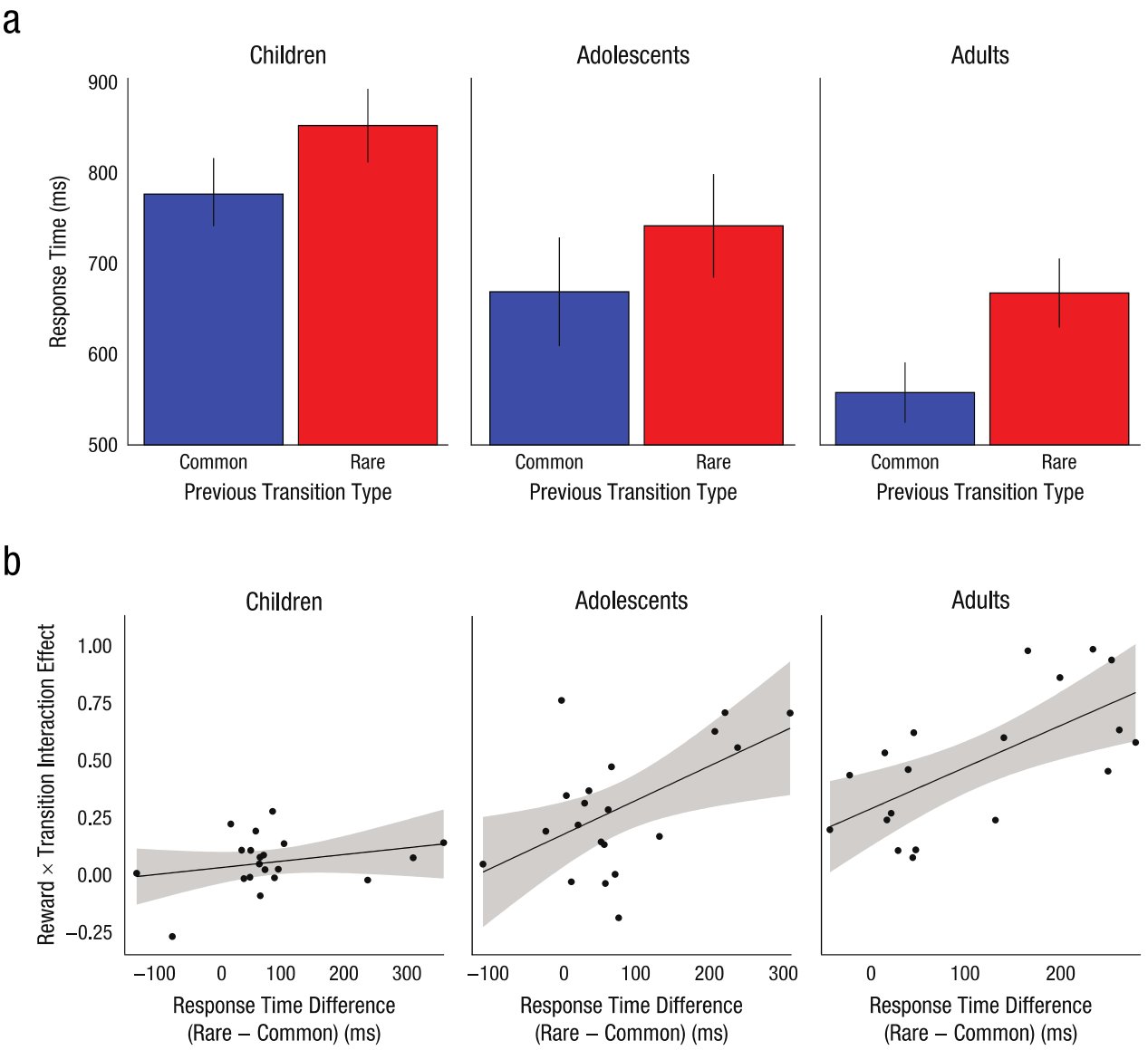



see also Eppinger et al. (2013) on aging

(Decker, Otto, Daw & Hartley, Psych Science 2016)

# not a failure to build model

explicit ("which planet did the blue rocket ship usually visit?") & implicit (RT) effects of transition model

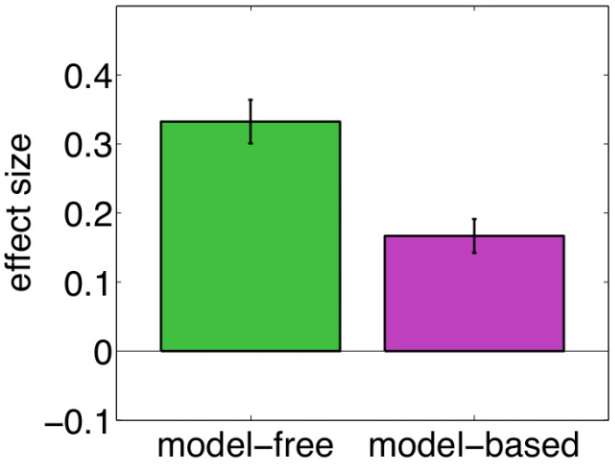



(Decker, Otto, Daw & Hartley, Psych Science 2016)

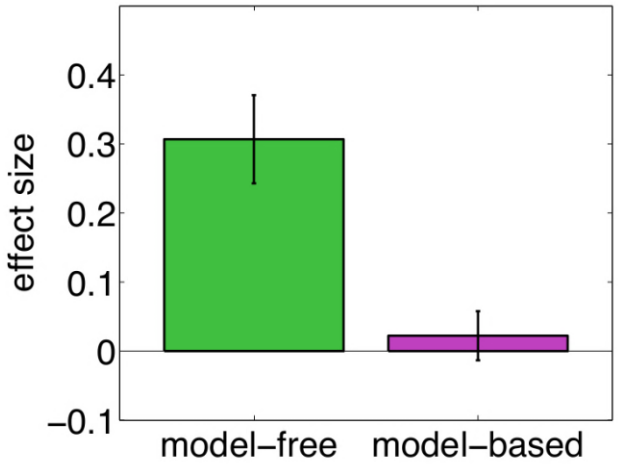
# Is model-based learning related to disorders of compulsion?

Claire Gillan

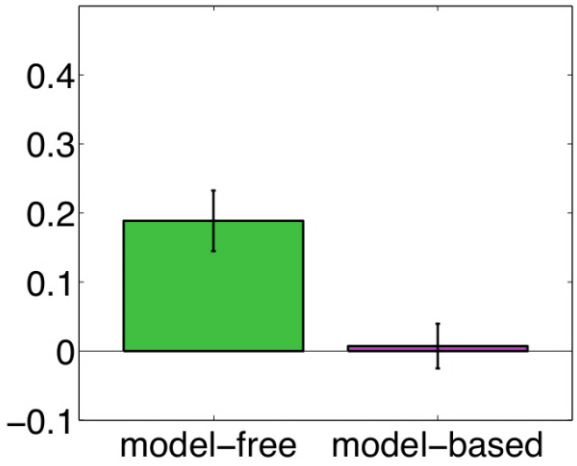
Healthy volunteers, n=106



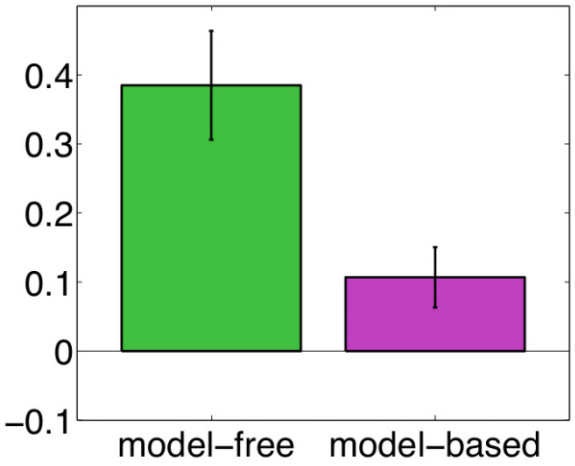

Binge eating disorder, n=30




Stimulant abusers, n=36




OCD, n=35




Methamphetamine/cocaine
Abstinent at least 1 wk

(Voon et al., Biological Psychiatry, 2014)

# however...

OPEN ACCESS Freely available online

PLOS ONE

## Impairments in Goal-Directed Actions Predict Treatment Response to Cognitive-Behavioral Therapy in Social Anxiety Disorder

**Gail A. Alvares, Bernard W. Balleine, Adam J. Guastella***

Brain & Mind Research Institute, The University of Sydney, Sydney, New South Wales, Australia

Biological Psychiatry

**Archival Report**

## Corticostriatal Control of Goal-Directed Action Is Impaired in Schizophrenia

Richard W. Morris, Stephanie Quail, Kristi R. Griffiths, Melissa J. Green, and Bernard W. Balleine

Journal of Abnormal Psychology
2016, Vol. 125, No. 6, 777–787


0021-843X/16/$12.00 http://dx.doi.org/10.1037/abn0000164

## Reduced Model-Based Decision-Making in Schizophrenia

Adam J. Culbreth and Andrew Westbrook
Washington University in Saint Louis

Nathaniel D. Daw and Matthew Botvinick
Princeton University

Deanna M. Barch
Washington University in Saint Louis

# the crisis in psychiatry

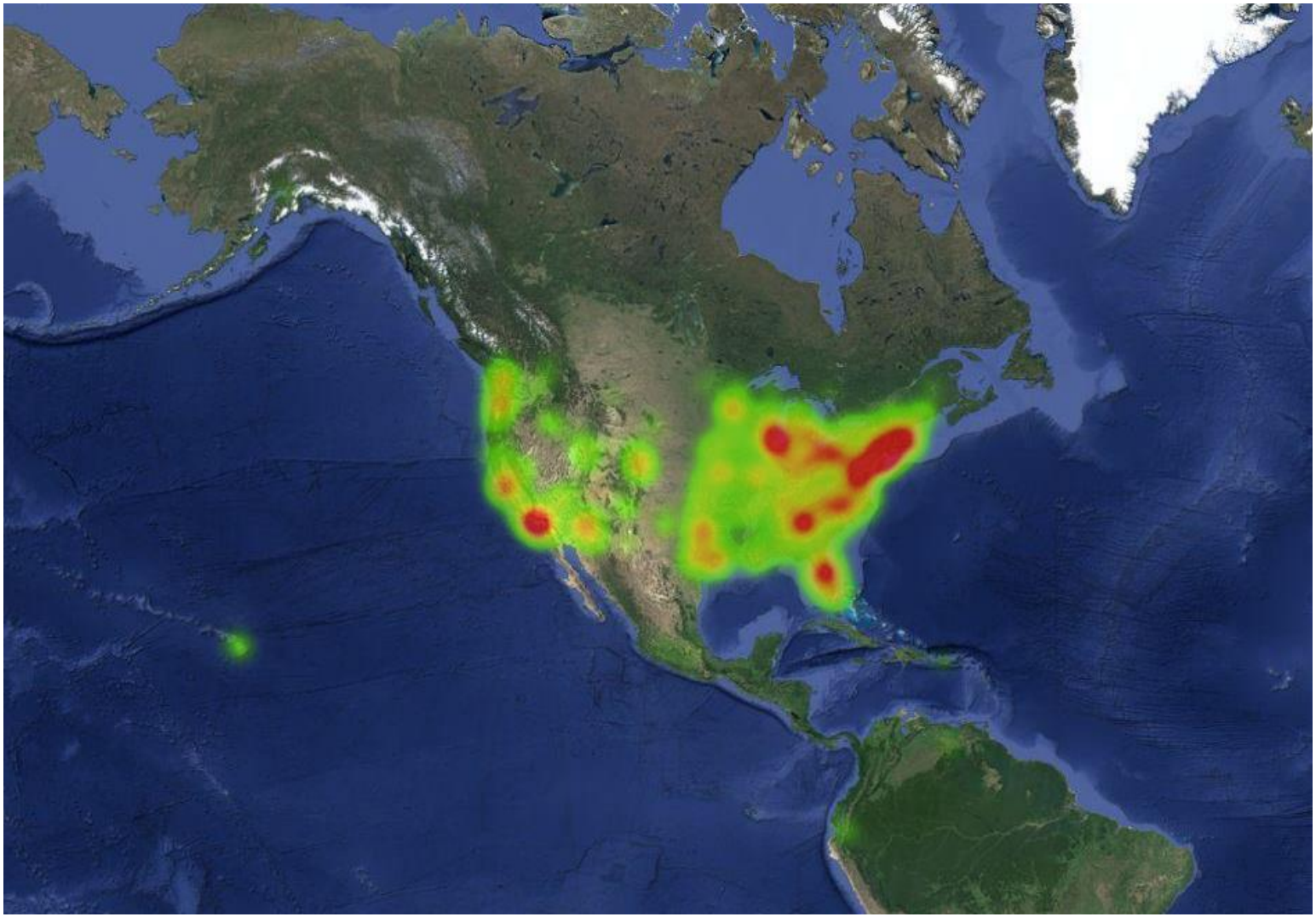
this may reflect a more general problem with psychiatric research – and psychiatric diagnoses

- co-morbidity, heterogeneity
- push toward dimensional, symptom-based view
- hope this will clarify etiology, neural basis

in a general population sample, look for evidence that this relationship is:

- graded/dimensional
- generalizes across diagnoses (“transdiagnostic”)
- yet is also specific to compulsive aspect

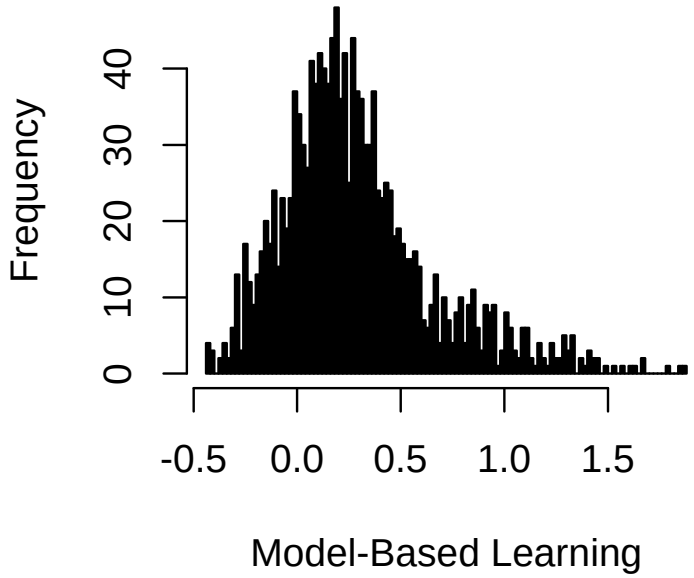
# Large-scale online testing



Amazon Mechanical Turk

# Experiment

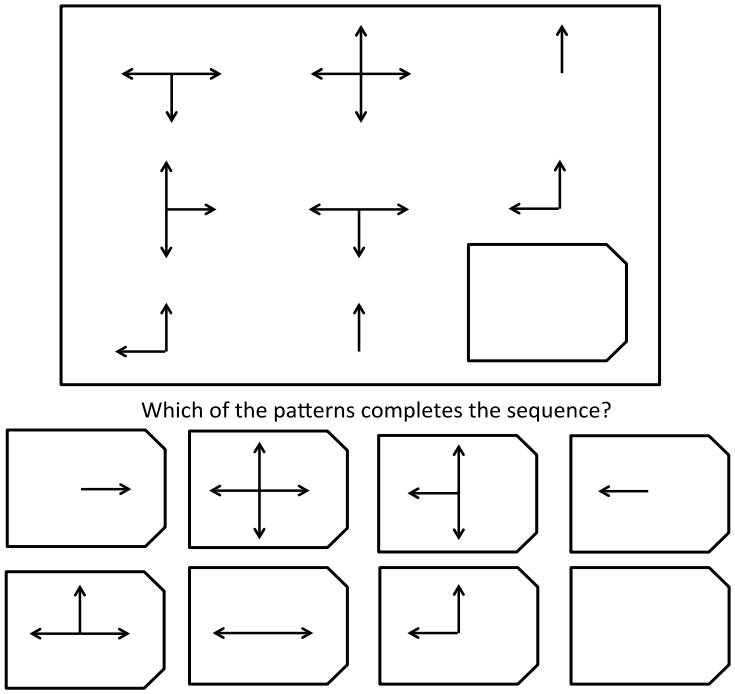
## 1. Model-based learning task




## 2. Self-Report Clinical Scales

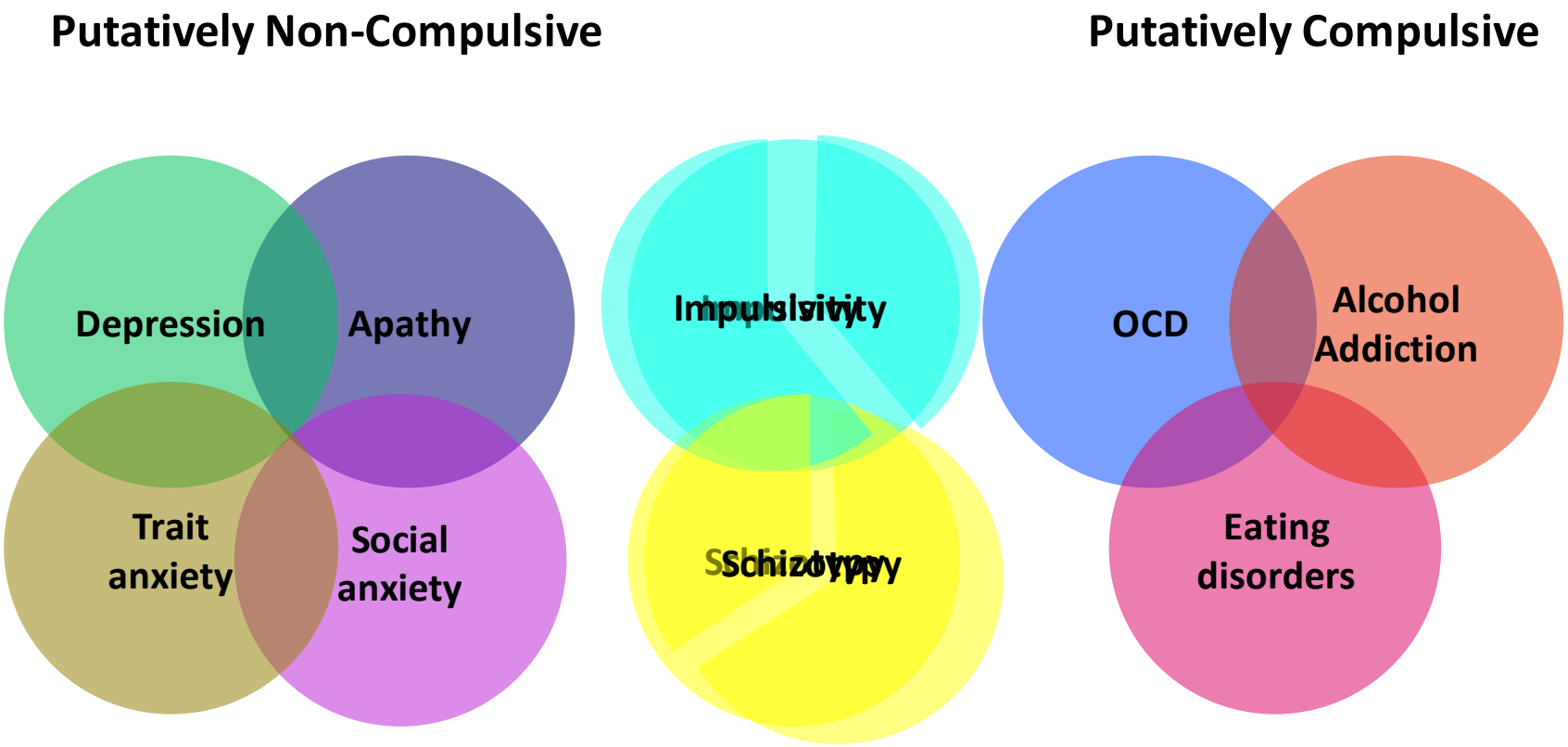
**OCD:** OCI-r (foa et al, 2002)
**Depression:** SDS (Zung, 1965)
**Anxiety:** STAI-trait (Spielberger, 1983)
…

## 3. IQ, age and gender



N=1413

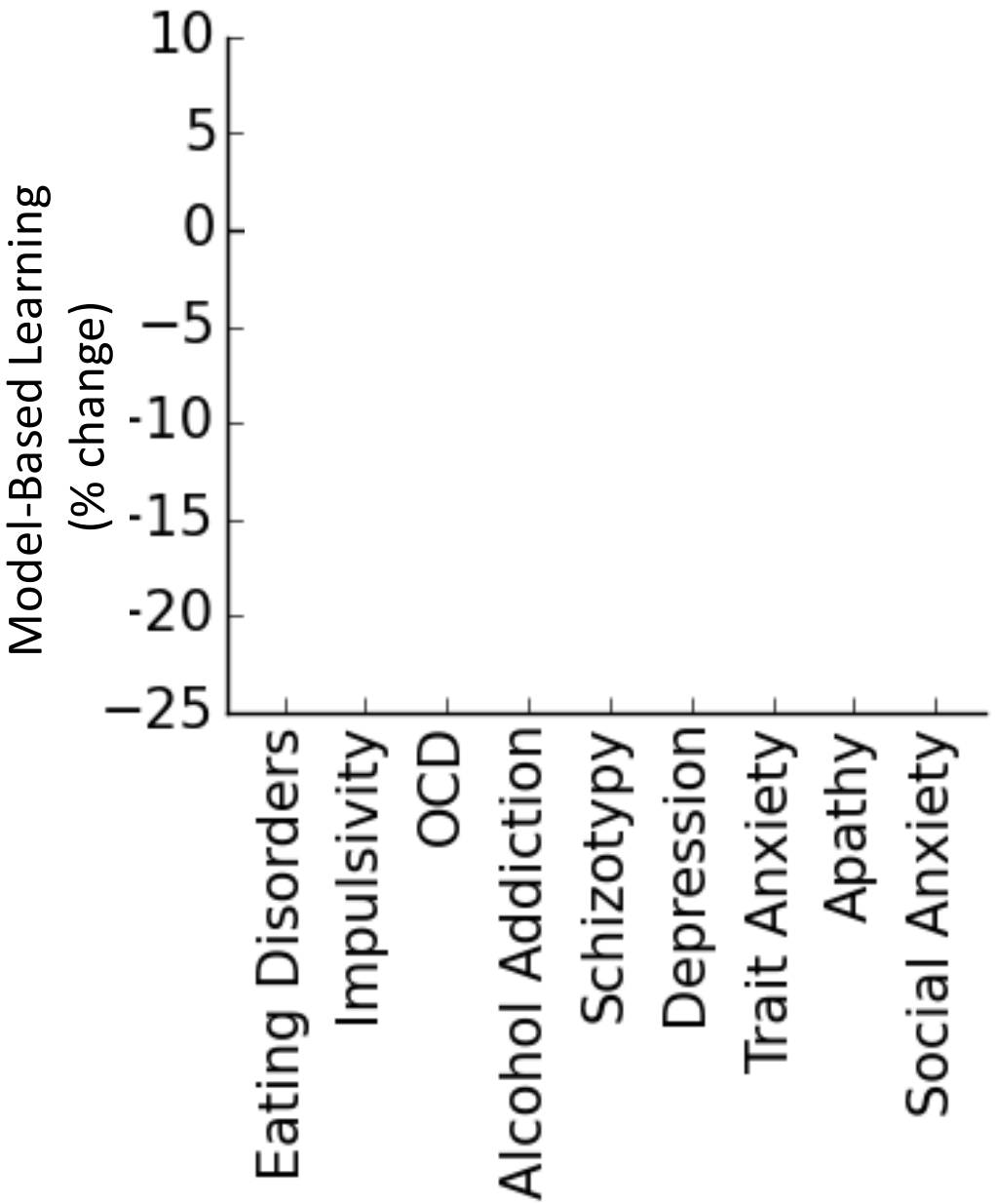
# Measures




(Gillan, Kosinski, Whelan, Phelps & Daw, eLife 2016)

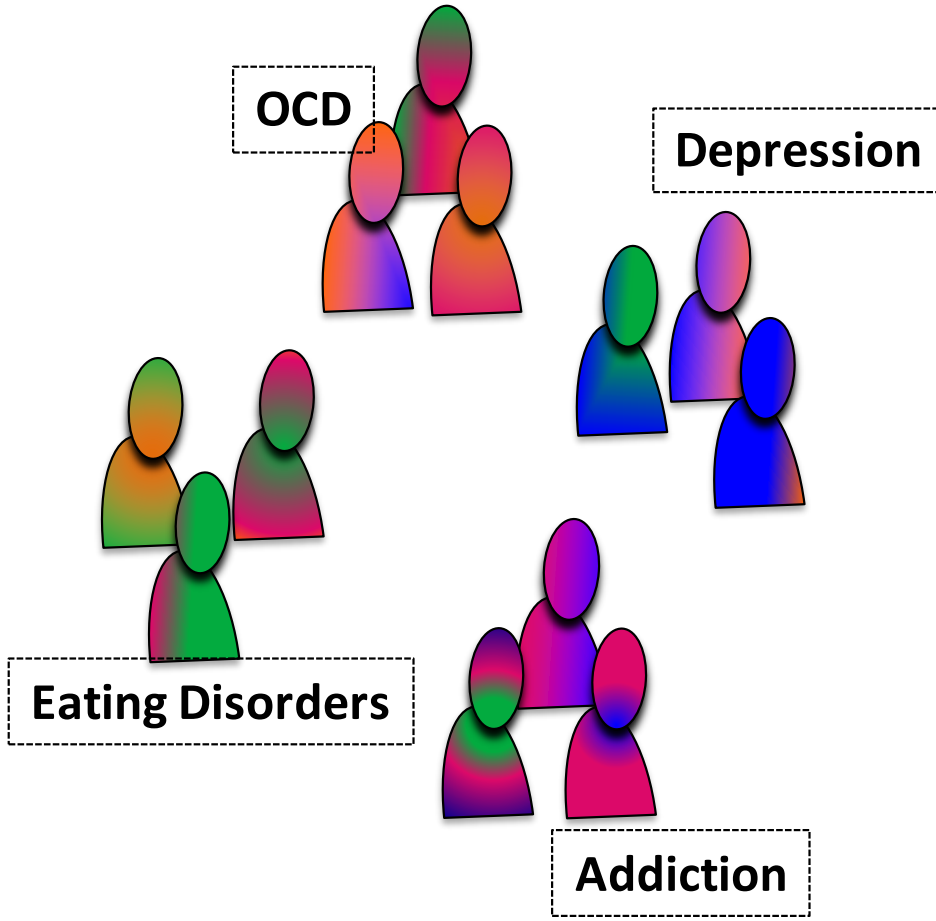
N=1413

# Experiment 2




*$p$<.05 ** $p$<.01 ***$p$<.001

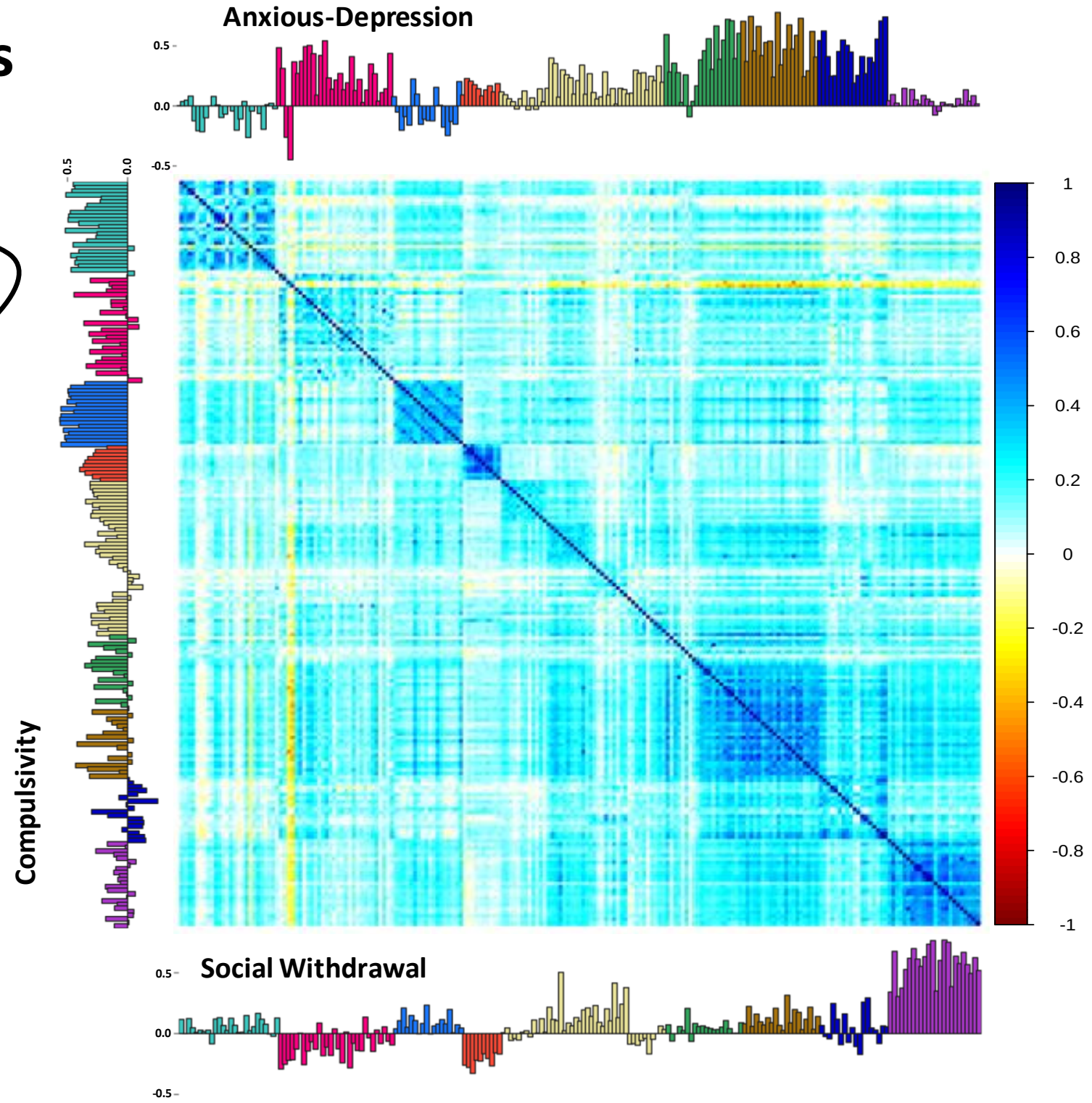
**But… this just illustrates the categorization problem**




(Gillan, Kosinski, Whelan, Phelps & Daw, eLife 2016)

# Factor Analysis

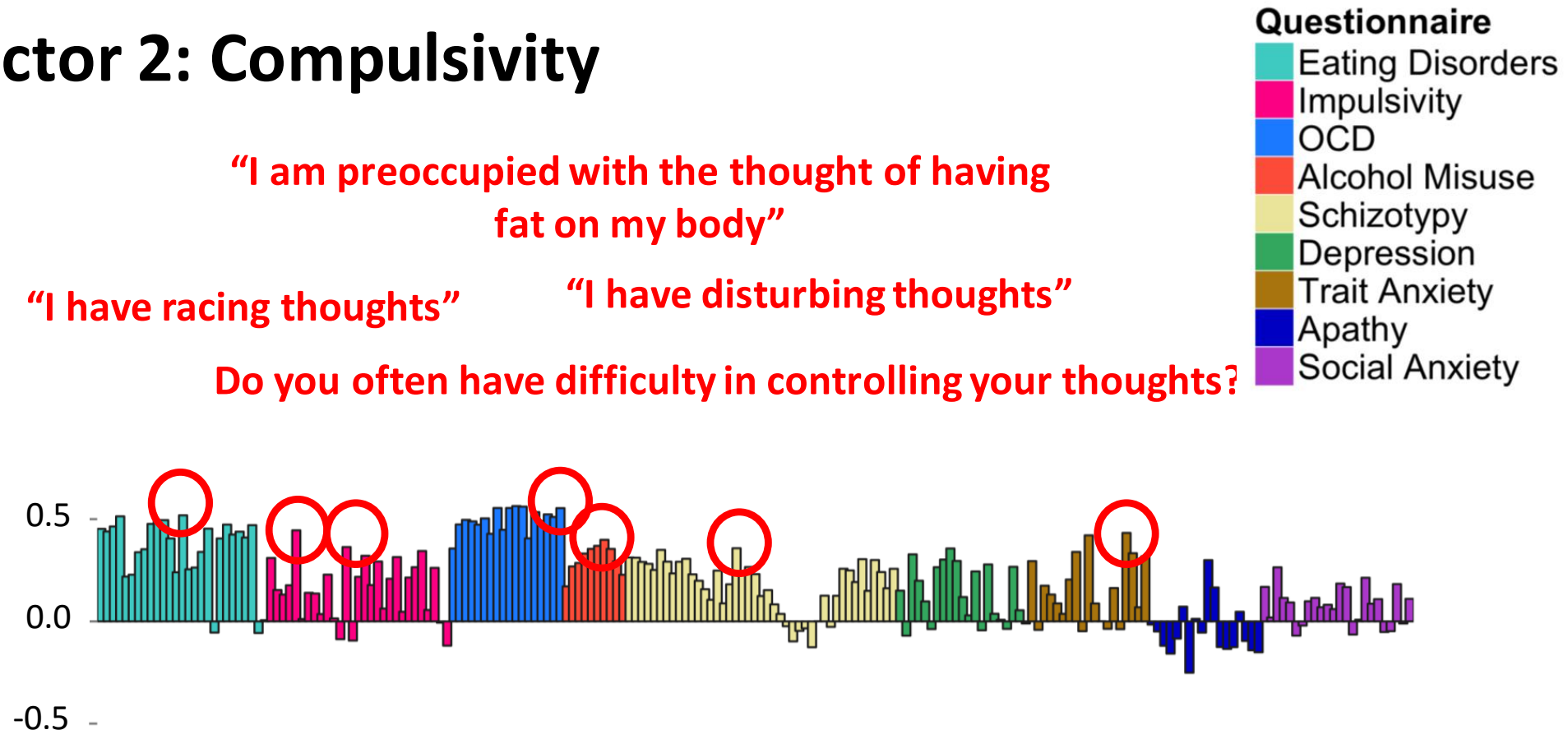
Inter-correlation of 209 individual self-report items




**Questionnaire**

- Eating Disorders
- Impulsivity
- OCD
- Alcohol Misuse
- Schizotypy
- Depression
- Trait Anxiety
- Apathy
- Social Anxiety

# Factor 2: Compulsivity




"I am preoccupied with the thought of having fat on my body"

"I have racing thoughts"

"I have disturbing thoughts"

Do you often have difficulty in controlling your thoughts?

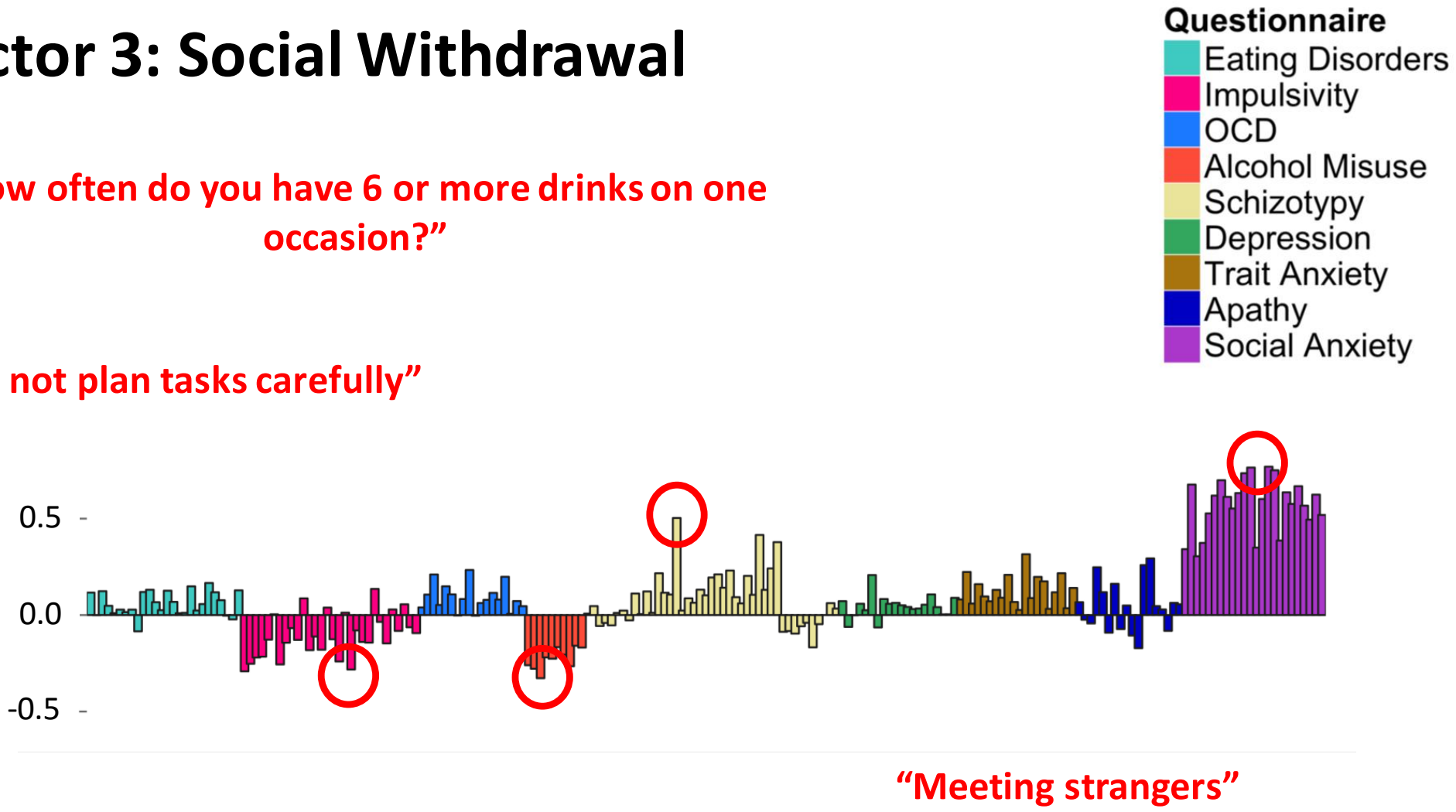
"I repeatedly check doors, windows, drawers, etc."

"I have gone on eating binges where I feel that I may not be able to stop"

"I buy things on impulse"

"How often … have you needed a first drink in the morning to get yourself going …?"

# Factor 3: Social Withdrawal




"How often do you have 6 or more drinks on one occasion?"

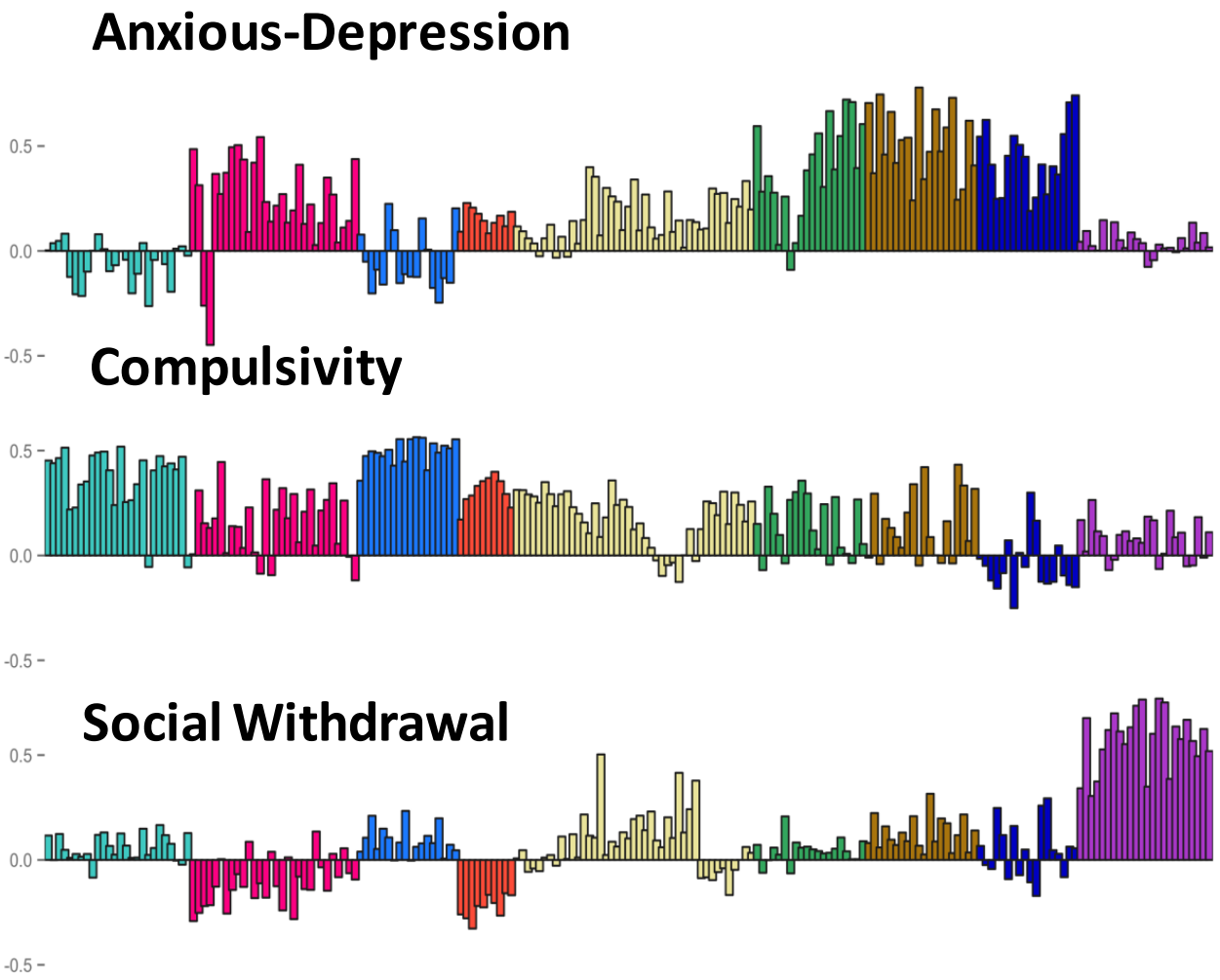
"I do not plan tasks carefully"

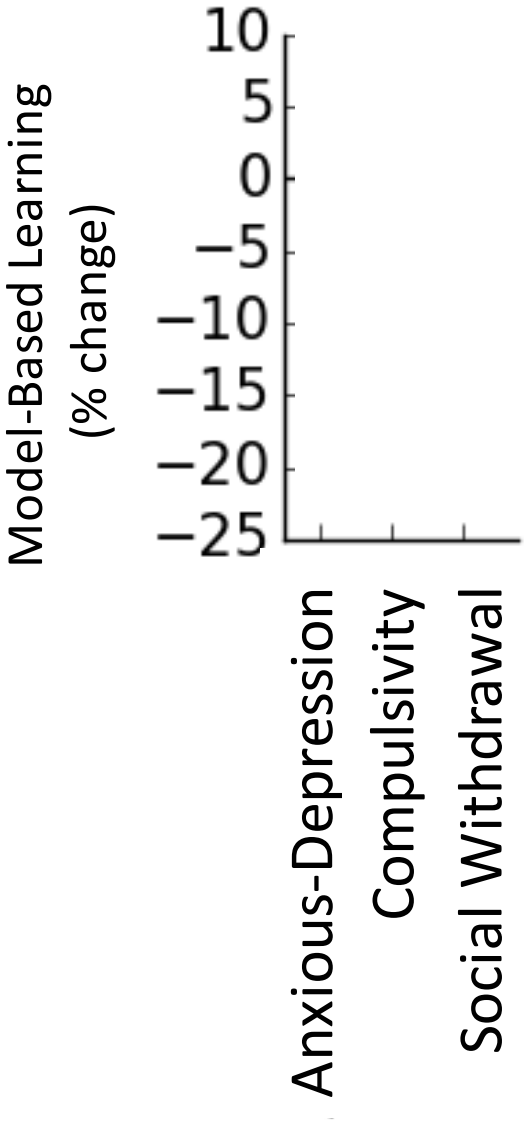
"Meeting strangers"

"Being the center of attention"

Do you dread going into a room when other people have gathered and are talking?

*$p$<.05 ** $p$<.01 ***$p$<.001




(Gillan, Kosinski, Whelan, Phelps & Daw, eLife 2016)

MB learning is selectively linked to compulsion, across diagnoses

- similar results from fully supervised, item-level analysis
- also effects of age, IQ

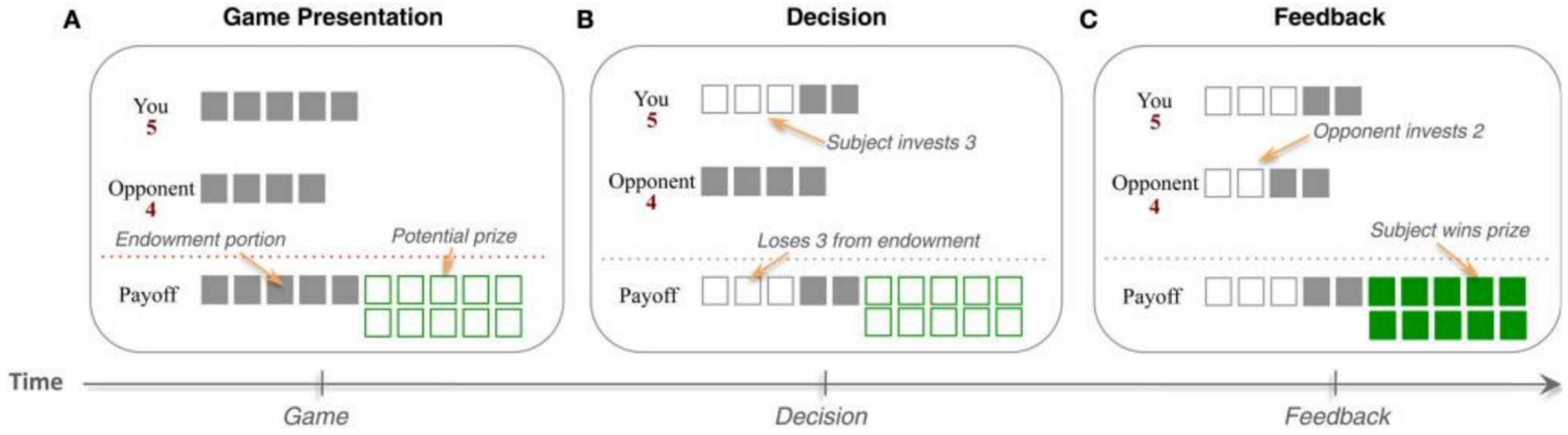
Compulsive thoughts and behaviors cluster in factor analysis

- relevant to obsessions vs. compulsions?

Of course these are just some symptom scales, and just one behavioral task

- progressively refine both sides
- promise of large-scale online testing more broadly

# patent race game

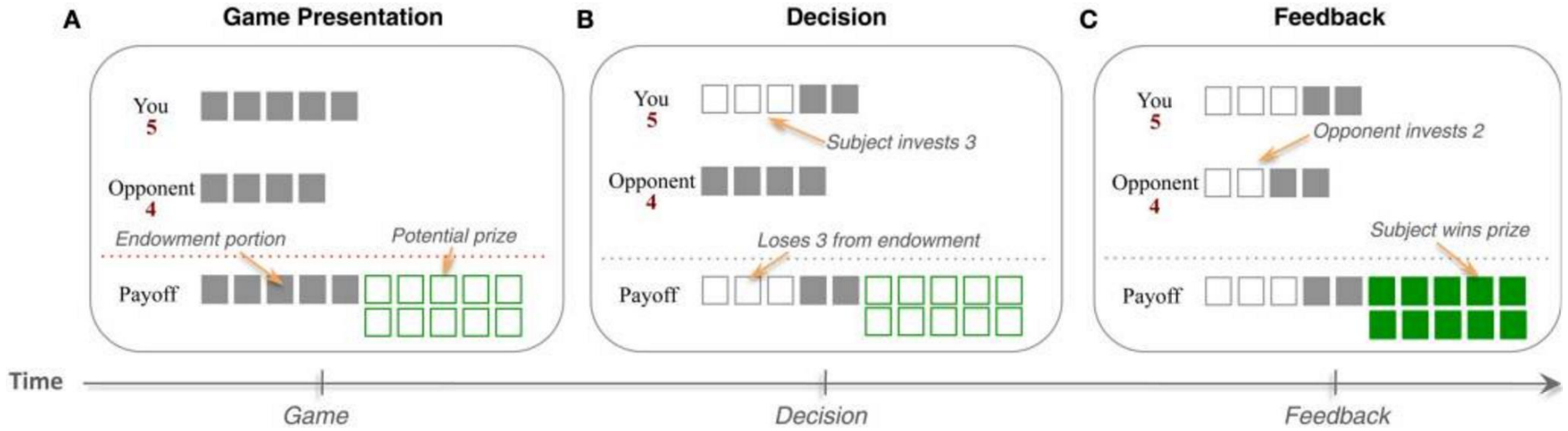



invest a portion of endowment, win prize if you invest more than opponent

- repeated play (80 trials) against replayed investments from previous subjects
- mixed strategy equilibrium
- learning (e.g. about opponents' move distribution, or which moves work)

Elana Meer; Lindsay Hunter; Ming Hsu

# patent race game




theory, EWA (Camerer & Ho, 1999) nests:

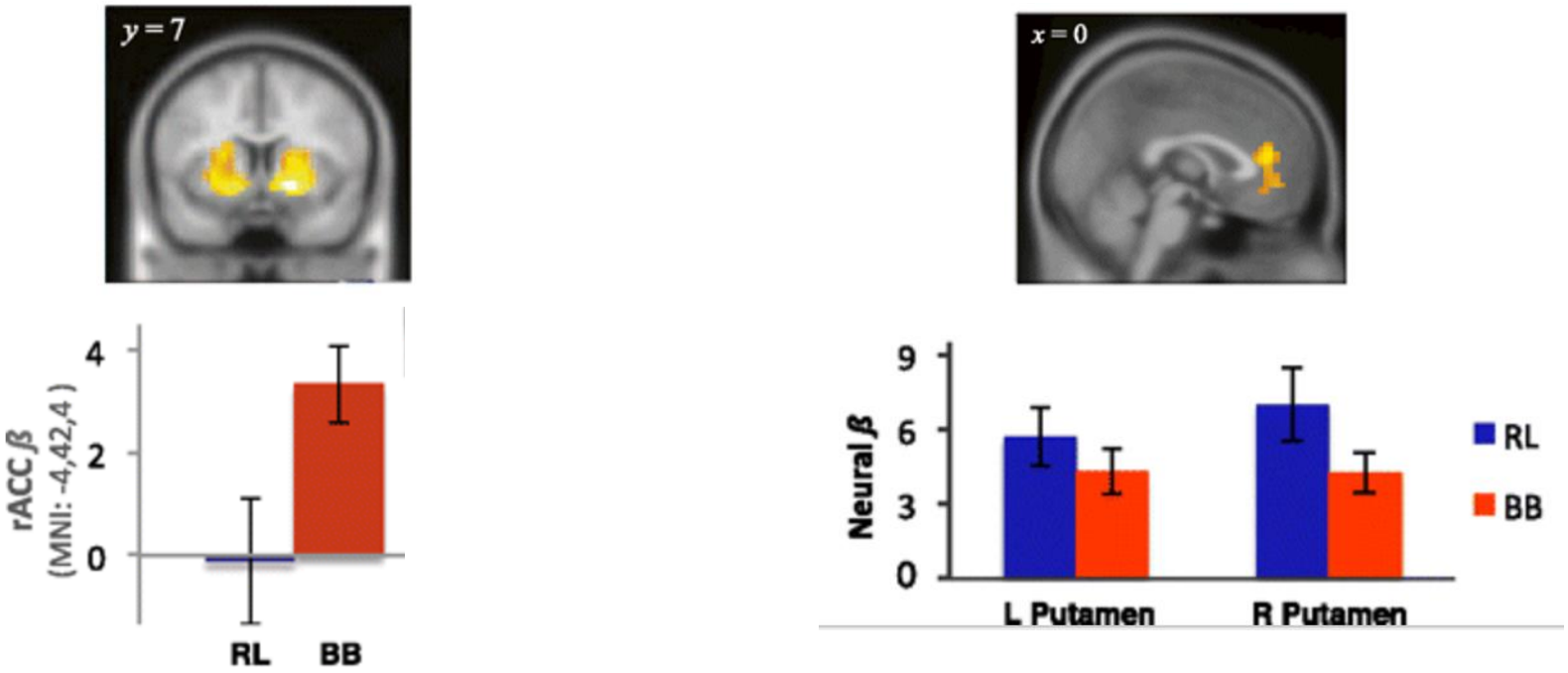
1. (model-free) "reinforcement learning", about reward received (or not) after actions
2. (model-based) "belief learning" about opponents' likely strategies, (& best-respond)

in this setting, (2) is algebraically equivalent to counterfactual learning about foregone rewards, governed by free parameter $\delta$:

$$Q_{t+1}(c_t) = \phi \cdot Q_t(c_t) + r(c_t) \qquad \textit{for chosen action}$$
$$Q_{t+1}(u_t) = \phi \cdot Q_t(u_t) + \delta \cdot r(u_t) \qquad \textit{for unchosen actions}$$

Elana Meer; Lindsay Hunter; Ming Hsu

# patent race game




(Zhu et al., PNAS 2012)

In a series of papers using EWA and games like this, Ming Hsu & colleagues (2012, 2014, 2015) have shown evidence for a similar two-system story as with MDPs

- fMRI dissociation between reward and belief learning (striatum, PFC)
- individual differences (striatal vs PFC dopamine genes, aging)

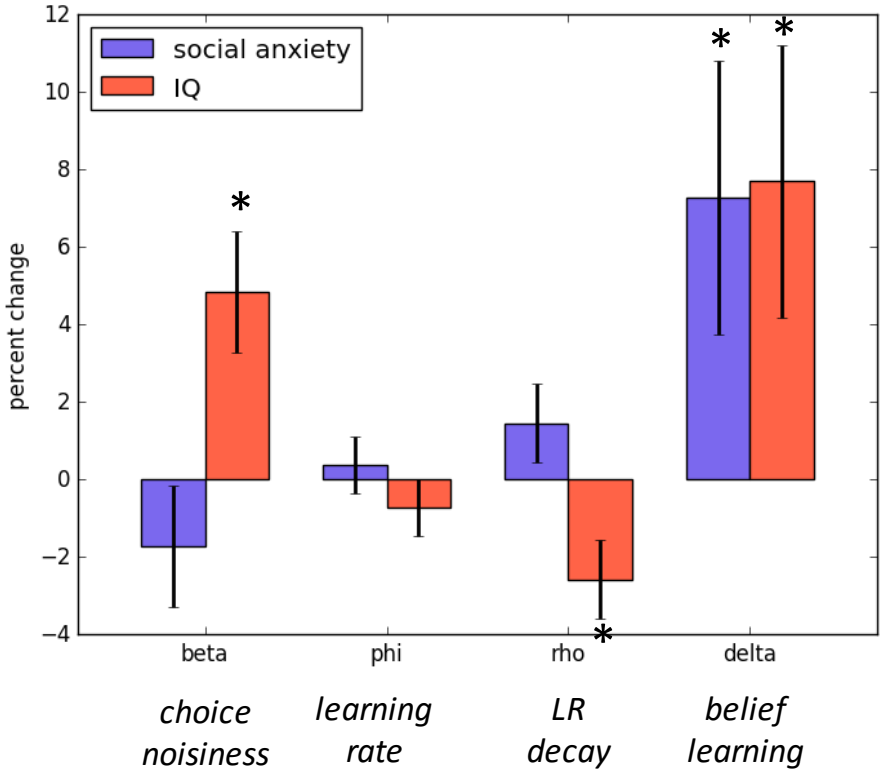
Due to social framing, this seemed like a strong candidate to follow up on social anxiety effects on model-based learning

# preliminary results

N=366, Turk sample

- social anxiety, IQ (ravens matrices), 80 trials of patent race
- fit EWA model

parameter $\delta$ (rel. strength of MB) increasing in anxiety (p<.05)




Elana Meer; Lindsay Hunter; Ming Hsu

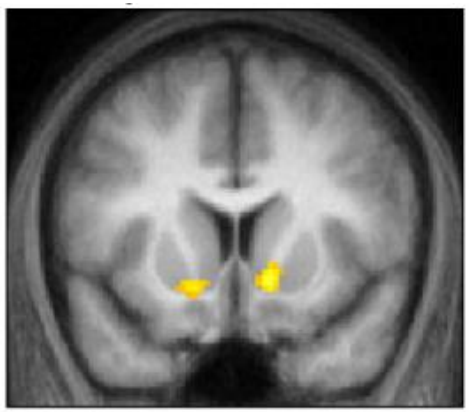
Question: how to account for the goal-directed nature of compulsion?

hybrids and MB/MF interactions

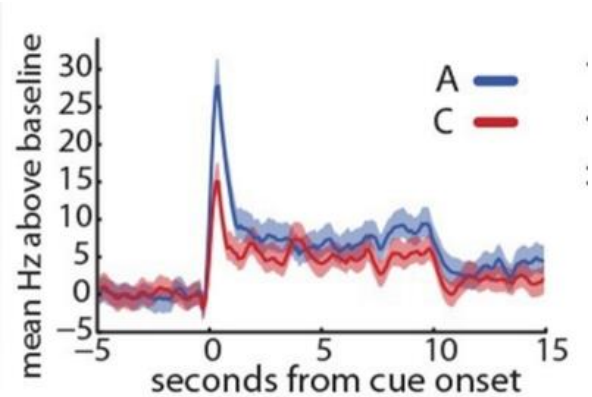
- Dyna & replay (Gershman et al. 2014)
- Pruning/truncation (Keramati & Dayan)
- Successor representation (Daw & Dayan 2015; Russek et al under review)
- MF goal selection (Cushman and Morris, 2015)
- MB as reoriented toward object of compulsion, rather than generally deficient (Voon et al. 2015)

# interactions

**MB valuations → Dopamine (& PEs)**

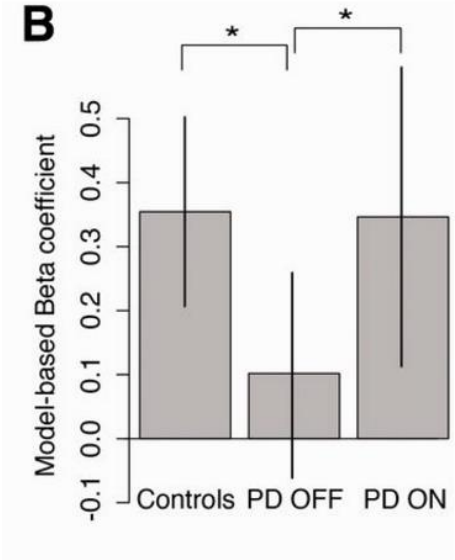


(Daw, Gershman et al, 2011)




(Sadacca et al., 2016)

**Dopamine → MB valuation**




- Parkinson's disease & meds (Sharp et al., 2016)
- COMT genotype (Doll et al., 2016)
- PET (Deserno et al., 2015)
- L-Dopa (Wunderlich et al., 2012)

# conclusions

1. distinguish two reinforcement learning computations in the human brain
   - linked with two distinct neural mechanisms
   - forward search vs error-driven updating
   - fills in detail behind important dual-system models

2. model-based learning is linked to compulsion (& tentatively, social anxiety)
   - generalizes across disorders but is specific to a subset
   - broad usefulness of large scale online testing in psychiatry

3. many future questions
   - can we understand neural mechanism for model-based computation in finer detail? (animals!)
   - how does interaction work? (important e.g. for drugs)
   - does this give us a handle on other dual-system phenomena and frameworks, e.g. self-control, time discounting?

**Lab:**
Bradley Doll
Claire Gillan
Ross Otto
Elana Meer
Lindsay Hunter
Evan Russek
Oliver Vikbladh

**Collaborators:**
Liz Phelps
Daphna Shohamy
Sam Gershman
Ming Hsu
Cate Hartley
Hugo Decker
Peter Dayan
Ray Dolan
Valerie Voon

**Funding:**
NIMH
NIDA
NINDS
McDonnell Foundation
Templeton Foundation
Google DeepMind