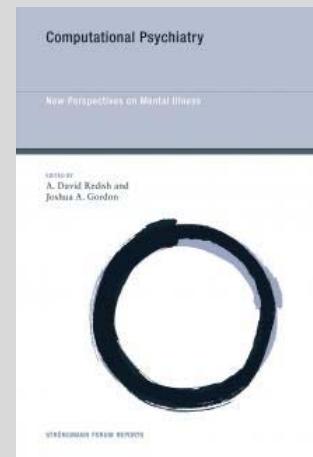
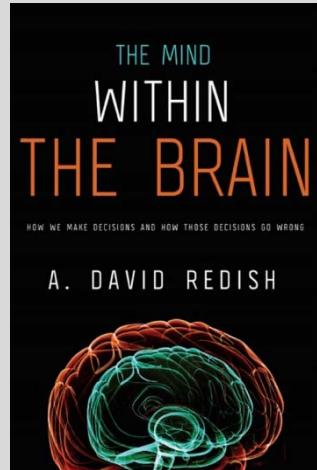
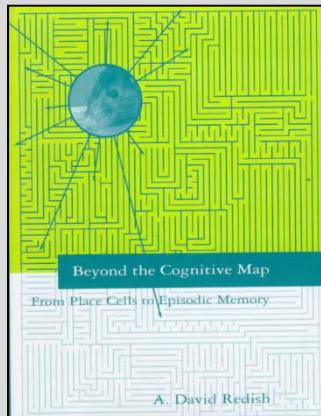


# Computational models of addiction

A. David Redish  
Department of Neuroscience  
University of Minnesota  
[redish@umn.edu](mailto:redish@umn.edu)  
<http://umn.edu/~redish>



RedishLab at the University of Minnesota  
<http://redishlab.neuroscience.umn.edu>



# Computational models of addiction

*assessing models and consequences*

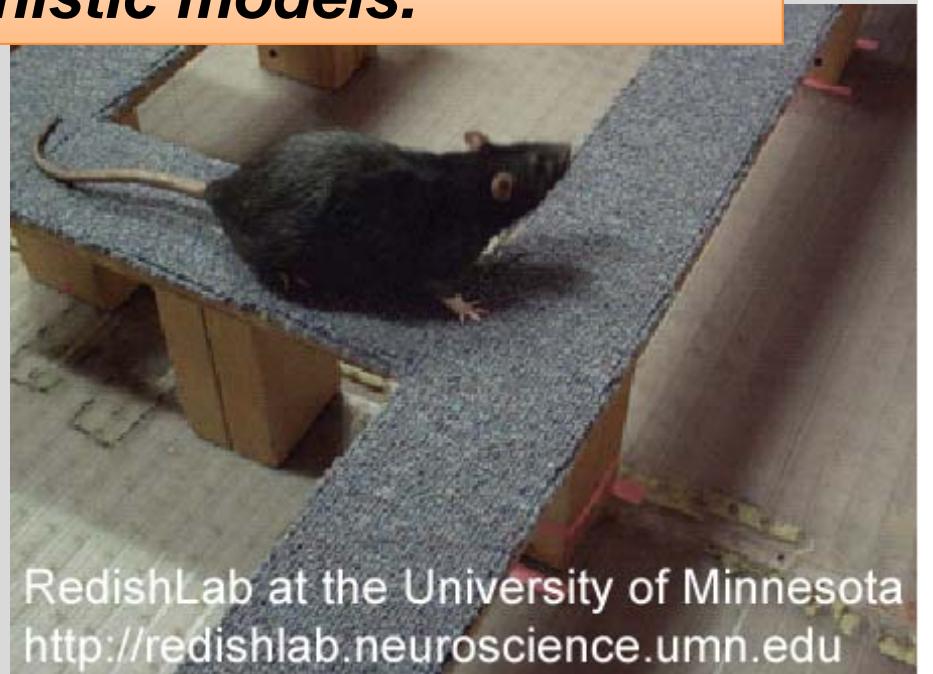
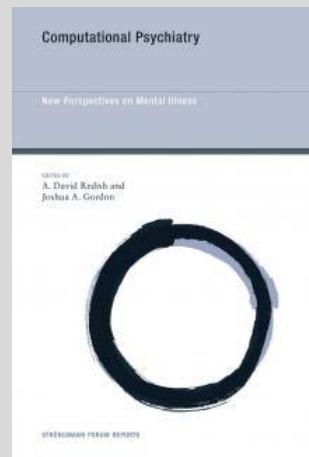
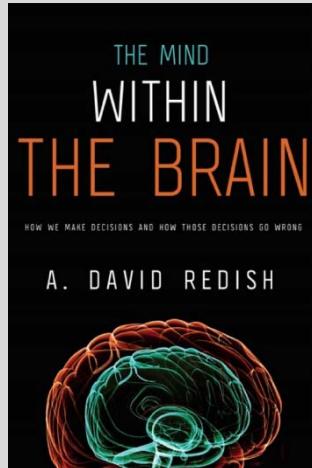
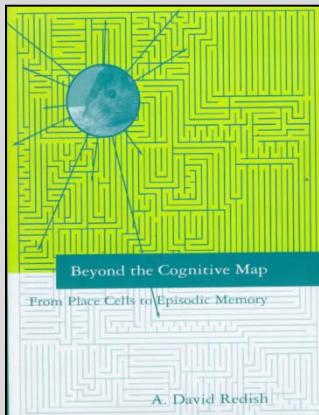
**We're going to be talking about  
mechanistic models.**

A. David Redish

Department of Neuroscience  
University of Minnesota

[redish@umn.edu](mailto:redish@umn.edu)

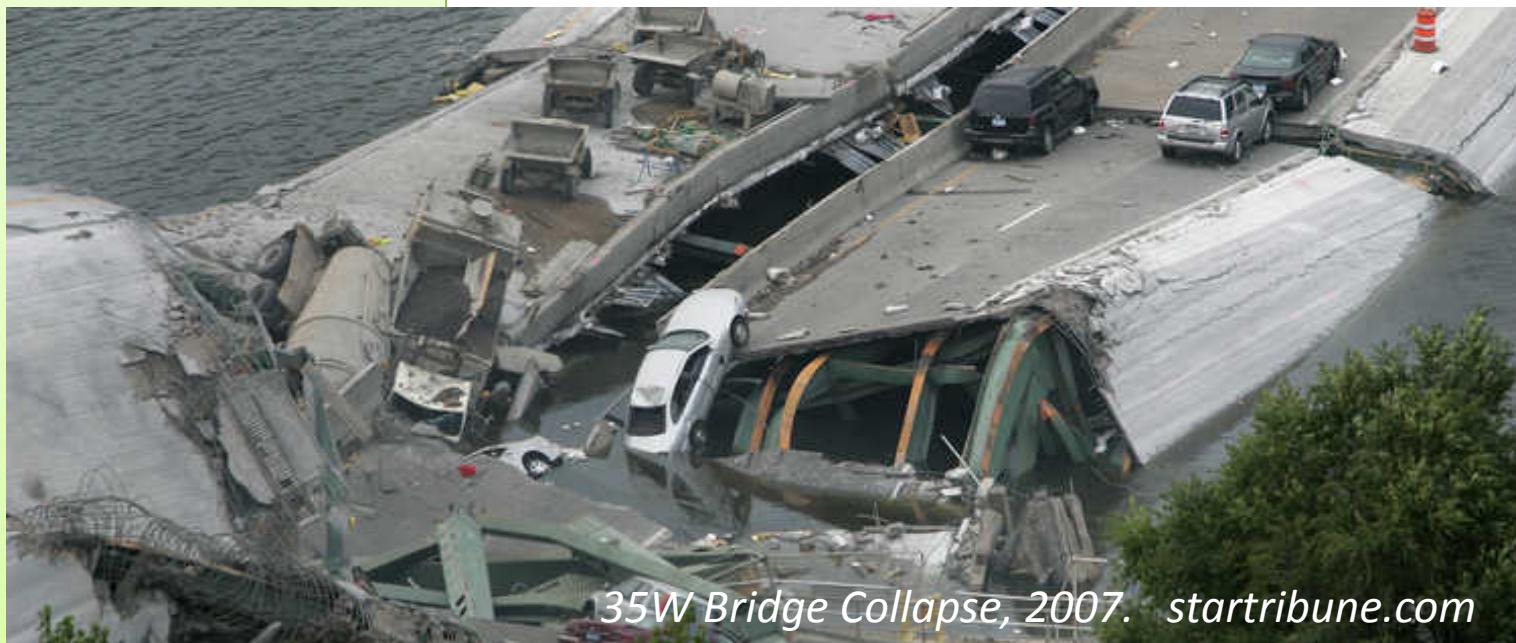
<http://umn.edu/~redish>



## *“Failure modes”*

### The concept of the “failure mode”

In engineering, a “failure mode” is a vulnerability inherent in the machinery.



*35W Bridge Collapse, 2007. startribune.com*

## The concept of the “failure mode”

In engineering, a “failure mode” is a vulnerability inherent in the machinery.

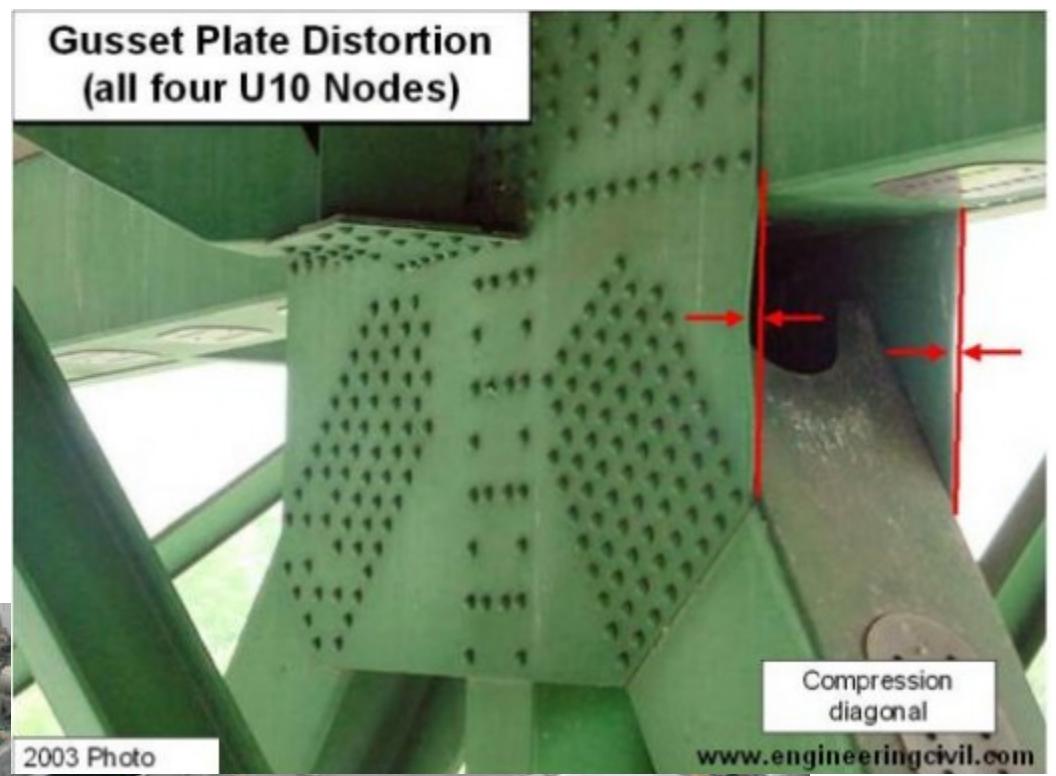


35W B



## “Failure modes”

### Gusset Plate Distortion (all four U10 Nodes)



# *A concrete example of a failure mode*

REPORTS

## Addiction as a Computational Process Gone Awry

A. David Redish

10 DECEMBER 2004 VOL 306 SCIENCE www.sciencemag.org

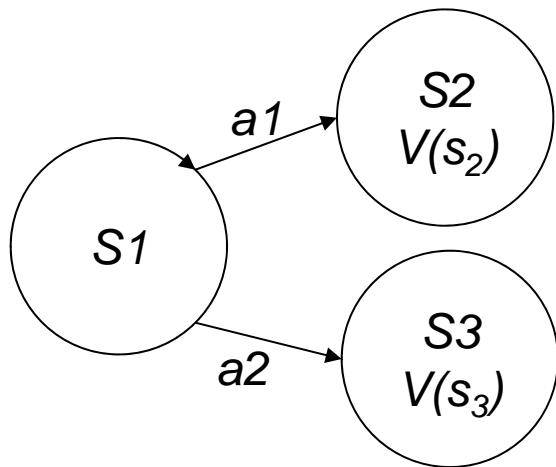
## The delta signal

Given a choice, if we learn the value of our choices, we can make decisions based on those learned values.

### “Failure modes”

Define “value” as the total expected (discounted) reward.

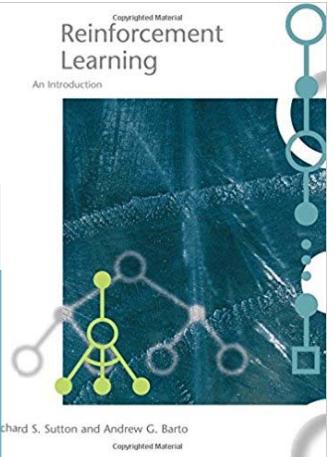
$$V(t) = \int_t^{\infty} \text{disc}(E[R(\tau)], \tau - t) d\tau$$



We can learn the value from the **value prediction error** signal  $\delta$ .

$$\delta(t) = \gamma^d(R(S_l) + V(S_l)) - V(S_k)$$

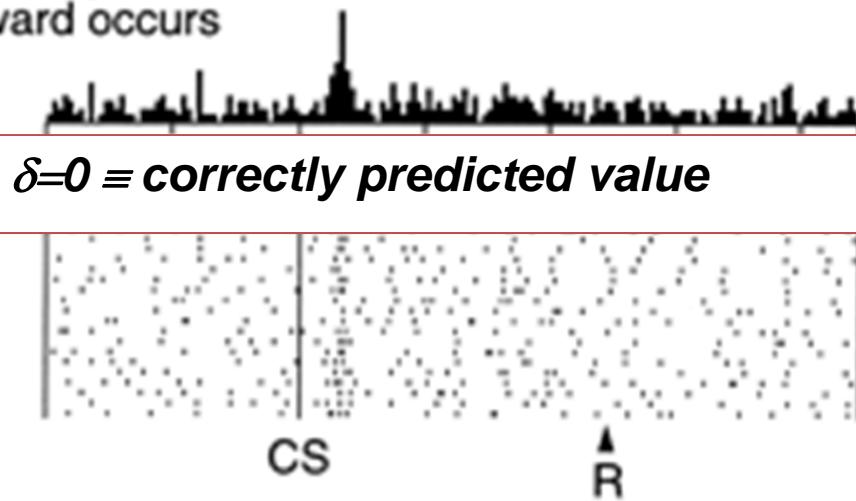
$$V(S_k) \leftarrow V(S_k) + \eta_V \cdot \delta$$



## Dopamine as delta

Phasic dopamine signals appear to track  $\delta$  surprisingly well.

Reward predicted  
Reward occurs



## A Neural Substrate of Prediction and Reward

Wolfram Schultz, Peter Dayan, P. Read Montague\*

\* See all authors and affiliations

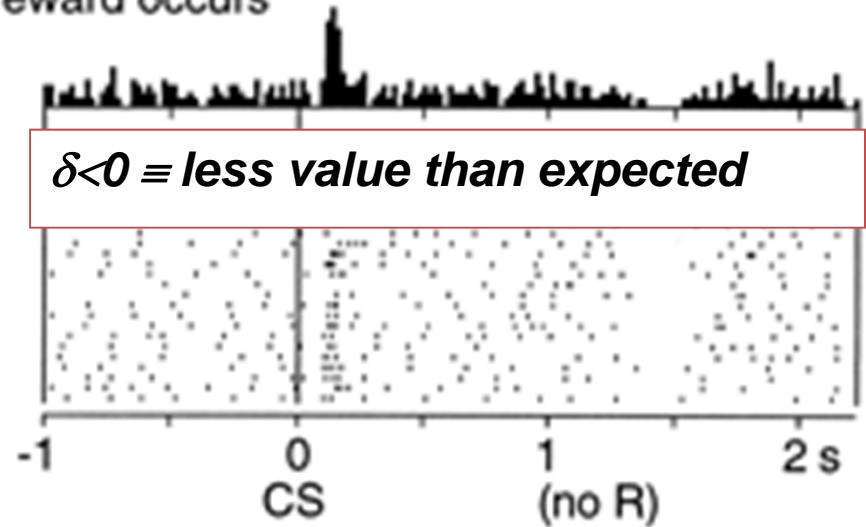
Science 14 Mar 1997;  
Vol. 275, Issue 5306, pp. 1593-1599  
DOI: 10.1126/science.275.5306.1593

No prediction  
Reward occurs

$\delta>0 \equiv \text{more value than expected}$



Reward predicted  
No reward occurs



# What if dopamine is delta?



# What if dopamine is delta?

## Hypothesis

*Dopamine = value prediction-error ( $\delta$ ) in reinforcement learning.*

## Hypothesis

*(Some) drugs of abuse produce dopamine neuropharmacologically.*

## Implication

*Drugs of abuse produce non-compensable value prediction error ( $\delta$ ) signals.*

$$\delta = \max(\gamma^d(R(S_l) + V(S_l)) - V(S_k) + D(S_l), D(S_l))$$

## Conclusion

*At least part of the addictive process is caused by this non-compensable dopamine signal.*

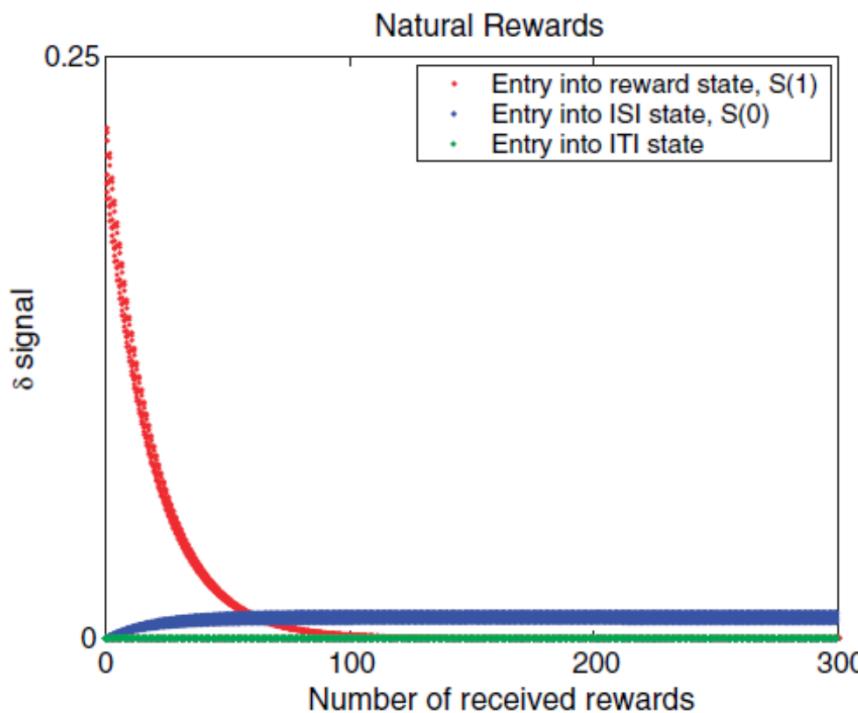
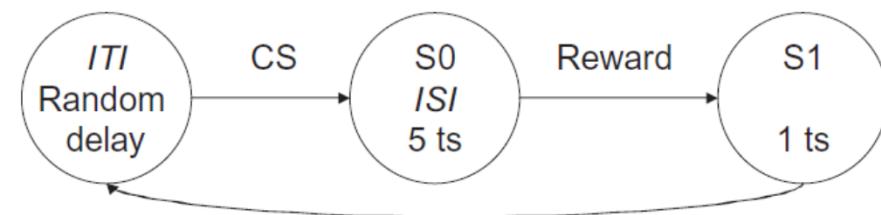
## Addiction as a computational process gone awry

In normal TDRL, learning drives  $\delta$  to 0, but in this equation  $\delta \geq D$ .

### “Failure modes”

Redish (2004) [Science](#)

$$\delta = \max(\gamma^d(R(S_l) + V(S_l)) - V(S_k) + D(S_l), D(S_l))$$



## Addiction as a computational process gone awry

A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives
2. Make drugs inelastic
3. Not show Kamin blocking

### *“Failure modes”*

Redish (2004) Science

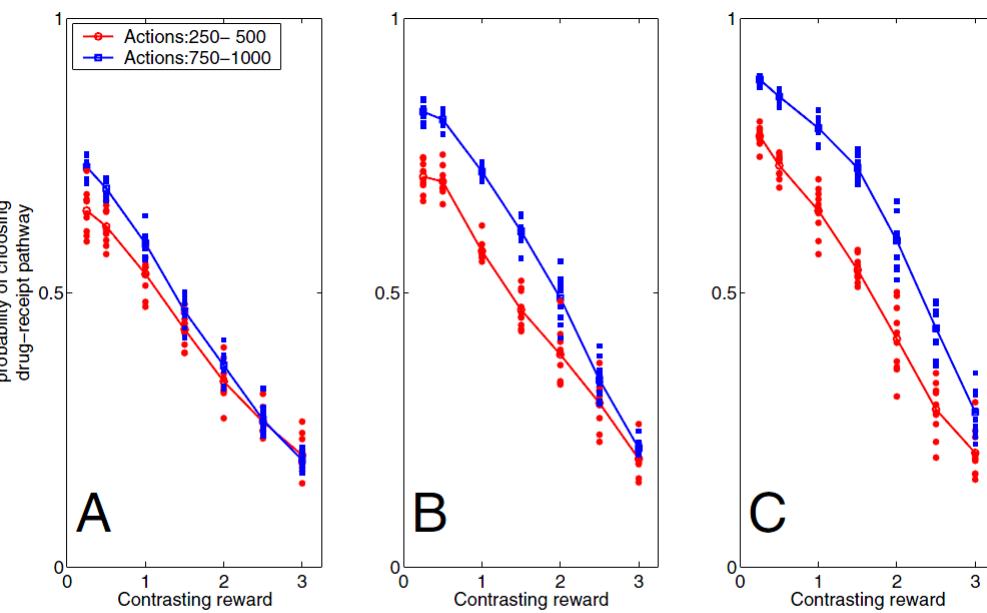
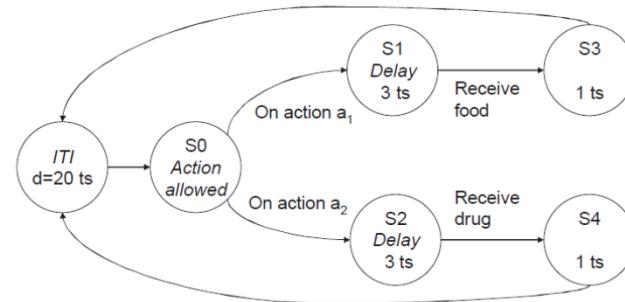
# Addiction as a computational process gone awry

A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives
2. Make drugs inelastic
3. Not show Kamin blocking

## “Failure modes”

Redish (2004) [Science](#)



**FIGURE S2:** Sensitivity of selection to number of drug experiences, size of contrasting food reward, and size of drug-receipt forced-dopamine signal (i.e. strength/dose of the drug). (A)  $R(S_4) = 1.0, D(S_4) = 0.010$ ; (B)  $R(S_4) = 1.0, D(S_4) = 0.025$ ; (C)  $R(S_4) = 1.0, D(S_4) = 0.040$ .

# Addiction as a computational process gone awry

A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives

2. Make drugs inelastic

3. Not show Kamin blocking

## “Failure modes”

Redish (2004) [Science](#)

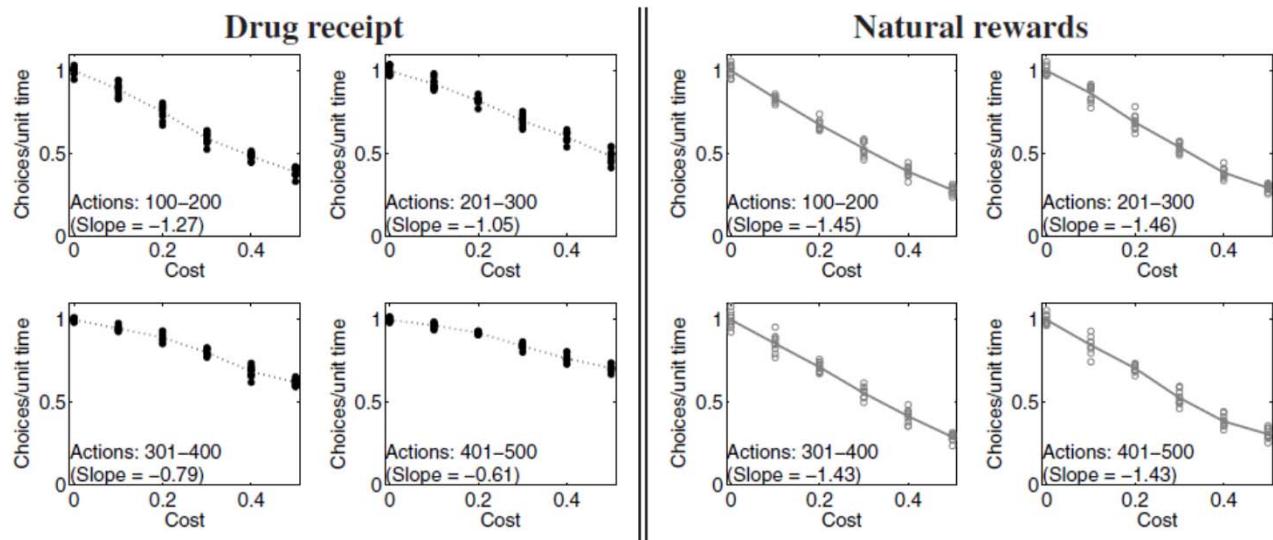
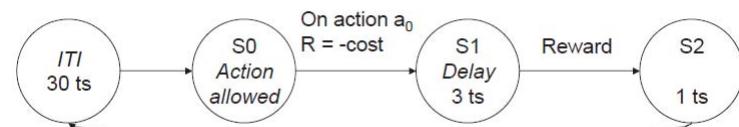


FIGURE S4: Elasticity decreases for drug-receipt but not reward-receipt.

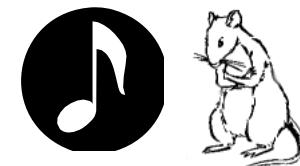
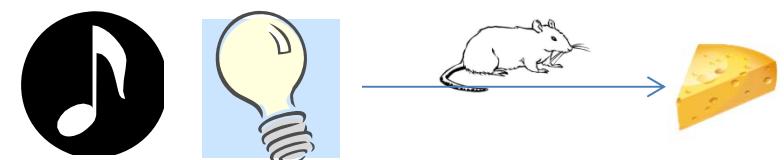
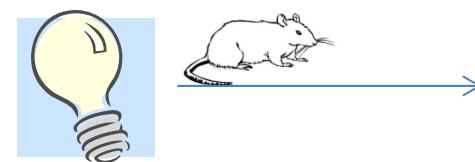
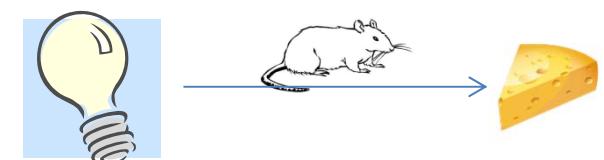
# Addiction as a computational process gone awry

A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives
2. Make drugs inelastic
3. Not show Kamin blocking

## *“Failure modes”*

Kamin 1969



## Addiction as a computational process gone awry

A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives
2. Make drugs inelastic
3. Not show Kamin blocking

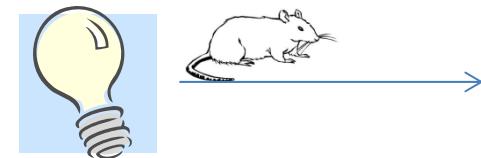
### “Failure modes”

$$\delta(t) = \gamma^d(R(S_l) + V(S_l)) - V(S_k)$$

$$\delta > 0$$



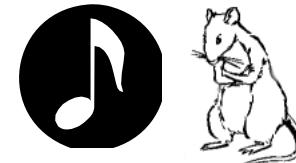
$$V(\text{light}) > 0$$



$$\delta = 0$$



$$V(\text{tone}) = 0$$



## Addiction as a computational process gone awry

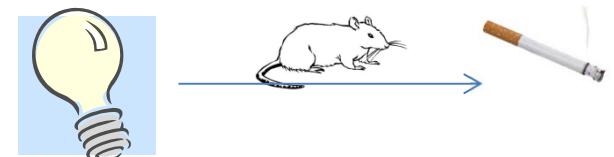
A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives
2. Make drugs inelastic
3. Not show Kamin blocking

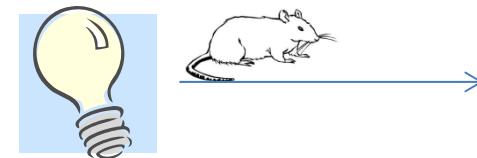
### *“Failure modes”*

$$\delta = \max(\gamma^d(R(S_l) + V(S_l)) - V(S_k) + D(S_l), D(S_l))$$

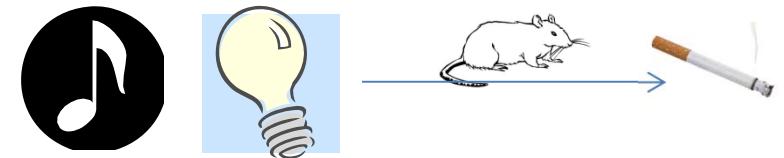
$$\delta > 0$$



$$V(\text{light}) > 0$$



$$\delta > 0$$



$$V(\text{tone}) > 0$$



# Addiction as a computational process gone awry

A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives
2. Make drugs inelastic
3. Not show Kamin blocking

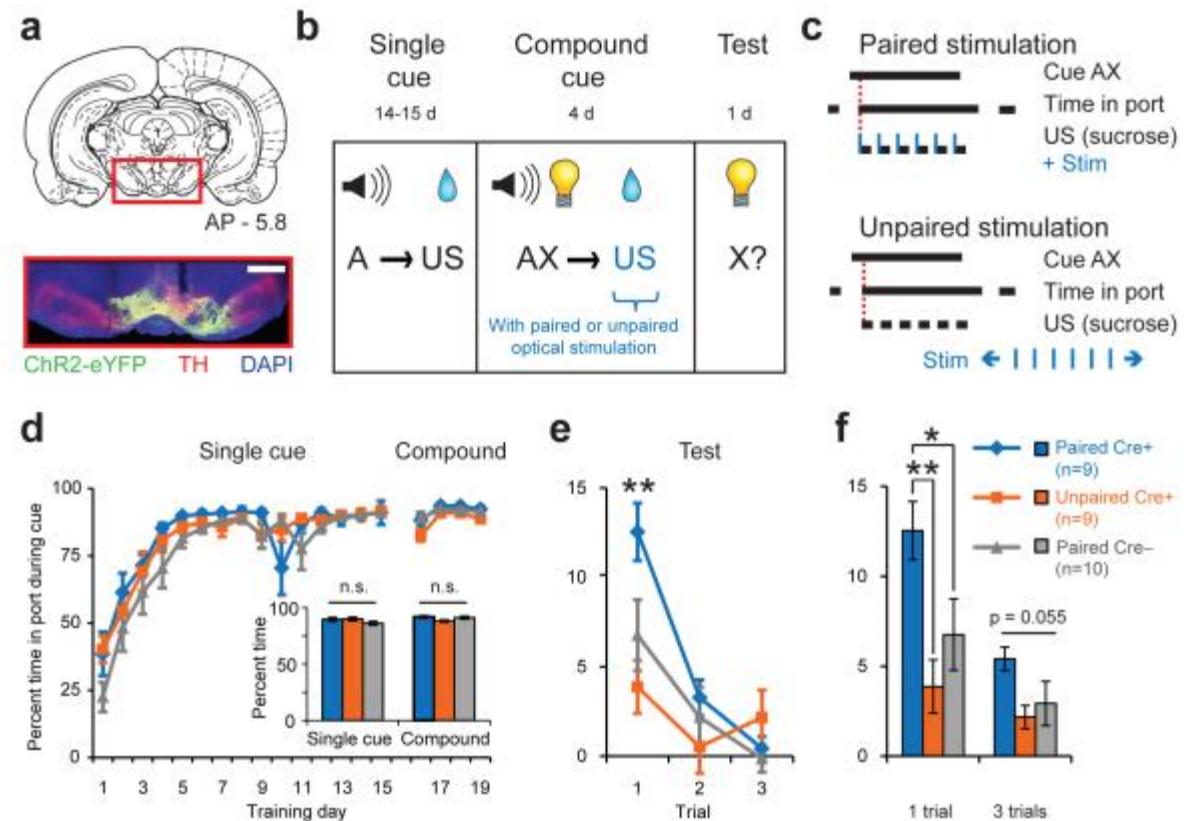
## "Failure modes"

### A causal link between prediction errors, dopamine neurons and learning

nature  
neuroscience

Elizabeth E Steinberg<sup>1,2,11</sup>, Ronald Keiflin<sup>1,11</sup>, Josiah R Boivin<sup>1,2</sup>, Ilana B Witten<sup>3,4</sup>, Karl Deisseroth<sup>5-8</sup> & Patricia H Janak<sup>1,2,9,10</sup>

*Direct manipulations of dopamine disrupt blocking.*



# Addiction as a computational process gone awry

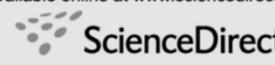
A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives
2. Make drugs inelastic
3. Not show Kamin blocking

**“Failure modes”**

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

 ELSEVIER

 ScienceDirect

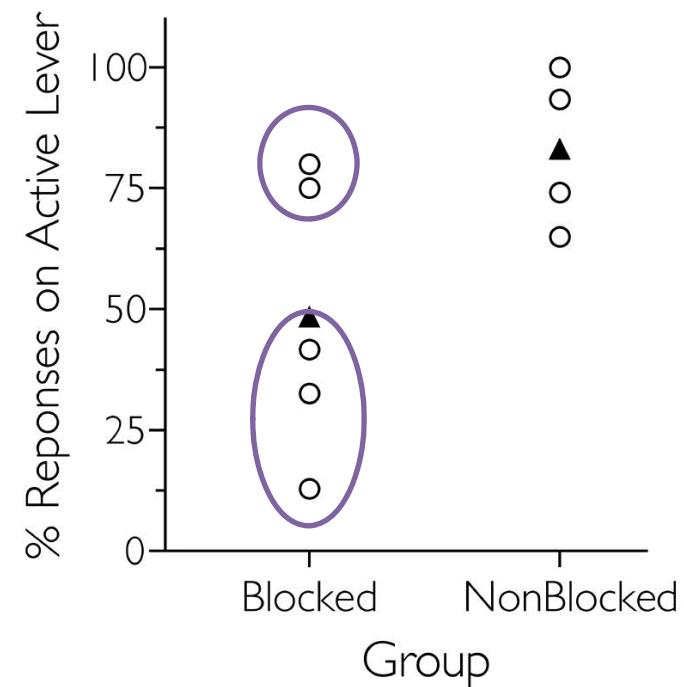
Pharmacology, Biochemistry and Behavior 86 (2007) 774–777

[www.elsevier.com/locate/phambiochembeh](http://www.elsevier.com/locate/phambiochembeh)

Blocking of conditioning to a cocaine-paired stimulus: Testing the hypothesis that cocaine perpetually produces a signal of larger-than-expected reward

Leigh V. Panlilio \*, Eric B. Thorndike, Charles W. Schindler

*But cocaine does not block.*



# Addiction as a computational process gone awry

A non-compensable dopamine signal would...

1. Increase the likelihood of taking drugs over alternatives
2. Make drugs inelastic
3. Not show Kamin blocking

*When Jaffe et al looked at the high-responders only, they saw no blocking to nicotine.*

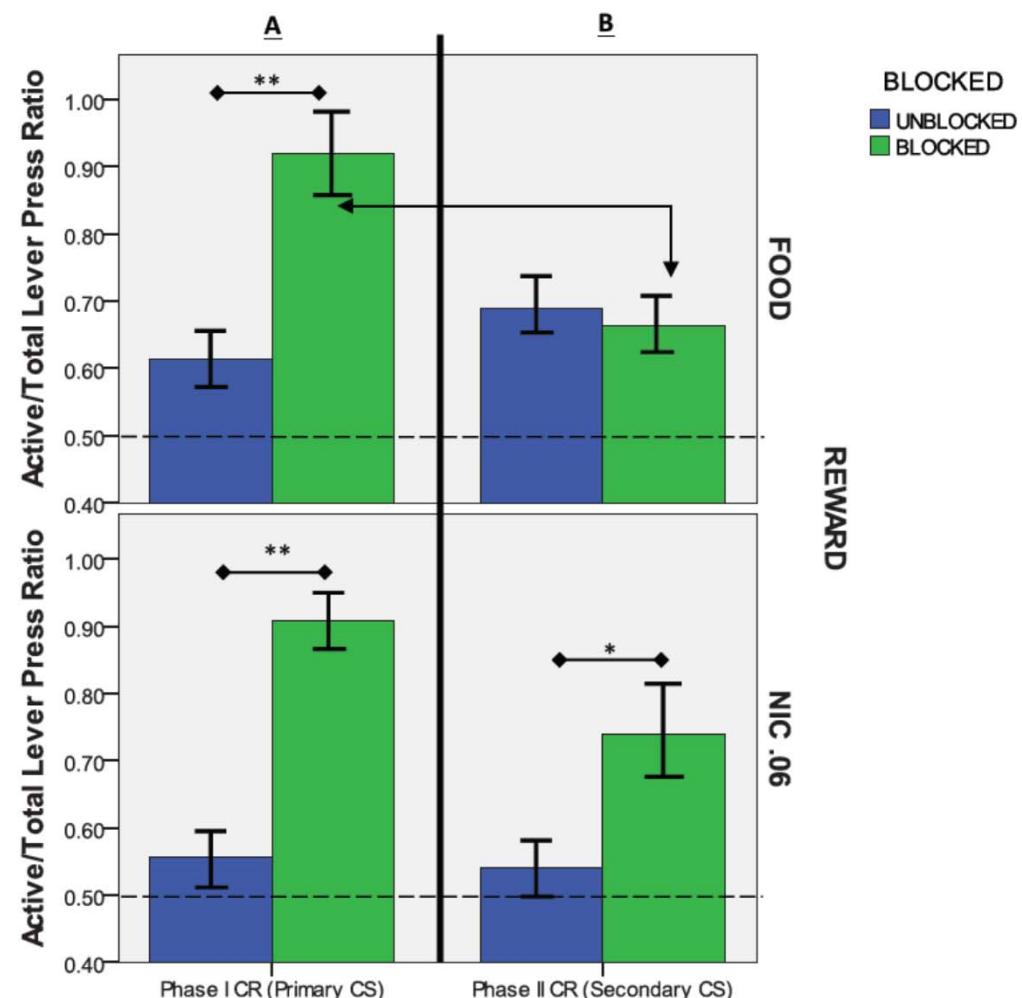
## "Failure modes"

The Open Addiction Journal, 2014, 7, 8-16

Open Access

### The Absence of Blocking Innicotine High-Responders as a Possible Factor in the Development of Nicotine Dependence?

Adi Jaffe<sup>\*1</sup>, J. Aurora Z. Pham<sup>1</sup>, Igal Tarash<sup>1</sup>, Sasha S. Getty<sup>1</sup>, Michael S. Fanselow<sup>1</sup> and J. David Jentsch<sup>1,2</sup>



Note: The dashed line represents the point at which active presses = inactive presses

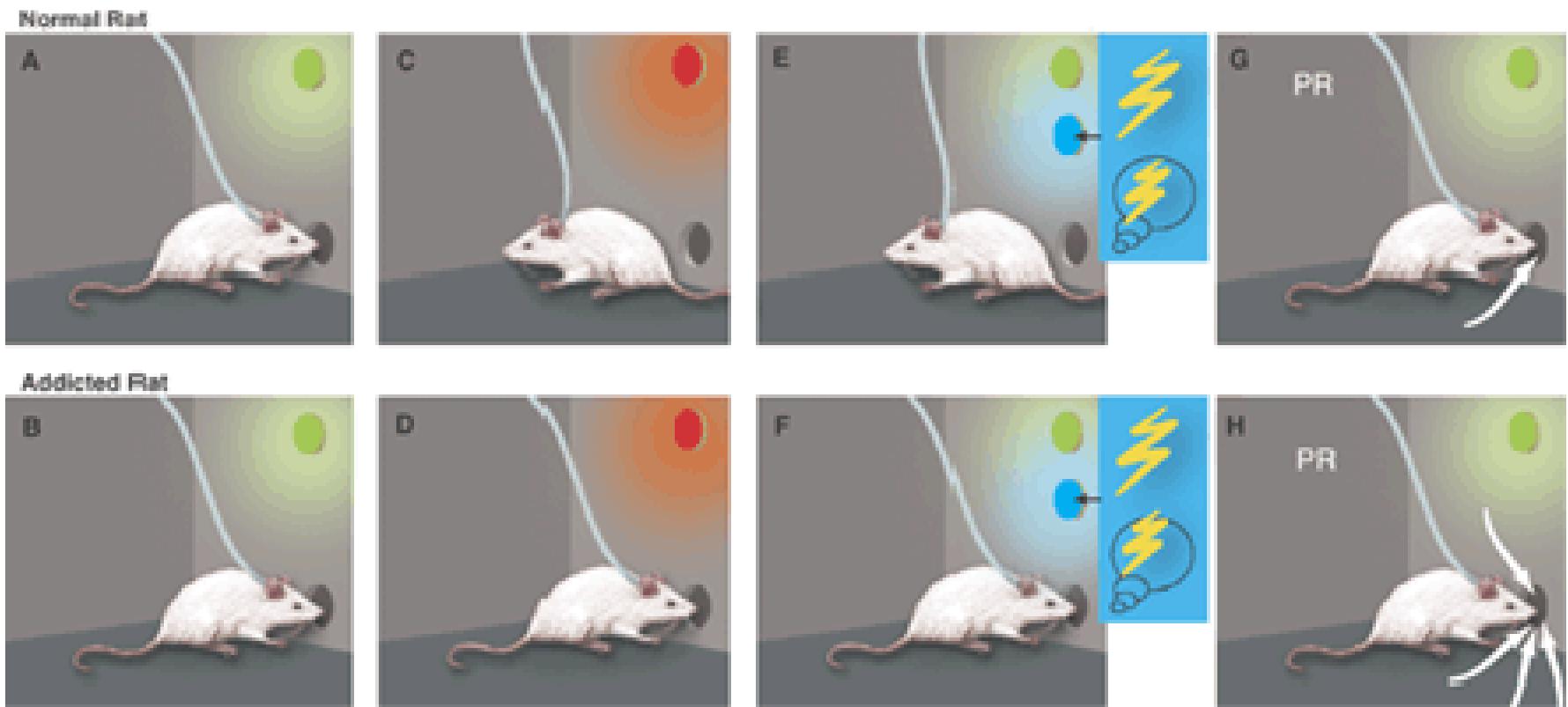
\*  $p < 0.05$

\*\*  $p < 0.01$

## *“Failure modes”*

Addiction ≠ drug use

Not everyone who takes drugs is addicted.



# Not all failure modes are habits



# From addiction to computational psychiatry

*This is but one failure mode of decision that can lead to addiction.*

*Addiction is a symptom, not a disease.*

BEHAVIORAL AND BRAIN SCIENCES (2008) 31, 415–487

*Printed in the United States of America  
doi:10.1017/S0140525X0800472X*

## A unified framework for addiction: Vulnerabilities in the decision process

### **A. David Redish**

*Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455*  
*redish@umn.edu*  
*<http://umn.edu/~redish/>*

### **Steve Jensen**

*Graduate Program in Computer Science, University of Minnesota, Minneapolis,  
MN 55455*  
*jens0491@umn.edu*

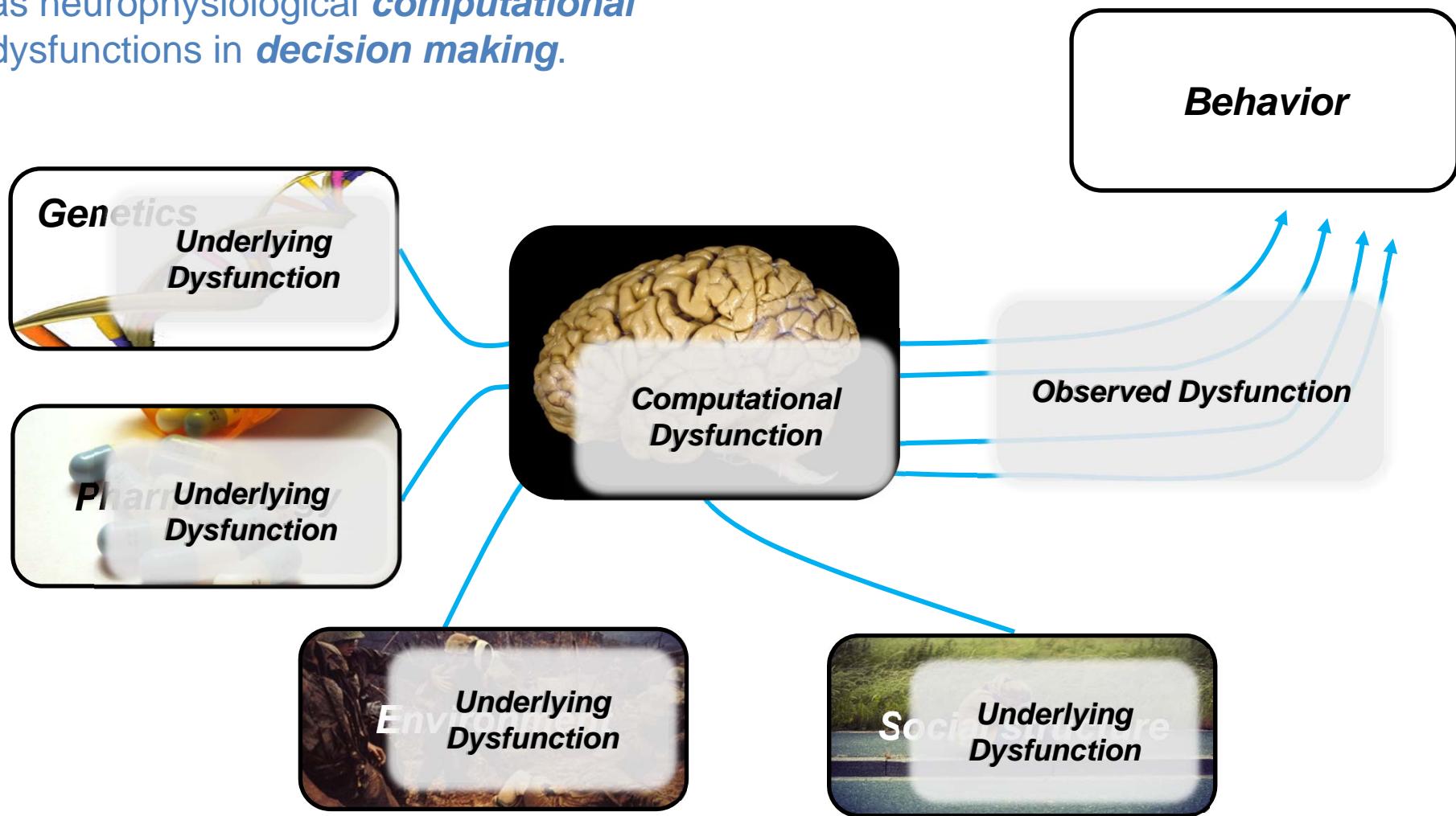
### **Adam Johnson**

*Graduate Program in Neuroscience and Center for Cognitive Sciences,  
University of Minnesota, Minneapolis, MN 55455*  
*john5726@umn.edu*

## Computational Psychiatry

We need to understand **how** decision-making works to begin to understand how it can go wrong.

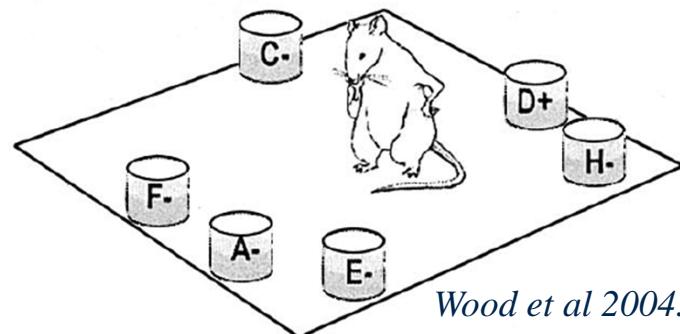
This suggests a new view on psychiatry as neurophysiological **computational** dysfunctions in **decision making**.



Let us operationalize decision-making as the process that leads to selecting an action.

## Decision-making systems

*What action should I take when?*



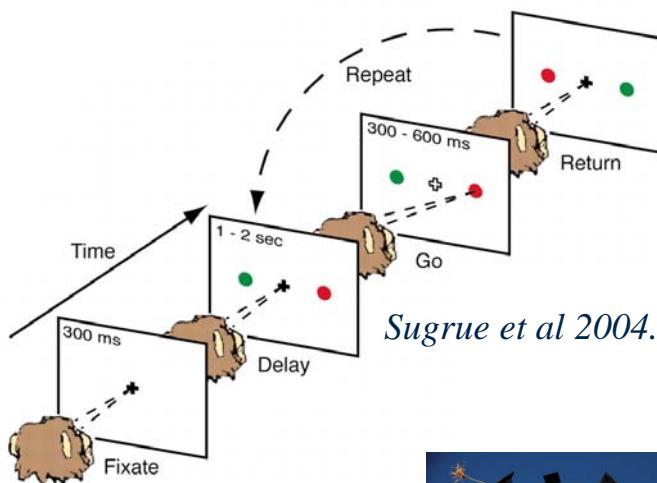
Wood et al 2004.



Rats in a maze... (Redish Lab)



Vending machines



Sugrue et al 2004.



Gambling



Taking drugs... or not...



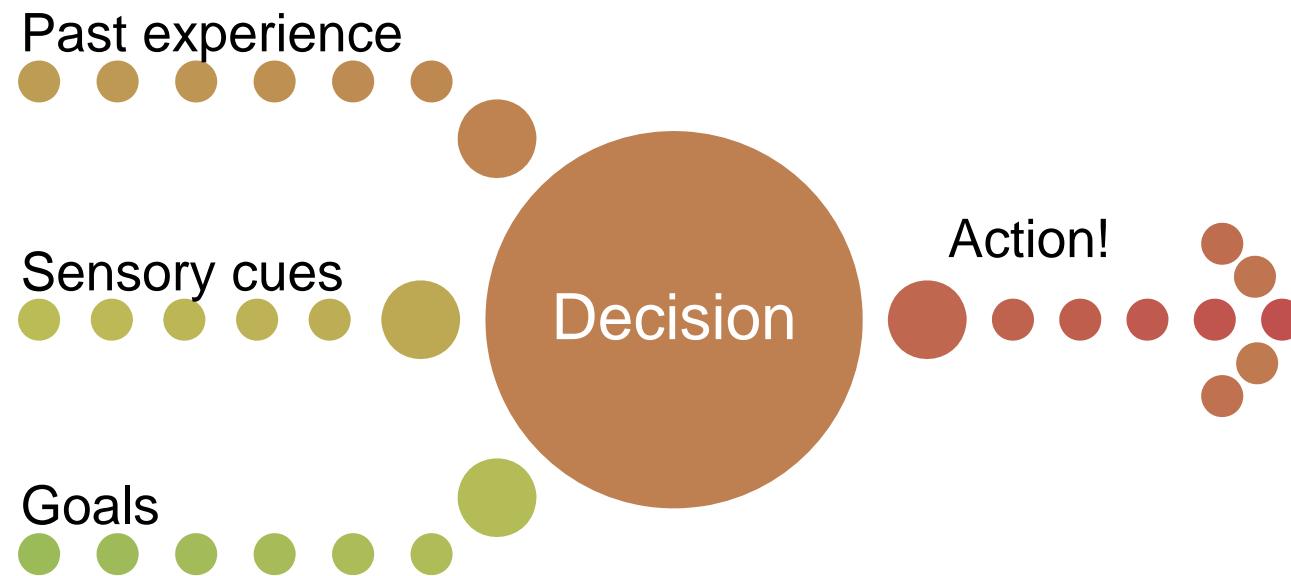
Where to go to college



Getting married...

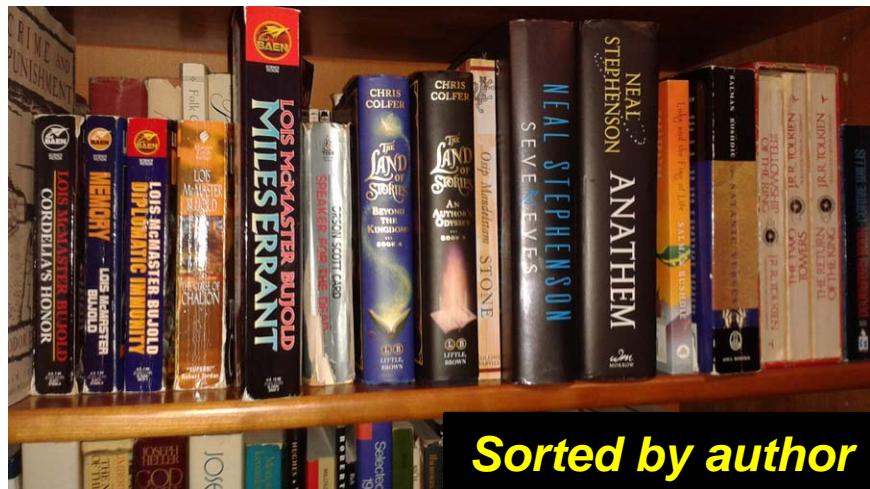
### Computation

Decision-making is about information processing.



## The computational perspective

How you represent the data  
changes the calculation



Sorted by author

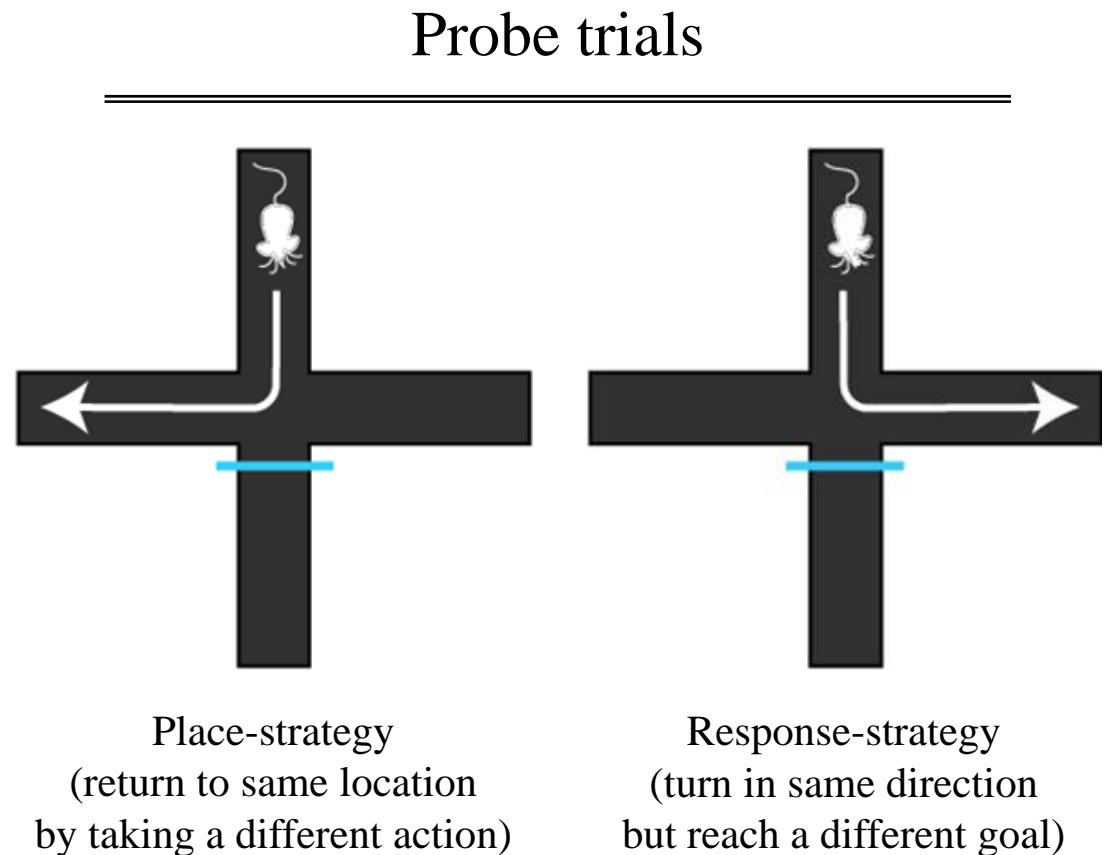


Sorted by size



The “correct” outcome can change, depending on how you represent the data

## *The computational perspective*

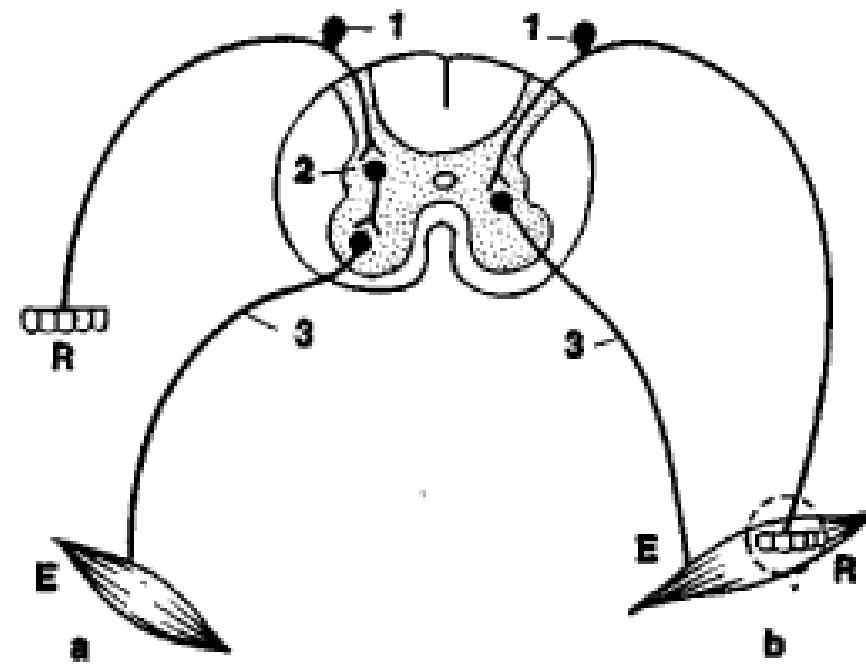


### Reflexes

Reflexes: prewired responses to stimuli.

Goals and stimulus-action pairs are learned over an *evolutionary timescale*.

Learning within the lifespan is limited to habituation, sensitization, and other simple threshold adjustments.



## Instinctual action-selection systems

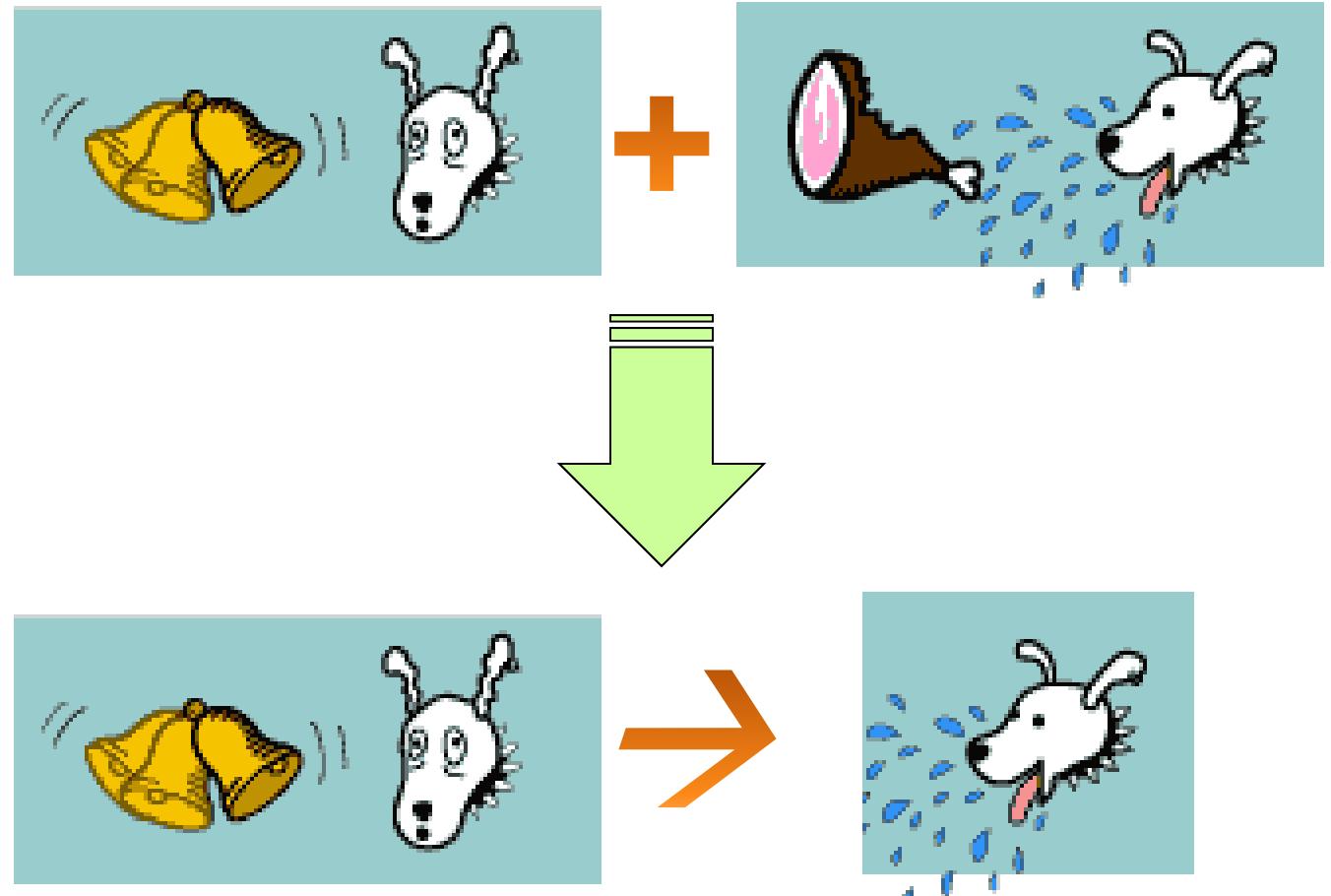
Reflexes: prewired responses to stimuli.

Instinctual (Pavlovian, emotional): learning the situation to release prewired actions.

## Decision-making systems

Learning to release prewired actions under the right conditions.

Emotionally driven.



## *Decision-making systems*

### Deliberation

Reflexes: prewired responses to stimuli.

Searching through potential outcomes.

Cognitively and computationally expensive.

Pavlovian (emotional): learning the situation to release prewired actions.

Requires mental time travel.

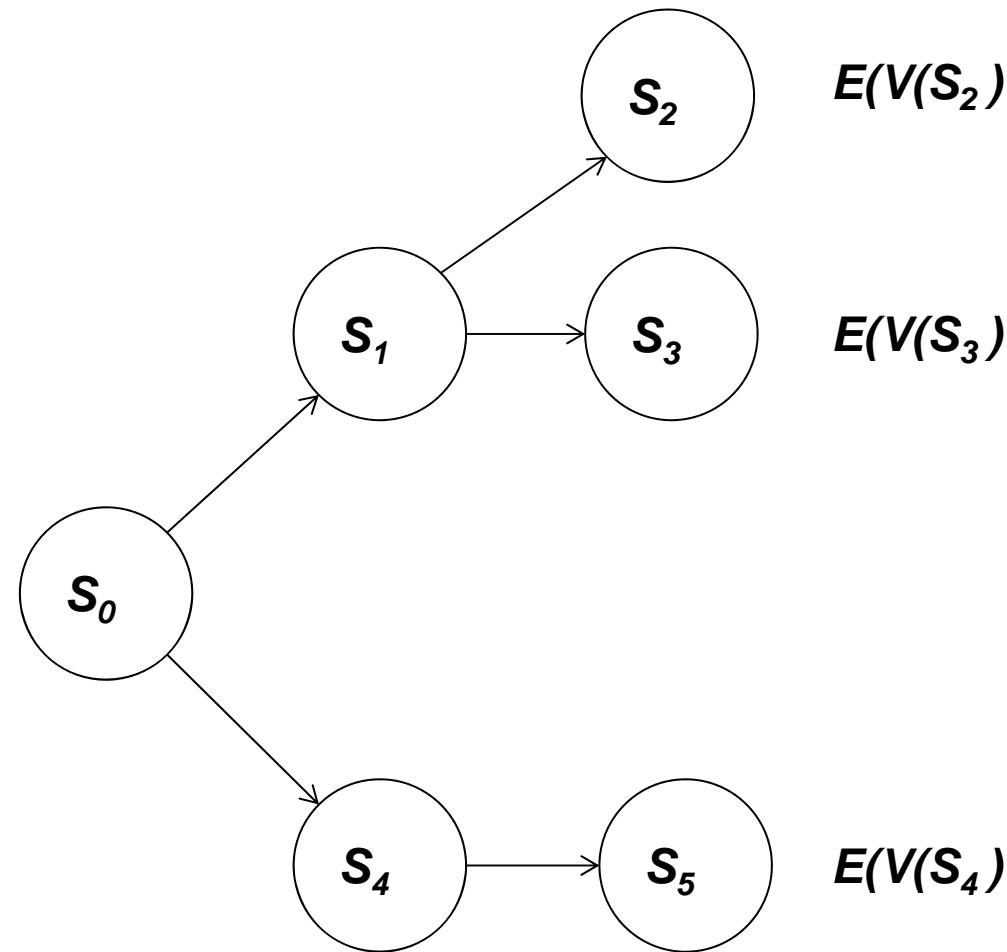
Deliberation: search and evaluate potential consequences.



## Models

Current models of planning suggests that prefrontal-hippocampal interactions allow the creation of simulated worlds that are subsequently evaluated.

### “Failure modes”



We can see failure modes of expectation.

Or of evaluation.

Failure modes of expectation

“cigarettes are cool”

“Failure modes”



*Joe Camel advertising to kids, 1987-1997*



*Humphrey Bogart and Lauren Bacall,  
To have and have not, 1944*

## Procedural Habits

Reflexes: prewired responses to stimuli.

Instinctual: learning the situation to release prewired actions.

Deliberation: search and evaluate potential consequences.

Procedural (habits): cached action-chain sequences.

## *Decision-making systems*

Fast but inflexible.

Learned slowly.



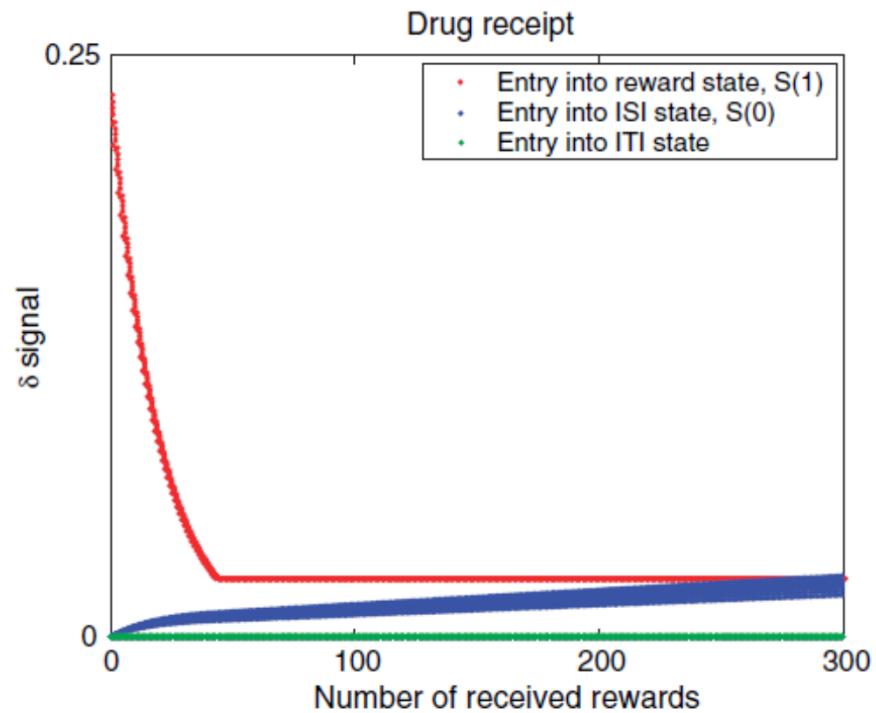
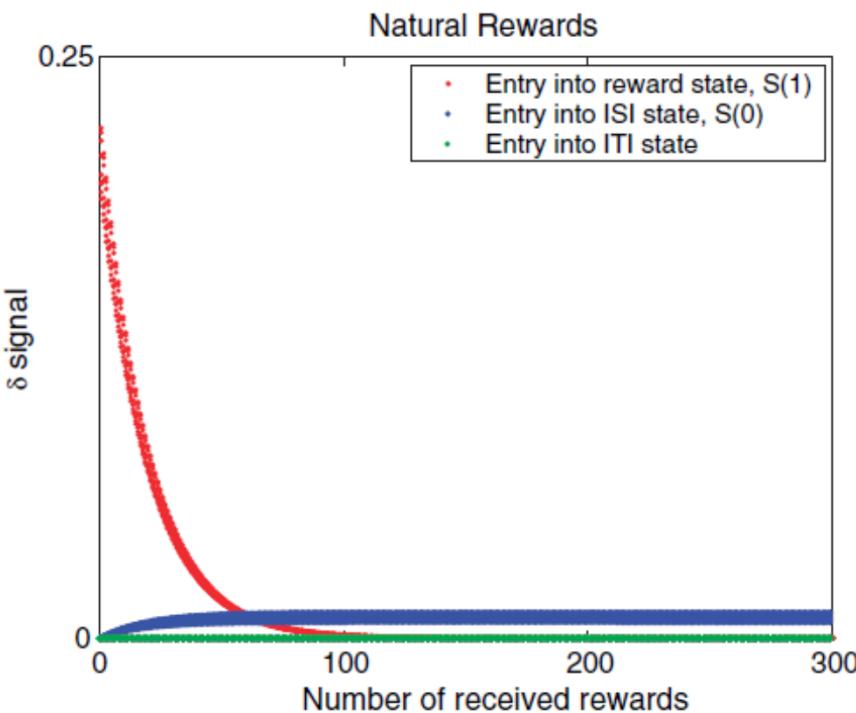
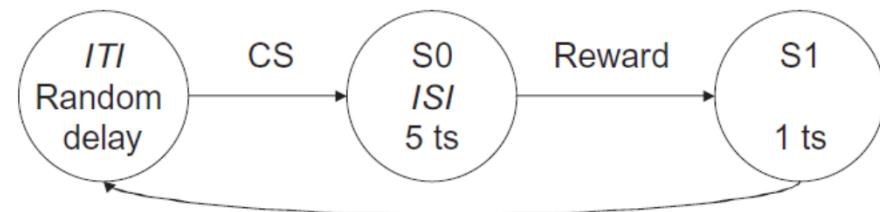
## Addiction as a computational process gone awry

In normal TDRL, learning drives  $\delta$  to 0, but in this equation  $\delta \geq D$ .

### “Failure modes”

Redish (2004) [Science](#)

$$\delta = \max(\gamma^d(R(S_l) + V(S_l)) - V(S_k) + D(S_l), D(S_l))$$



## Decision-making systems

Reflexes: prewired responses to stimuli.

Pavlovian (emotional): learning the situation to release prewired actions.

Deliberation: search and evaluate potential consequences.

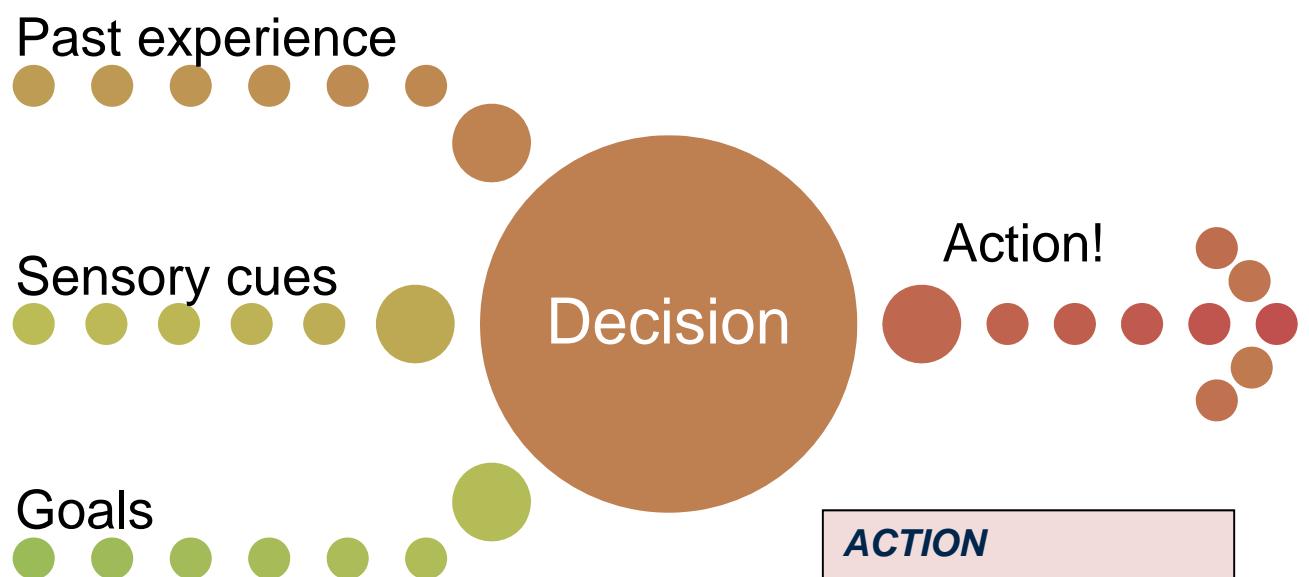
Procedural (habits): cached action-chain sequences.

### PERCEPTION

*Integrating information...*

### SITUATION RECOGNITION

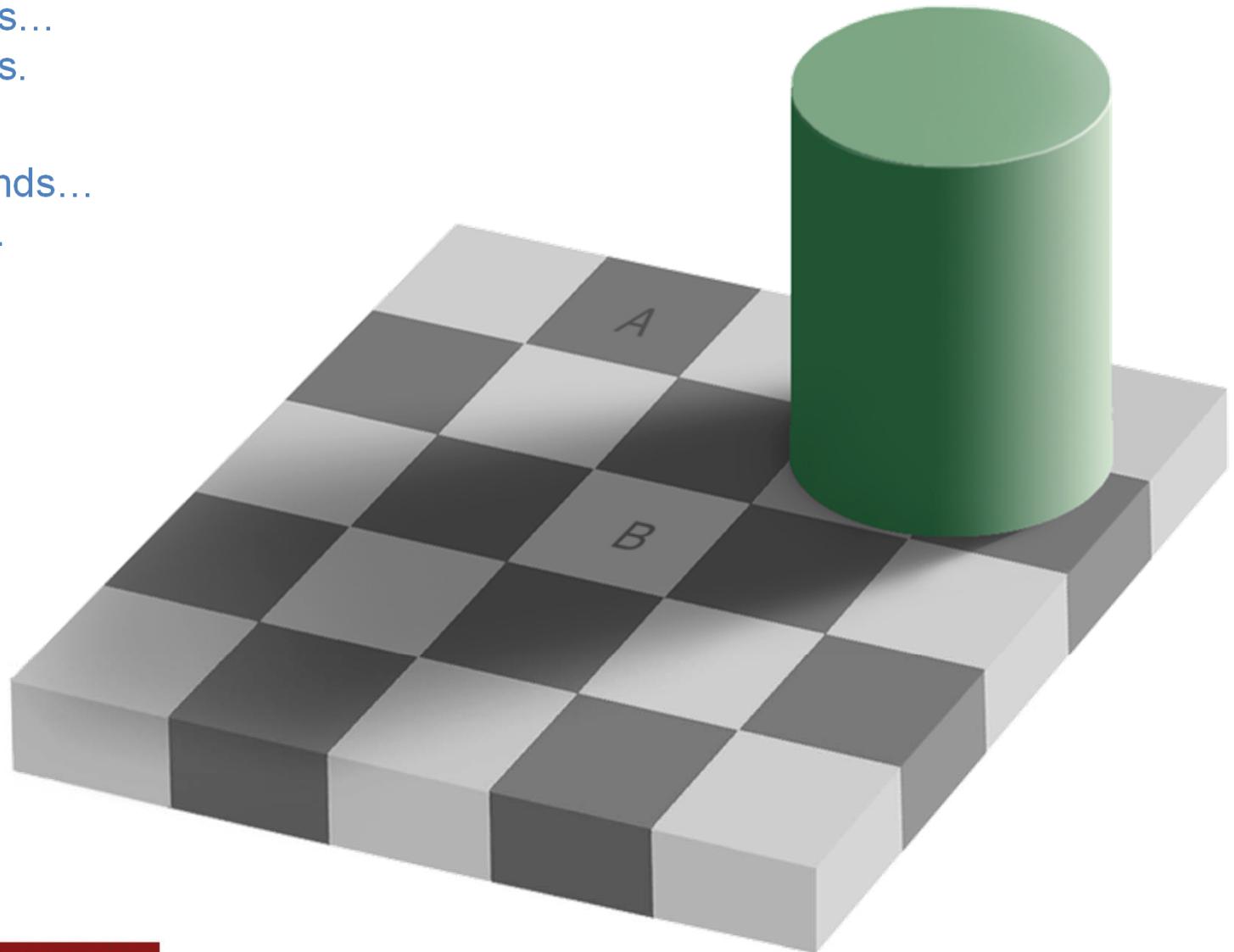
*Classification of the situation by cortical prototype mechanisms.*



### Situation recognition

We don't see colors...  
we perceive objects.

We don't hear sounds...  
we perceive music.



## Environments

Classical rat experiments  
put rats in “Skinner boxes”  
with nothing but a lever



*But real rats don't live in  
these controlled  
environments.*



## Cue-rich environments

We live in cue-rich environments, and have to learn to recognize the key factors of the situation.

*This state categorization process defines*

*... the space that one searches over in planning systems*

*... the parameters that habit systems respond to.*



# State of the world as a categorization problem

**“Failure modes”**

**Note: this is the inferential brain  
that you've been hearing about  
all week.**

## Hypothesis

*The world is ever changing.*

## Hypothesis

*Decisions are made on categorized representations.*

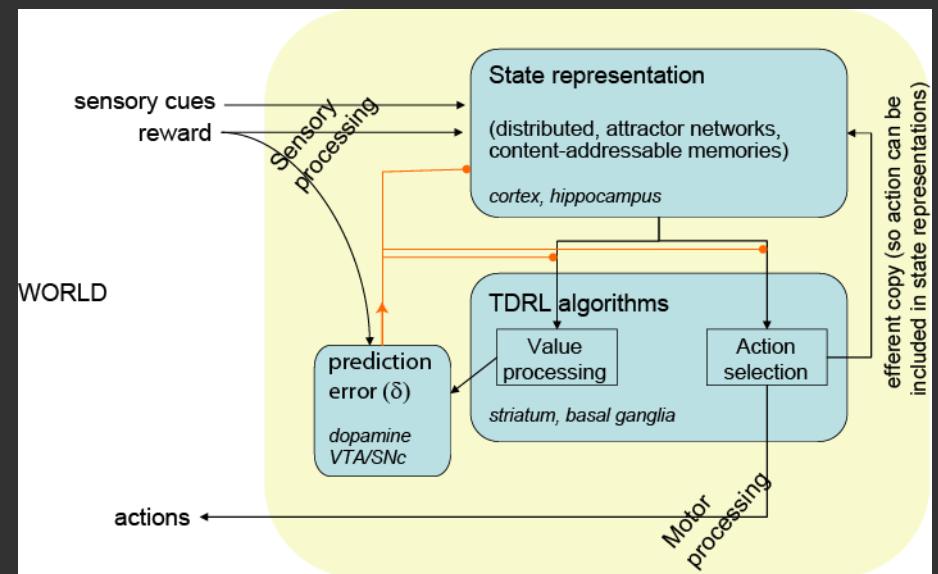
## Implication

*There is a categorization process.*

## Conclusion

*At least part of the addictive  
process can be caused by  
mistakes in this  
categorization process.*

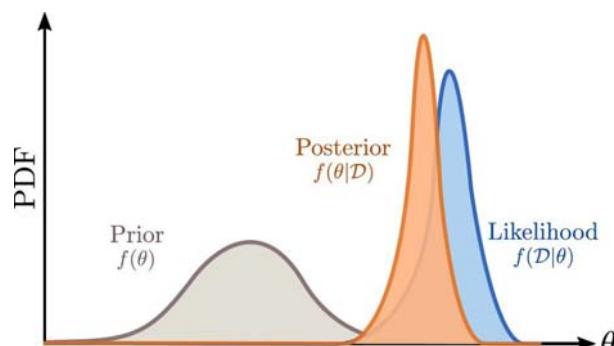
**In order to use a POMDP, you first  
need to identify the possible states.**



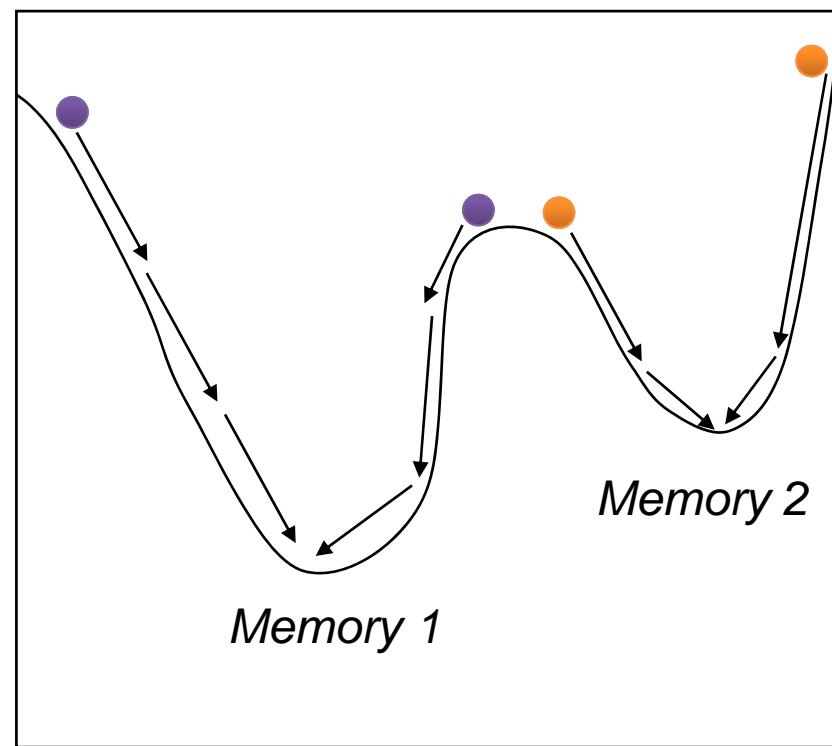
Two theories of categorization

## “Failure modes”

### ***Bayesian update***



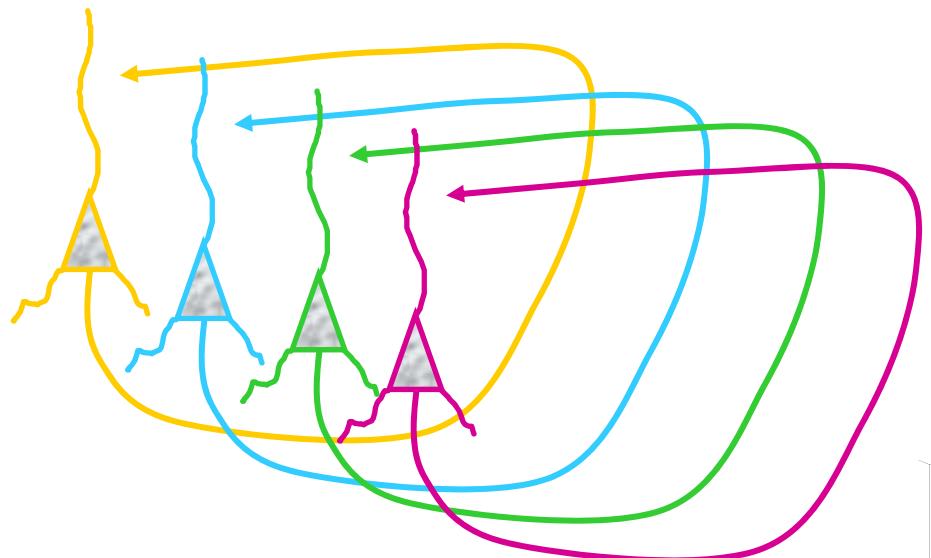
### ***Basins of attraction***



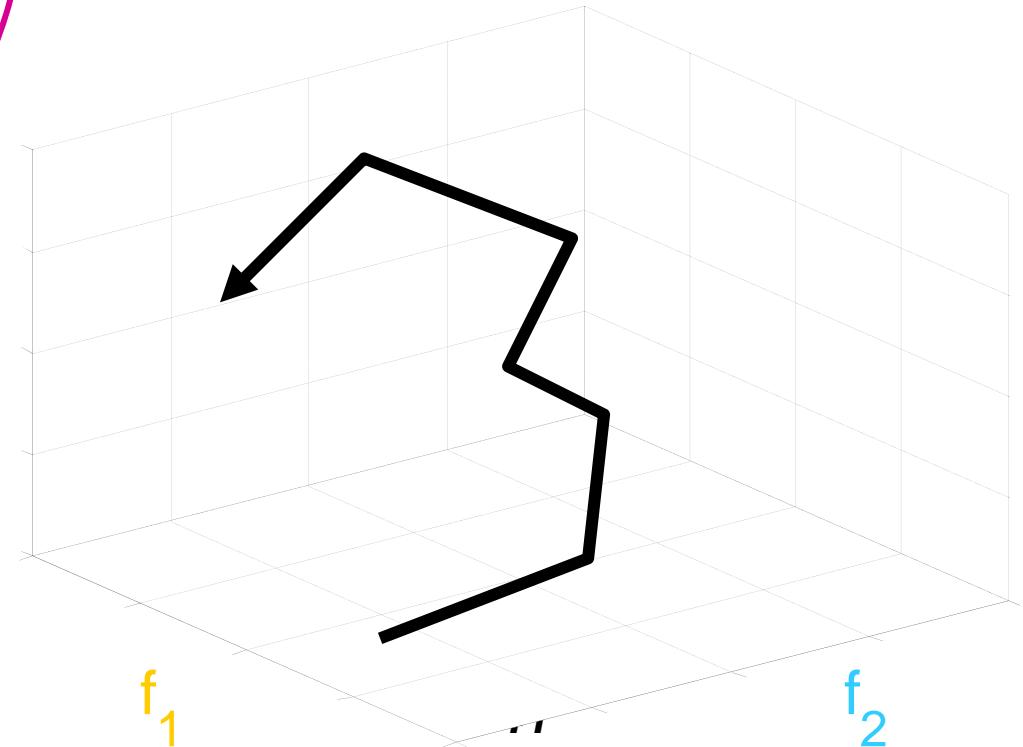
*Different sets of neural activity*

Cell assemblies as a point in N-dimensional space

## *“Failure modes”*

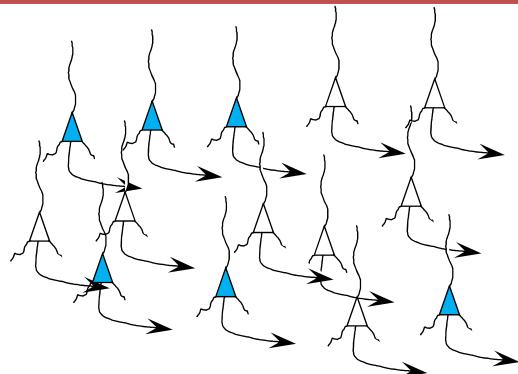


$$F = (f_1, f_2, f_3, f_4, \dots, f_n)$$

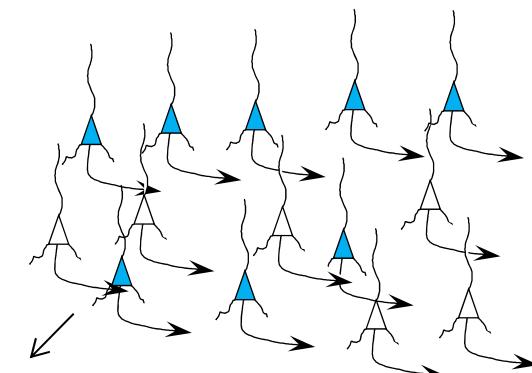
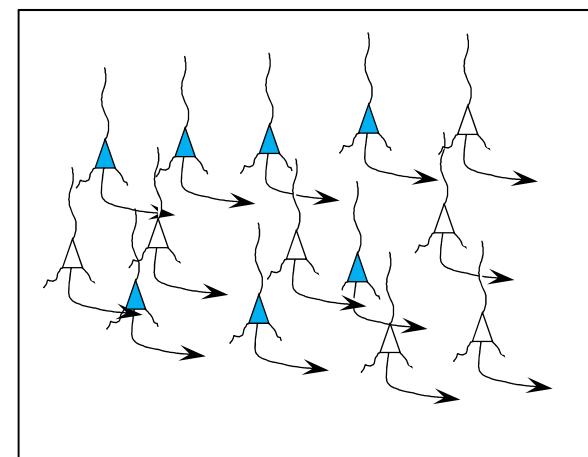


## *“Failure modes”*

A distance metric

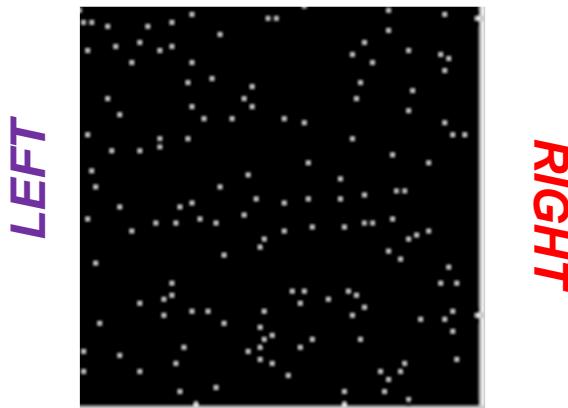


**3 cells  
changed**

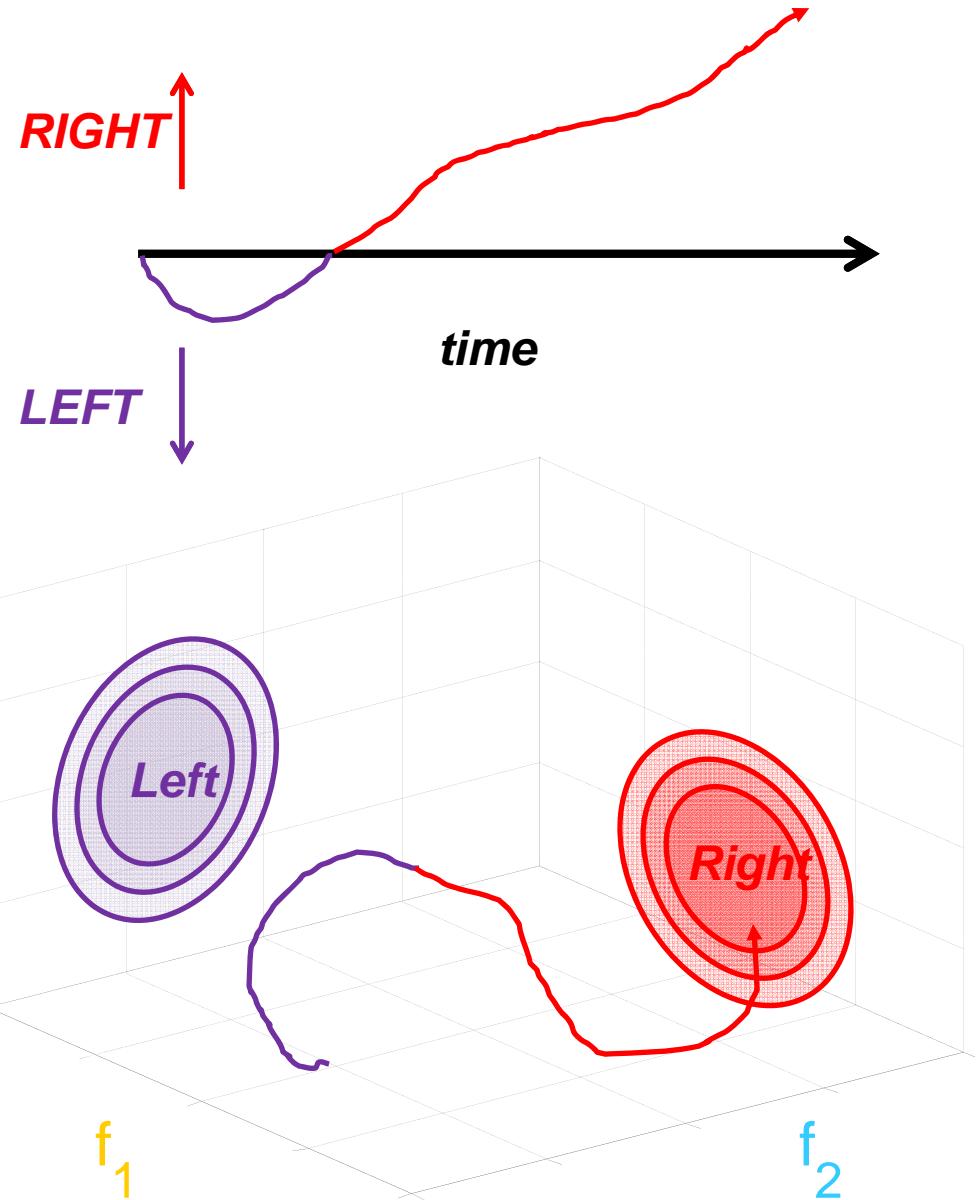


**1 cell  
changed**

## Basins of attraction



Integrating information is descent  
into a basin of attraction,  
by moving through neural space.



## *“Failure modes”*

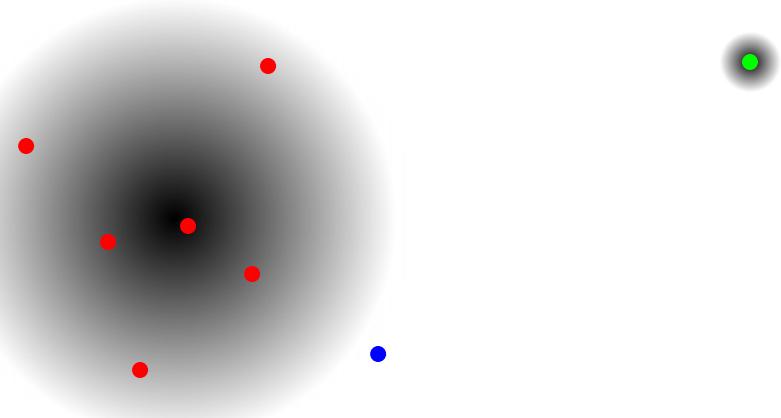
### A starting model

Let's hypothesize that the classification window depends on the level of dopamine present in the system.

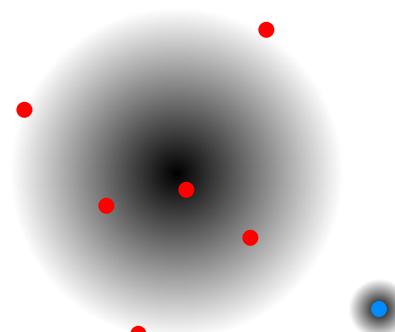
That is, the depth of the basin of attraction depends on dopamine.

***Note that depth of the basin depends on the “inverse temperature” – deep basins are cold (low noise), shallow basins are hot (high noise).***

High dopamine

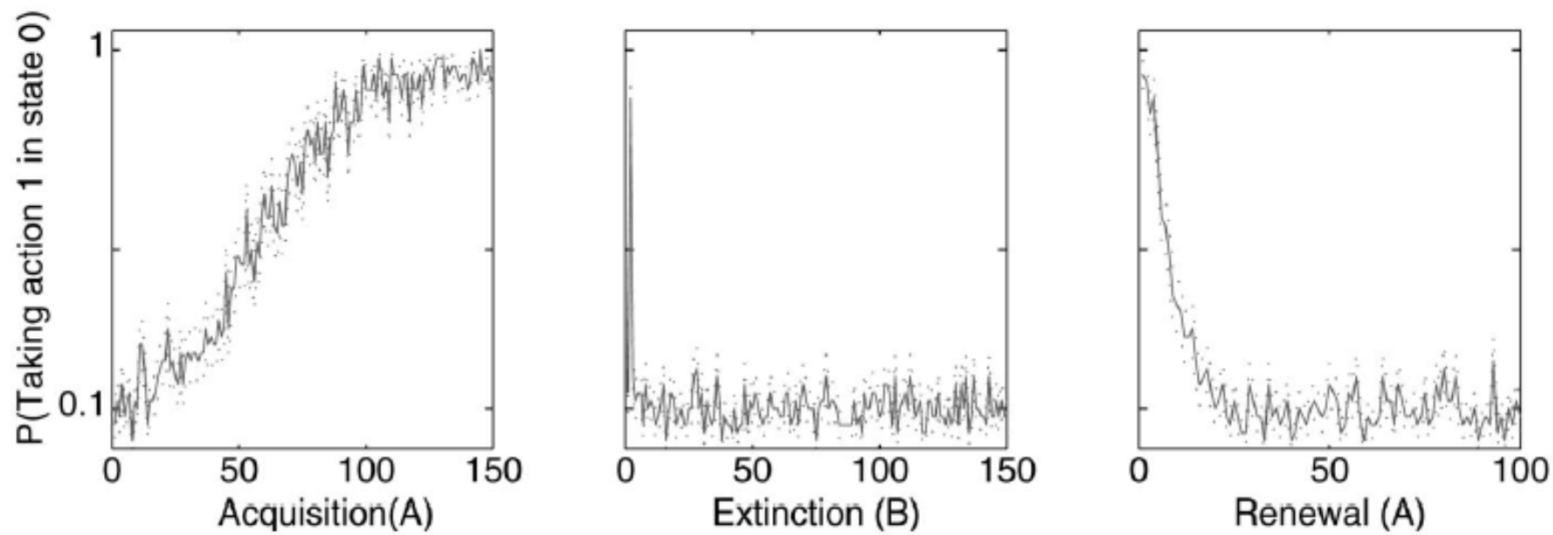
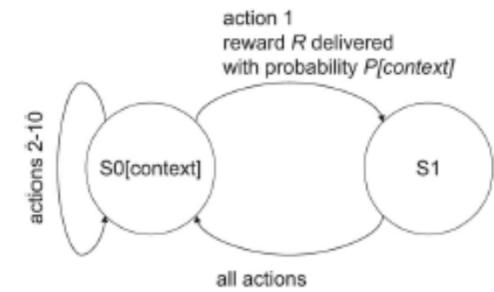


Low dopamine



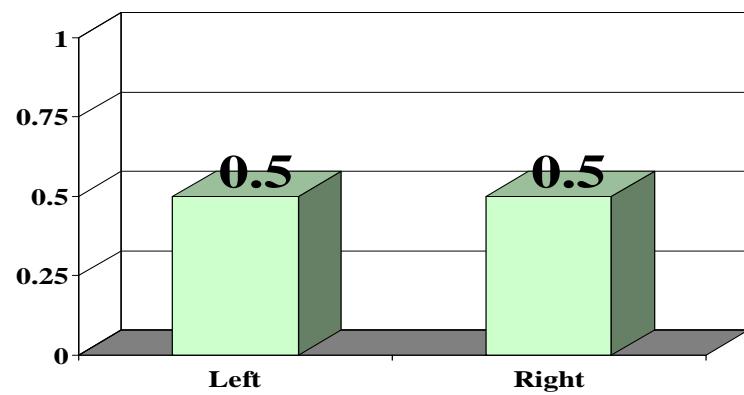
## *Situation recognition*

### Behavioral extinction



Brains work because  
information

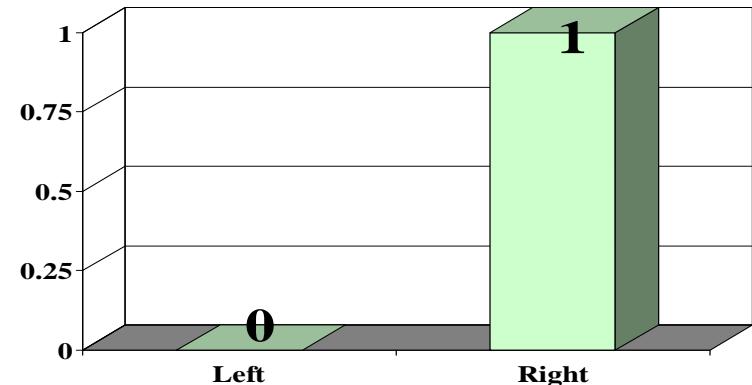
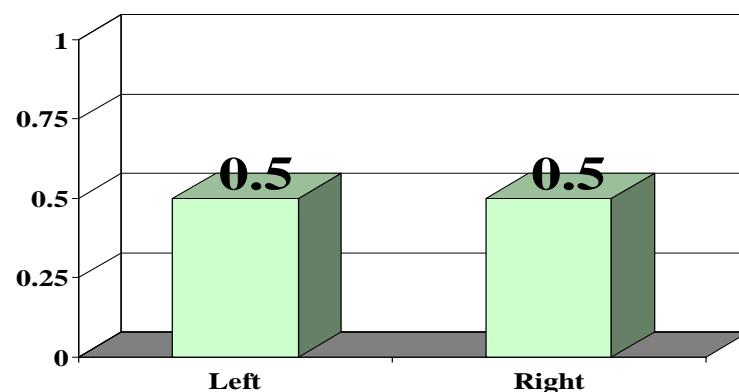
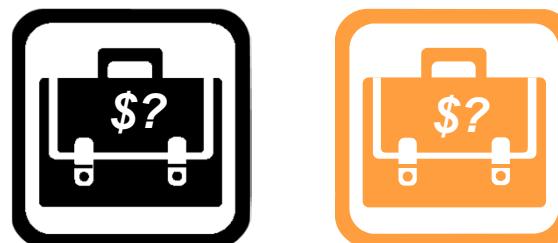
## *The computational perspective*



Brains work because information

Even if we can't measure the actual entropy, we can measure the **information gained** (entropy accounted for).

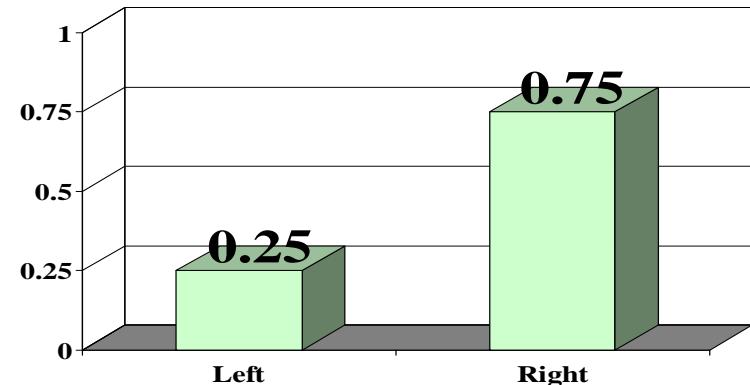
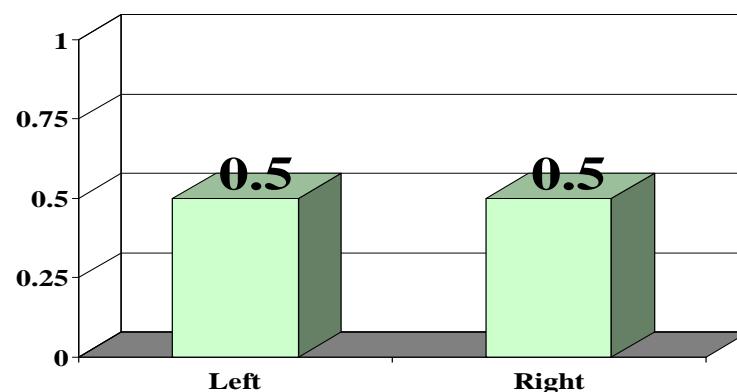
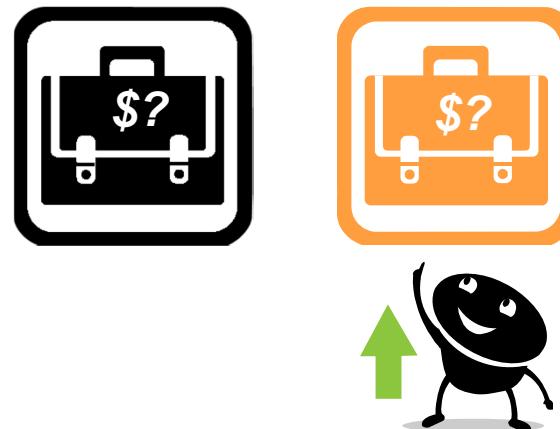
## The computational perspective



Brains work because information

Even if we can't measure the actual entropy, we can measure the **information gained** (entropy accounted for).

## The computational perspective

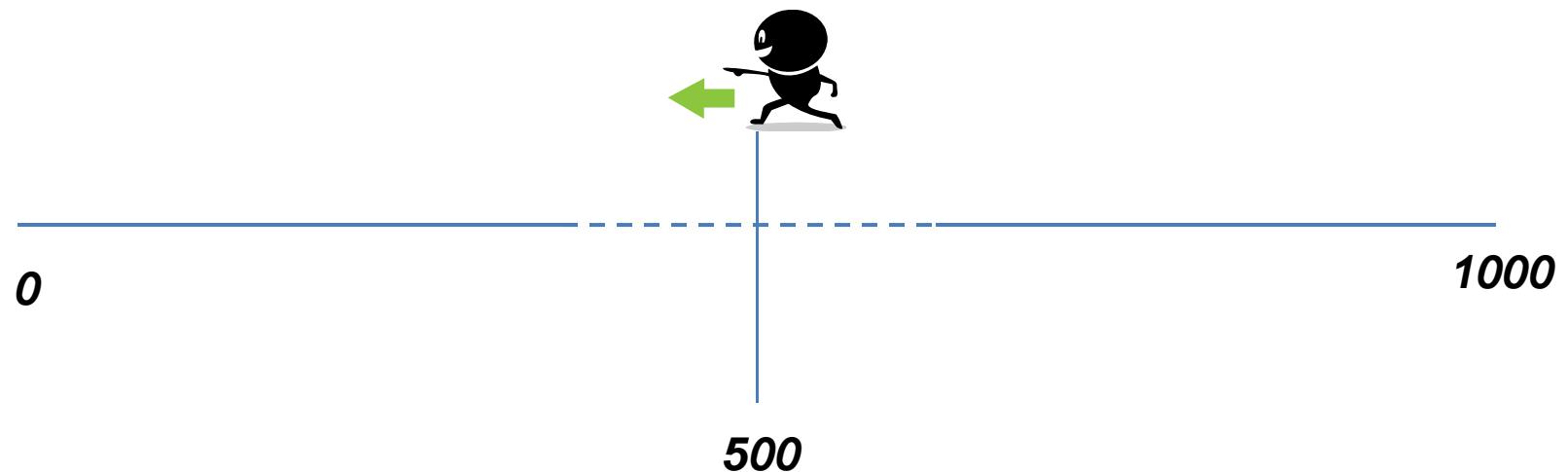


Brains work because  
information

Even if we can't measure  
the actual entropy, we  
can measure the  
information gained.

### The computational perspective

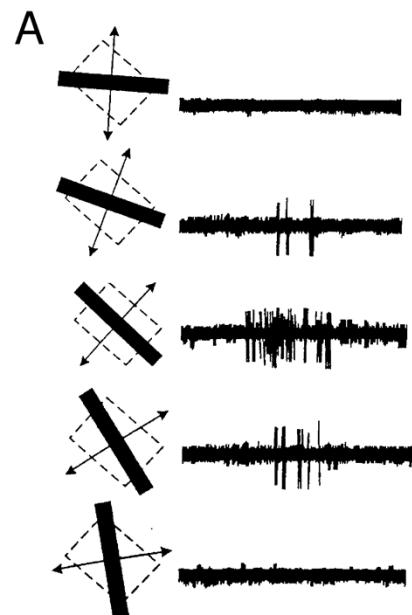
*This is called the  
mutual information  
between X and Y.*



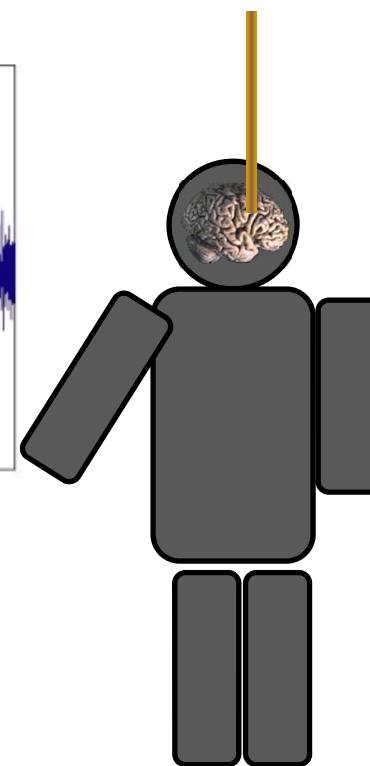
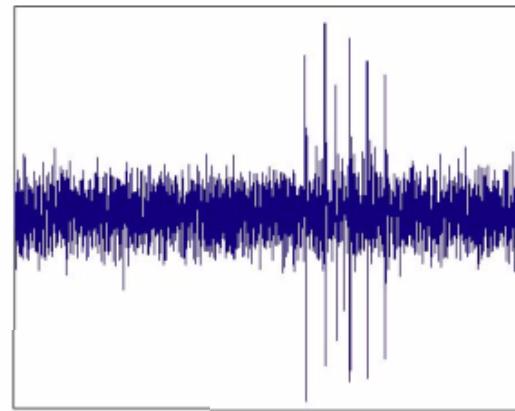
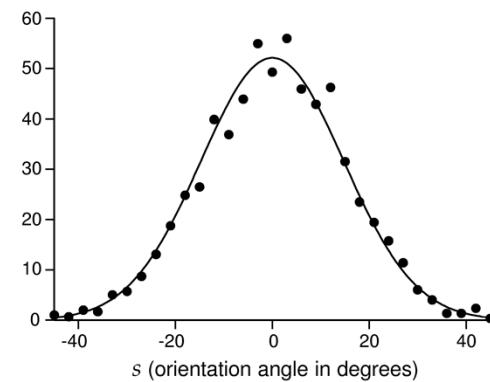
## The computational perspective

### Mutual information

Neural signals have mutual information with properties of the world.



**V1 responses**  
Hubel and Weisel 1968

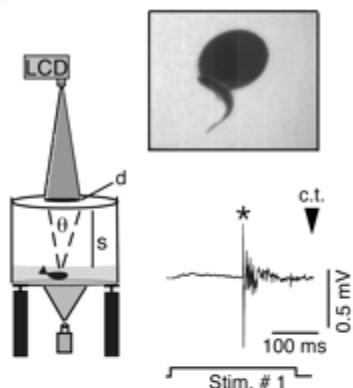


## The computational perspective

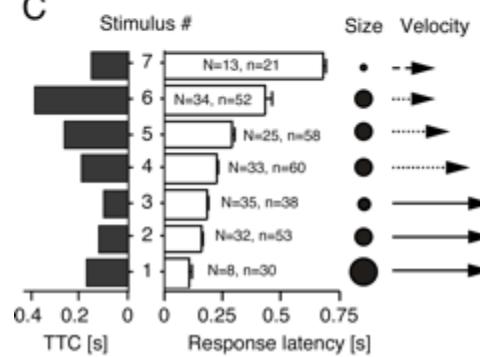
### Mutual information

The Mauthner cell in the goldfish.

A



C



Preuss...Faber (Jnsci 2006)



A1



Stimulus 2 reversed; receding disc

A2



Stimulus 1; looming black disk

50  
mV

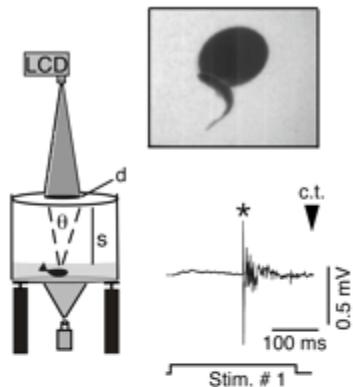


## The computational perspective

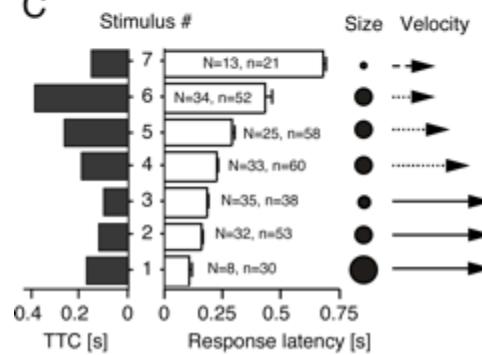
### Mutual information

The Mauthner cell in the goldfish.

A



C



Preuss...Faber (Jnsci 2006)

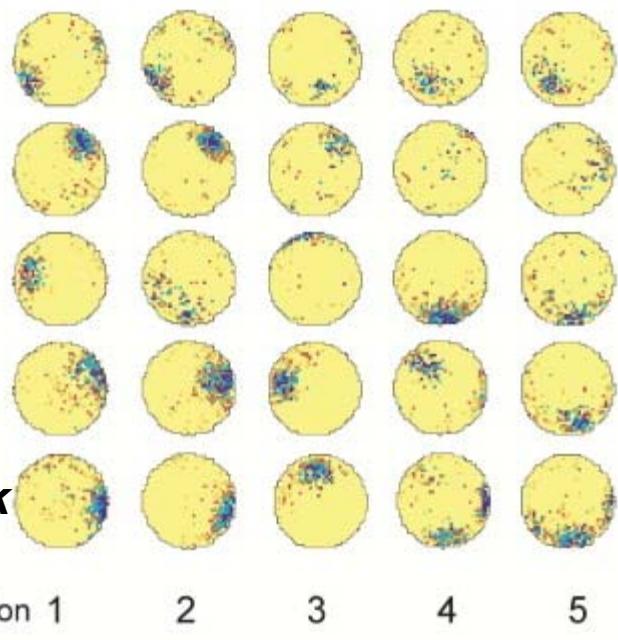


## Tuning curve stability and attention

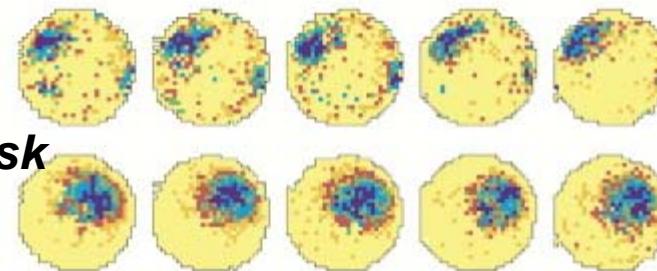
Place cells are more stable from day to day if the animal is “attending” to a task.

### “Failure modes”

Kentros et al. 2004



**Mice: no task**



**Mice: spatial task**

**Tuning curve stability = stable categorization = stable basins of attraction.**

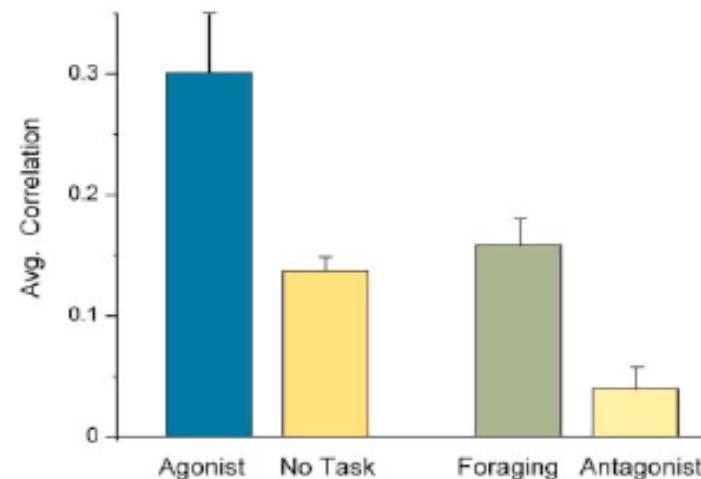
## *“Failure modes”*

### Tuning curves stability

Tonic levels of dopamine control the stability of tuning curves.

In mathematical terms, tonic dopamine strengthens the inertia of the basin of attraction in the content addressable memory.

D1/D5 Receptor Agonists and Antagonists Modulate Place Field Stability



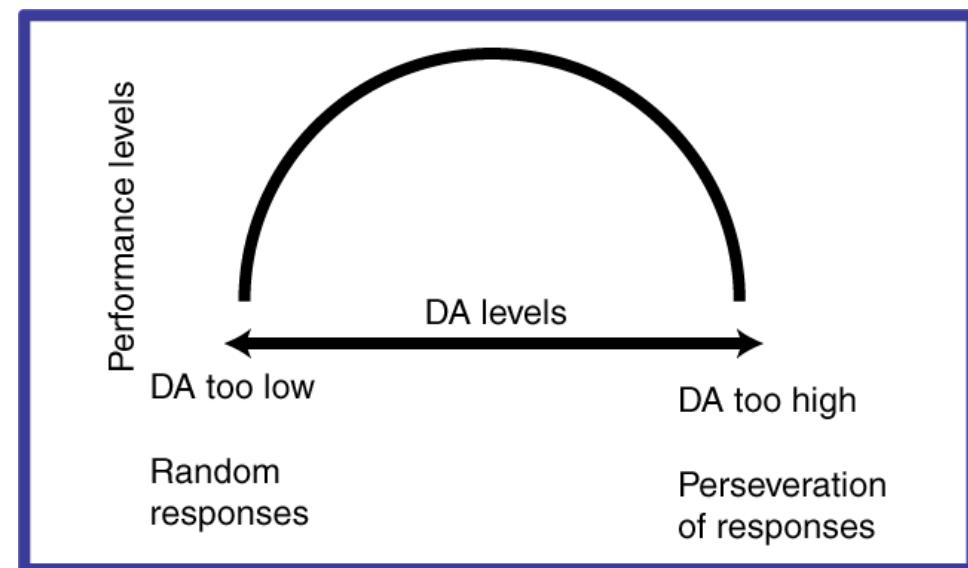
*Kentros et al. 2004*

***Tuning curve stability = stable categorization = stable basins of attraction.***

## The inverted U dopamine curve

Tonic DA (in prefrontal cortex) controls the perseveration of action choice.

### *“Failure modes”*



After Goldman-Rakic et al., 2000,  
Seamans and Yang, 2004.

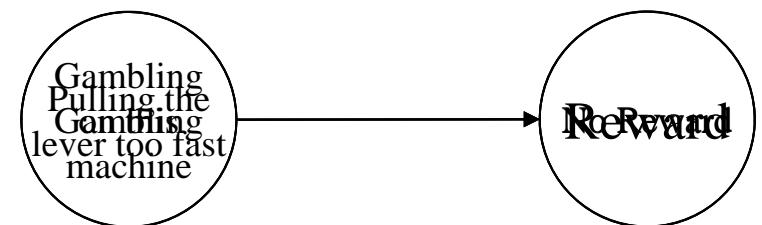
## Implications: Problem gambling and the illusion of control

### "Failure modes"

#### Hypothesis

Unexpected gains produce acquisition:  
 $\delta > 0 \Rightarrow$  store association

Unexpected losses produce state-splitting:  
 $\delta < 0 \Rightarrow$  transform state representation



#### Implication

A sequence of wins can produce anomalous expectations, that cannot be unlearned from subsequent losses.

#### This can explain

Hindsight bias in which losses are "explained away" (as a difference between s and s').

The illusion of control in which gamblers believe they can control statistical situations (by moving the state from s' to s).

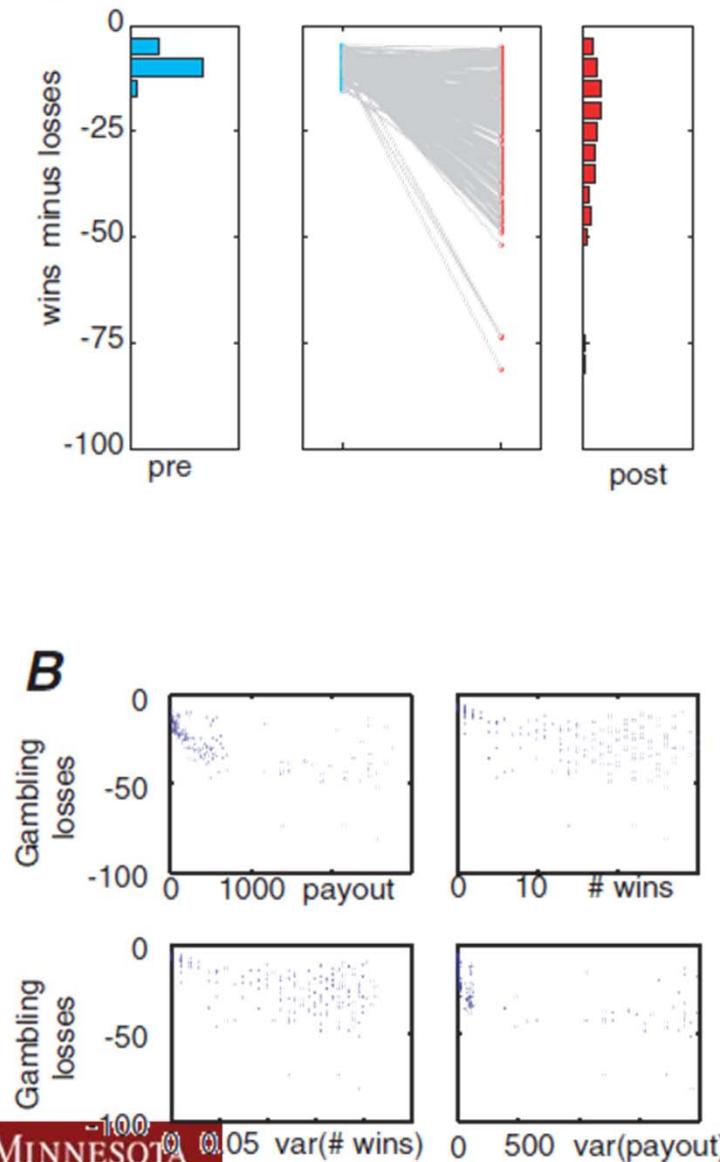
The path to problem gambling which begins with a statistically unlikely sequence of wins.

- Slot machines have gotten more complex, not less. Near misses drives one to find new ways to interpret problems.

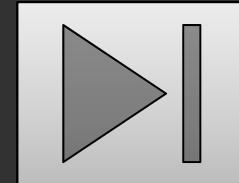
Langer and Roth, 1975. Custer, 1984. Wagenaar, 1988. Elster, 1999.

## *“Failure modes”*

### Simulations of problem gambling



# Implications for treatment



*State spaces matter ... precommitment*

frontiers in  
**BEHAVIORAL NEUROSCIENCE**

ORIGINAL RESEARCH ARTICLE  
published: 13 December 2010  
doi: 10.3389/fnbeh.2010.00184



A reinforcement learning model of precommitment in decision making

Zeb Kurth-Nelson and A. David Redish\*

frontiers in  
**NEUROSCIENCE**

FOCUSED REVIEW  
published: 08 October 2012  
doi: 10.3389/fnins.2012.00138



Don't let me do that! – models of precommitment

Zeb Kurth-Nelson<sup>1</sup> and A. David Redish<sup>2\*</sup>

*Contingency management*

frontiers  
in Psychiatry

HYPOTHESIS AND THEORY  
published: 01 June 2015  
doi: 10.3389/fpsyt.2015.00076

**Contingency management and deliberative decision-making processes**

Paul S. Regier<sup>1†</sup> and A. David Redish<sup>2\*</sup>

## Precommitment

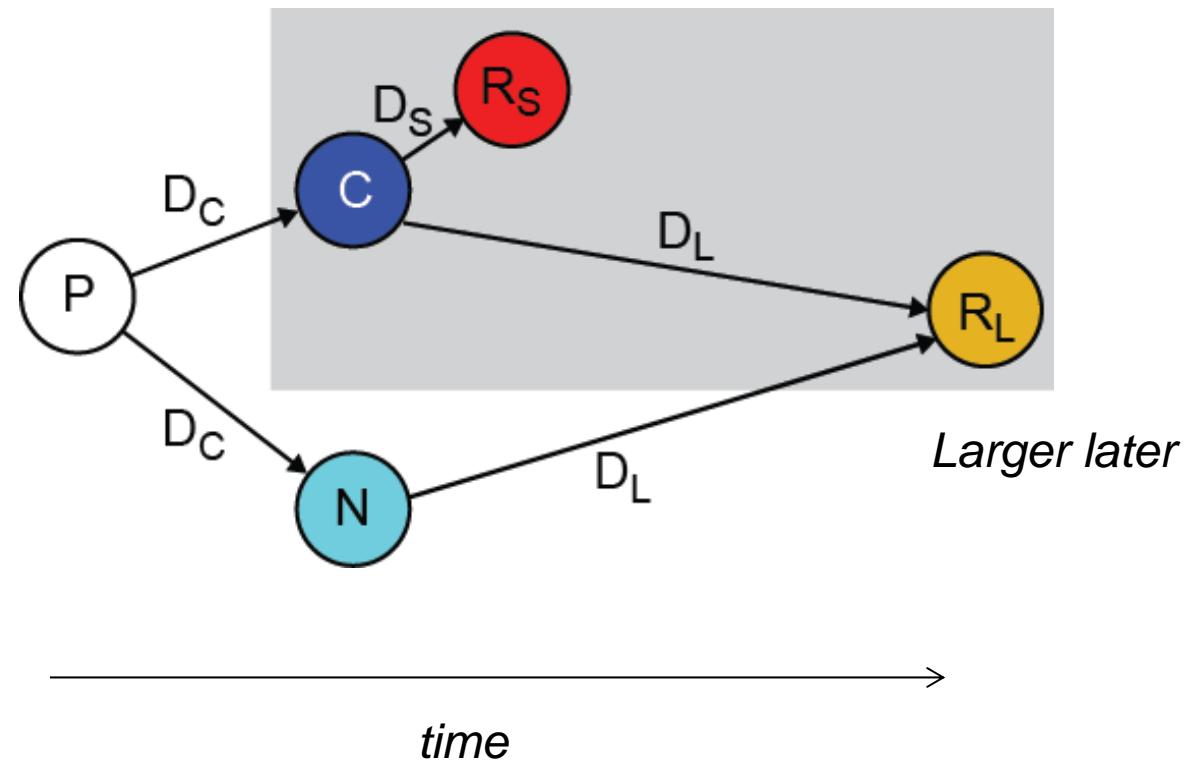
Precommitment entails rejecting a future choice to prevent an action.

This means the system has to prefer  $R_S \geq R_L$  but also to prefer  $N \geq C$ , which means  $R_S \geq R_L$  at  $T_{C/N}$  but also  $R_S \leq R_L$  at  $P$

## Implications for treatment

Zeb Kurth-Nelson, Redish (2010) [Frontiers in Behav Nsci](#)  
Zeb Kurth-Nelson, Redish (2012) [Frontiers in Nsci](#)

*Smaller sooner*



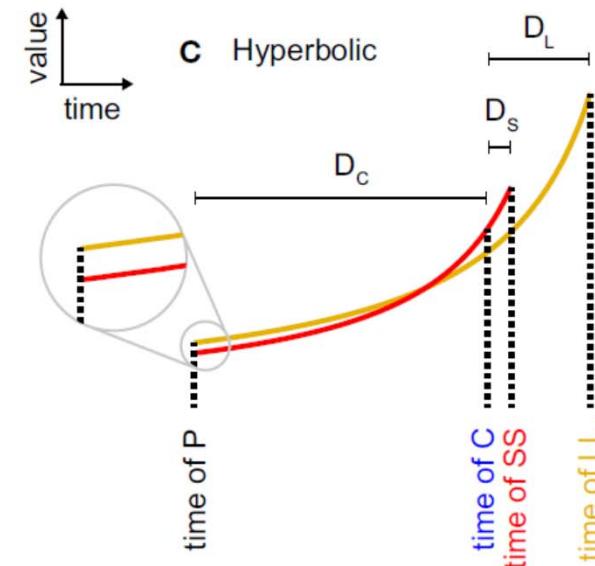
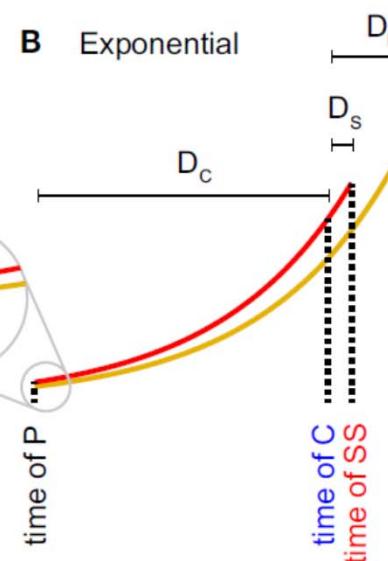
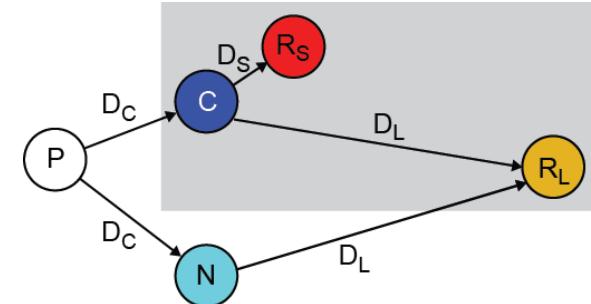
## Precommitment

Precommitment arises from non-exponential discounting functions.

### Implications for treatment

Zeb Kurth-Nelson, Redish (2010) Frontiers in Behav Nsci

Zeb Kurth-Nelson, Redish (2012) Frontiers in Nsci



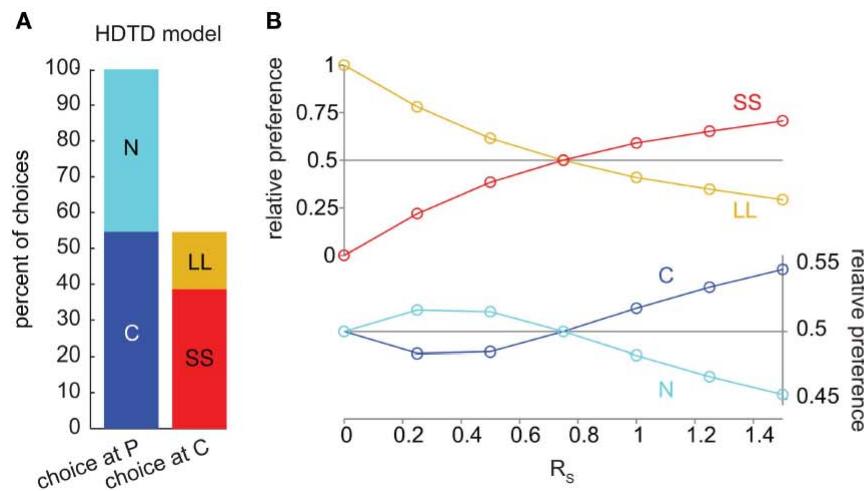
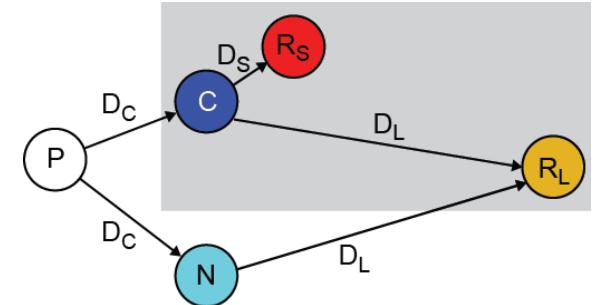
## Precommitment

But hyperbolic discounting  
is not enough.

## Implications for treatment

Zeb Kurth-Nelson, Redish (2010) Frontiers in Behav Nsci

Zeb Kurth-Nelson, Redish (2012) Frontiers in Nsci



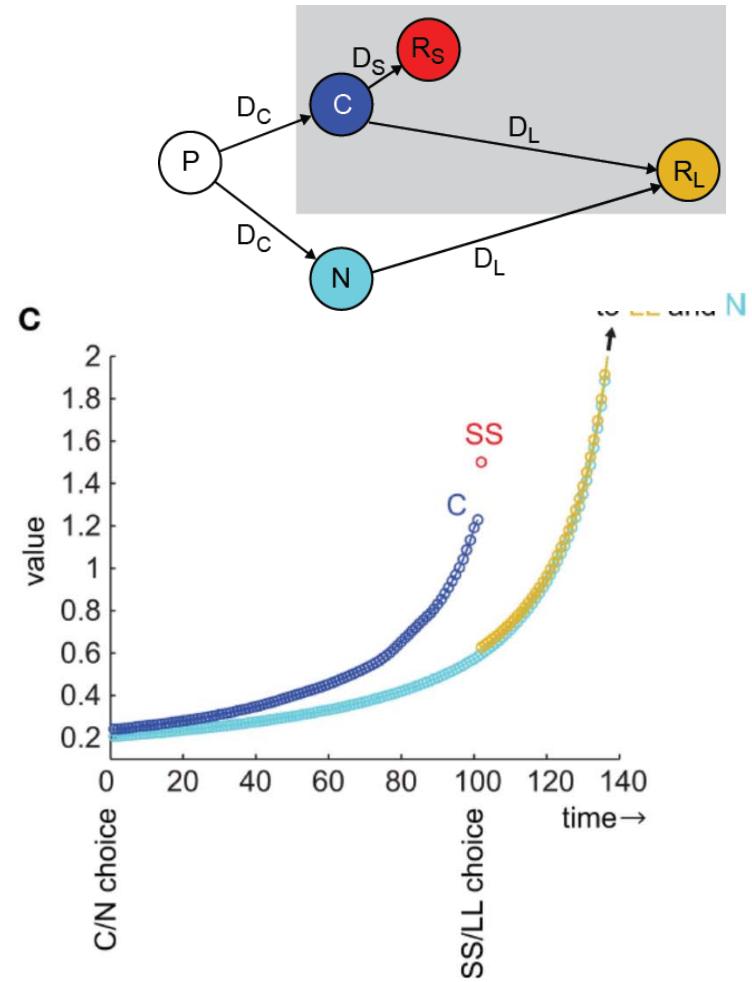
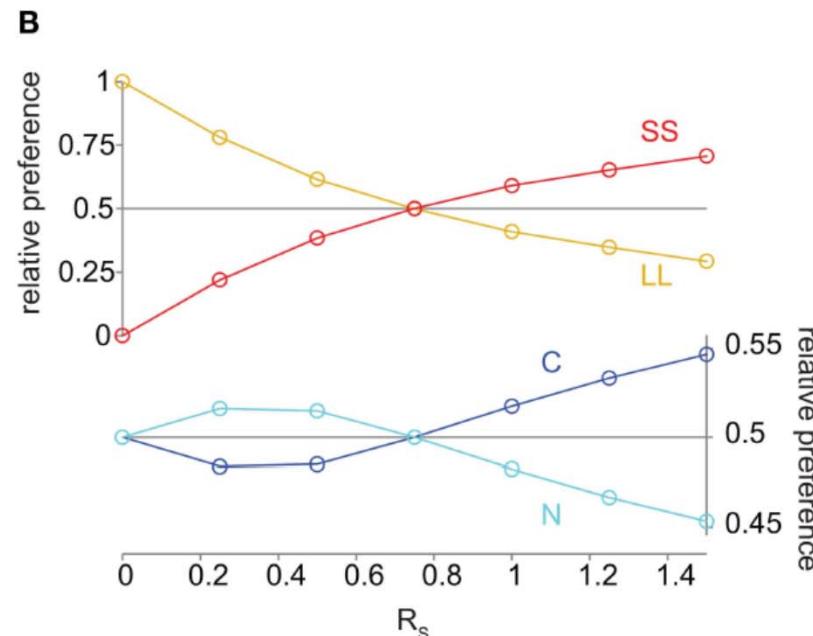
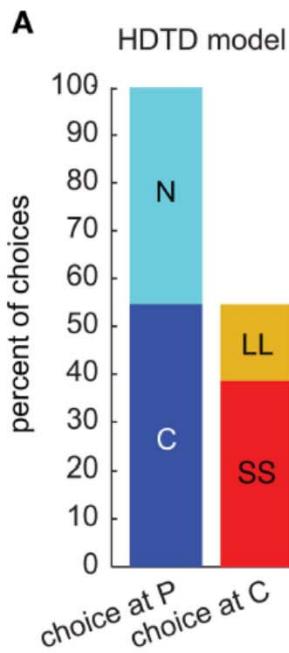
**Semi-markov model with hyperbolic discounting.**

## Precommitment

### Implications for treatment

Zeb Kurth-Nelson, Redish (2010) Frontiers in Behav Nsci  
 Zeb Kurth-Nelson, Redish (2012) Frontiers in Nsci

But hyperbolic discounting  
is not enough.



**HDTD model (Alexander and Josh Brown 2010).**

## Precommitment

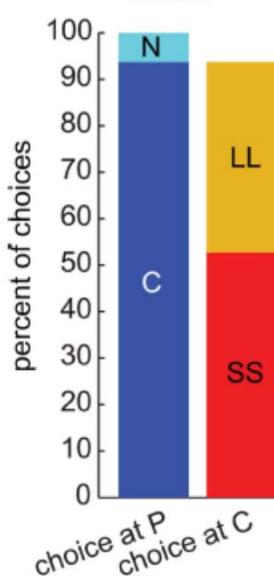
### Implications for treatment

Zeb Kurth-Nelson, Redish (2010) Frontiers in Behav Nsci

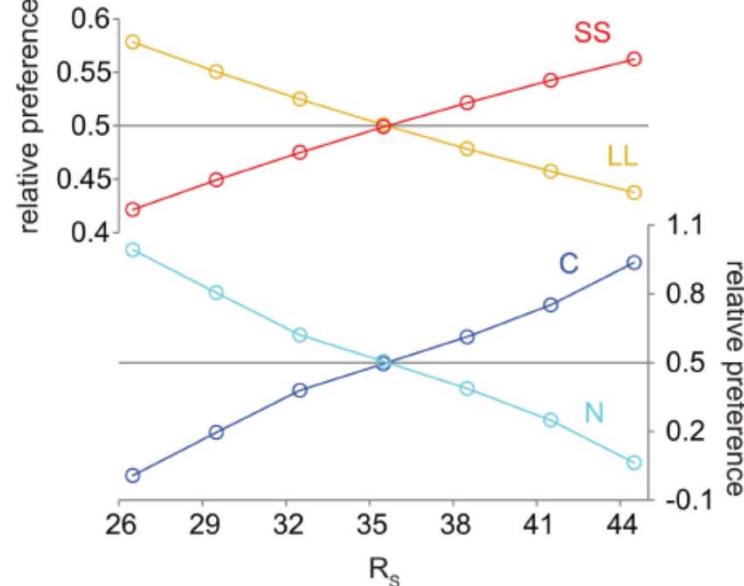
Zeb Kurth-Nelson, Redish (2012) Frontiers in Nsci

But hyperbolic discounting  
is not enough.

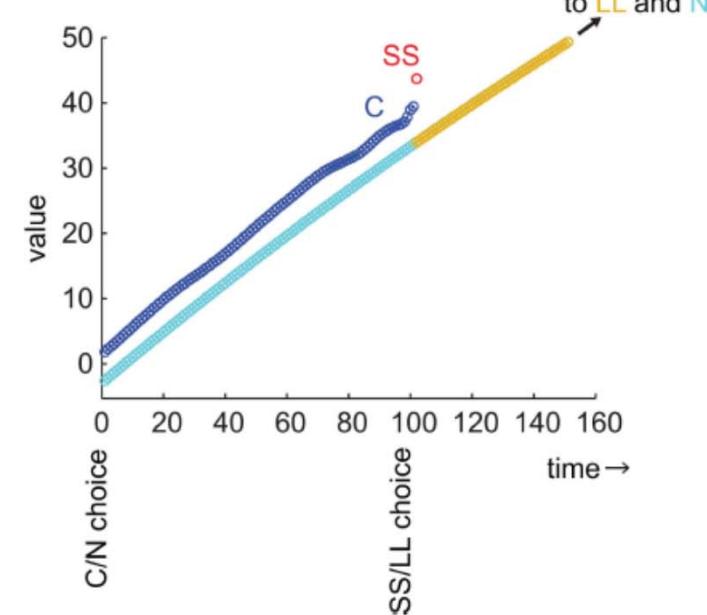
**A** Average reward model



**B** relative preference



**C**



**Average reward (hyperbolic) model (Daw and Touretzky 2000).**

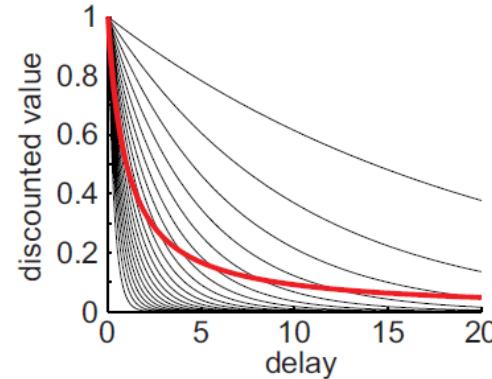
## Precommitment

Precommitment requires crossing value functions.

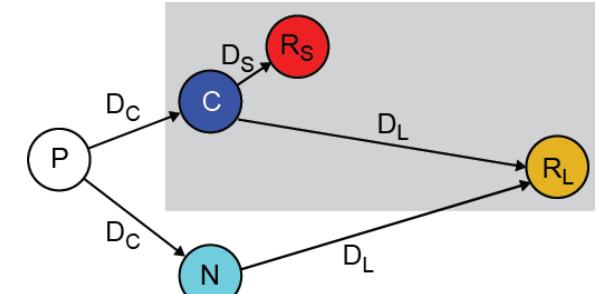
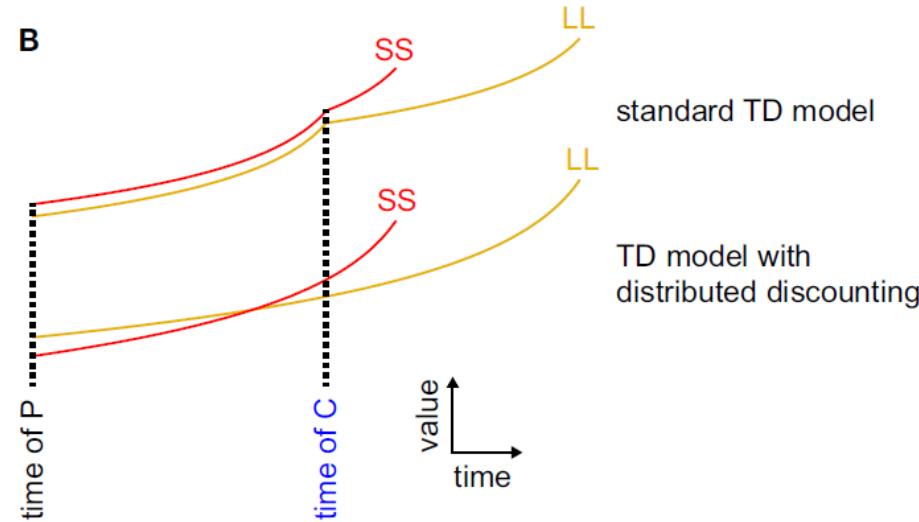
### Implications for treatment

Zeb Kurth-Nelson, Redish (2010) Frontiers in Behav Nsci  
Zeb Kurth-Nelson, Redish (2012) Frontiers in Nsci

A



B



## Precommitment

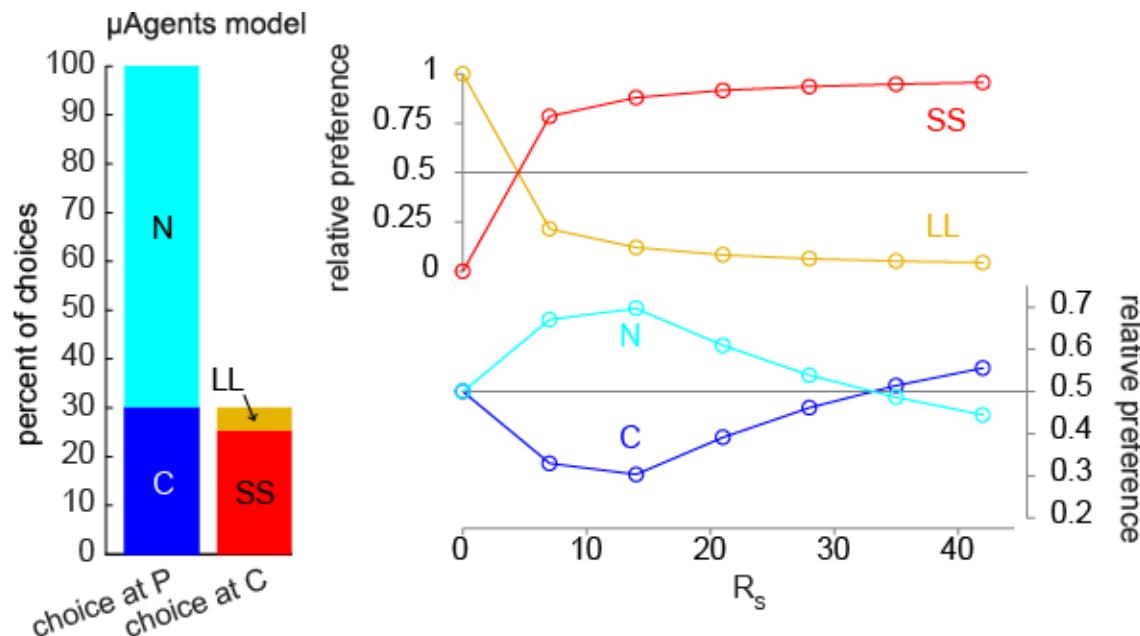
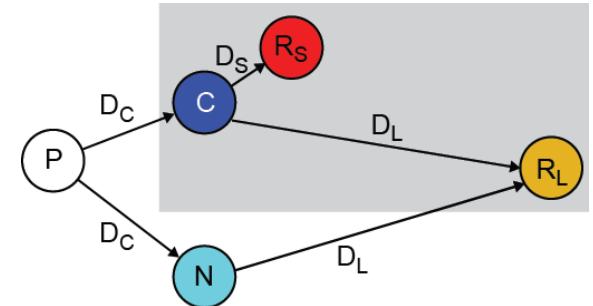
Precommitment entails rejecting a future choice to prevent an action.

Precommitment depends on having multiple value functions.

### Implications for treatment

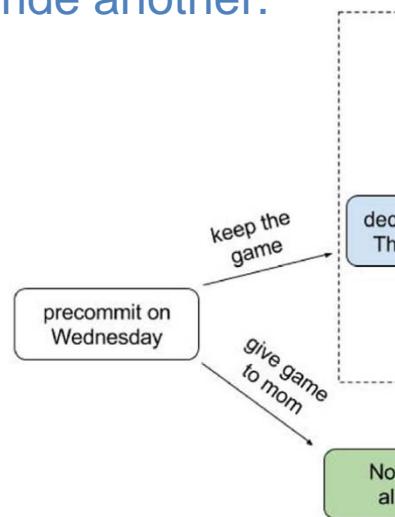
Zeb Kurth-Nelson, Redish (2010) Frontiers in Behav Nsci

Zeb Kurth-Nelson, Redish (2012) Frontiers in Nsci



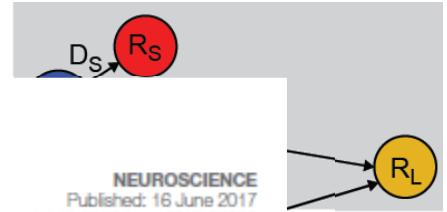
## Precommitment

Precommitment is a tool through which one decision system can override another.

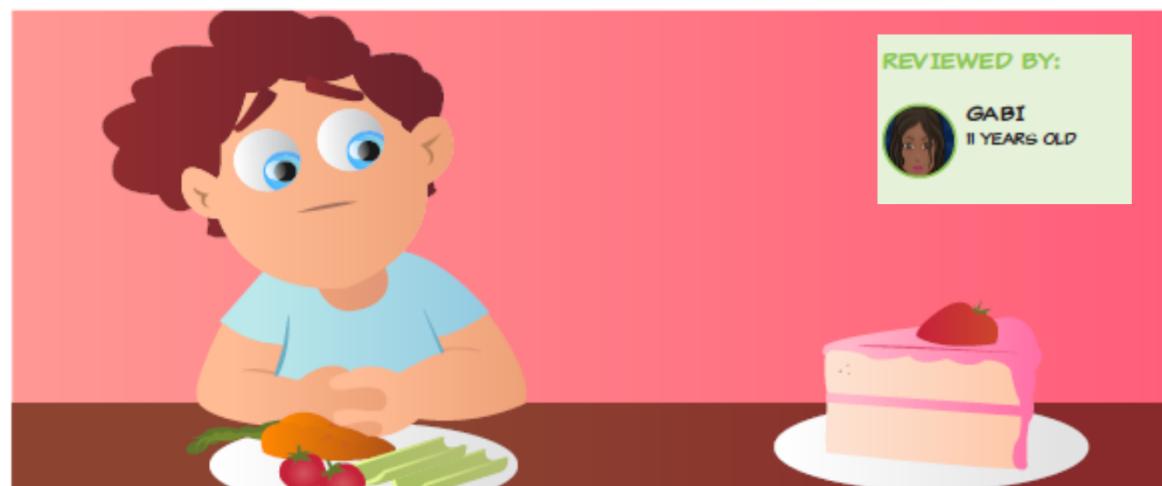


## Implications for treatment

Zeb Kurth-Nelson, Redish (2017) [Frontiers for young minds](#)



NEUROSCIENCE  
Published: 16 June 2017  
doi:10.3389/frym.2017.00026



## PRECOMMITMENT: A WAY AROUND TEMPTATION

Zeb Kurth-Nelson<sup>1,2</sup> and A. David Redish<sup>3\*</sup>



Together, they prefer studying on Wednesday, but prefer games on Thursday

# Implications for treatment

## Precommitment

frontiers in  
**BEHAVIORAL NEUROSCIENCE**

ORIGINAL RESEARCH ARTICLE  
published: 13 December 2010  
doi: 10.3389/fnbeh.2010.00184



A reinforcement learning model of precommitment in decision making

Zeb Kurth-Nelson and A. David Redish\*

frontiers in  
**NEUROSCIENCE**

FOCUSED REVIEW  
published: 08 October 2012  
doi: 10.3389/fnins.2012.00138



Don't let me do that! – models of precommitment

Zeb Kurth-Nelson<sup>1</sup> and A. David Redish<sup>2\*</sup>

## Contingency management

frontiers  
in Psychiatry

HYPOTHESIS AND THEORY  
published: 01 June 2015  
doi: 10.3389/fpsyt.2015.00076

Contingency management and deliberative decision-making processes

Paul S. Regier<sup>1†</sup> and A. David Redish<sup>2\*</sup>

## Nudging behavior

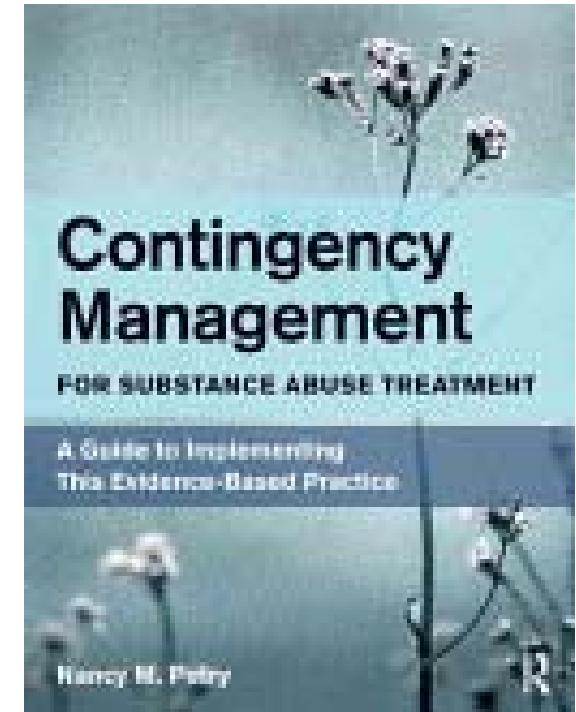
If you don't use drugs for a week (come in clean), then you receive a small reward.

### **Current theory:**

- The reward is an **alternate reinforcer**.
- Losing it increases the **opportunity costs** of the drug.

### **Implications for treatment**

Paul Regier, Redish (2015) *Frontiers in Psychiatry*

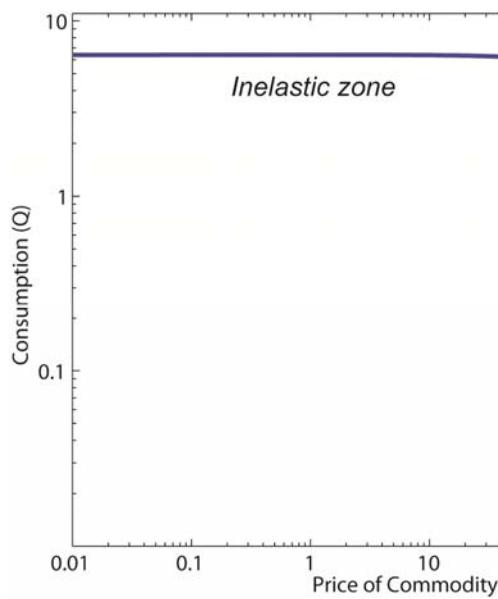


*But the rewards are small.*

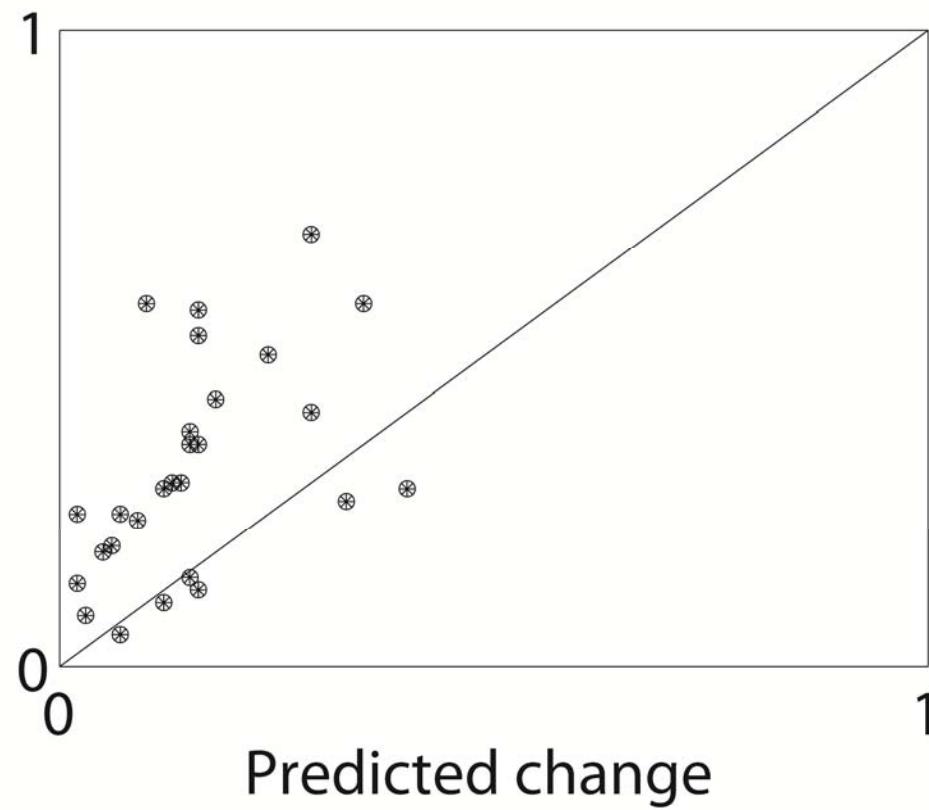
Increasing the cost of the drug on the street by that amount has little to no effect.

## Contingency management works too well

We measured the demand curves from a typical study and compared them to the effects of contingency management.

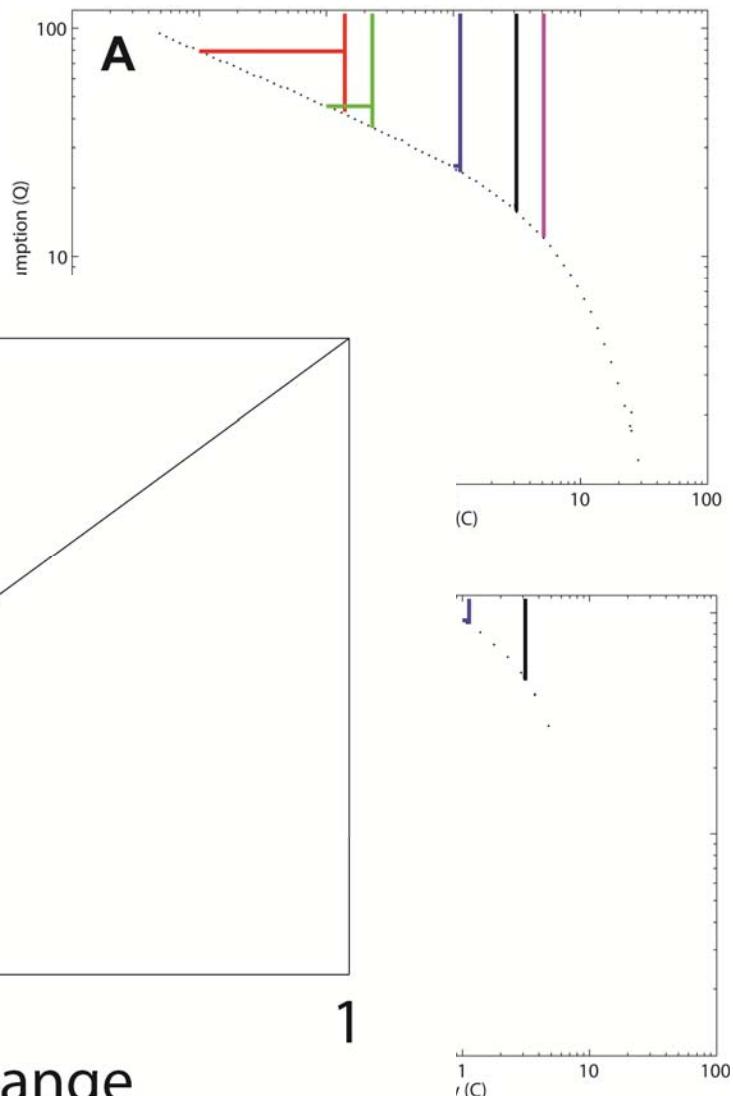


Observed effect size



## Implications for treatment

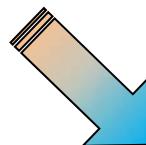
Paul Regier, Redish (2015) *Frontiers in Psychiatry*



## Nudging behavior

If you don't use drugs for a week (come in clean), then you receive a small reward.

**Is it worth it?**



**Which one?**

**How can we strengthen deliberation?**

**We can test for prefrontal-hippocampal integrity.**

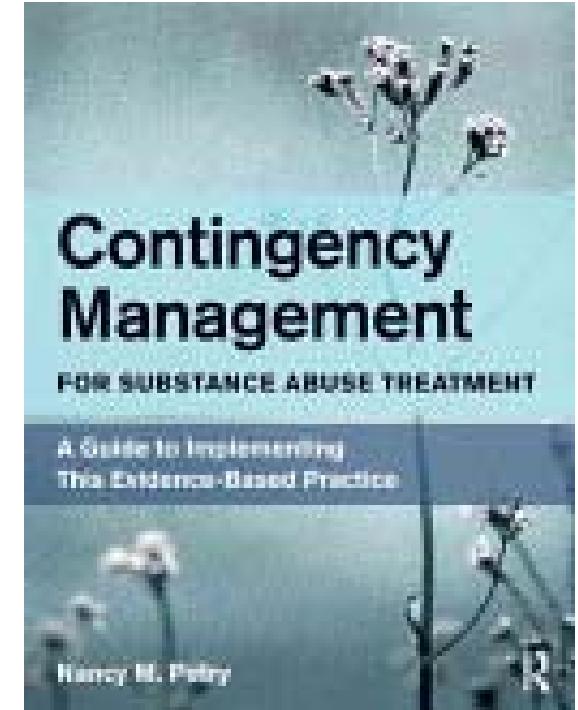
**We can train working memory.**

**We can make the second option more concrete.**

**We can provide reminders.**

## Implications for treatment

Paul Regier, Redish (2015) *Frontiers in Psychiatry*



## Review

We started with a discussion of failure modes...

The concept of the "failure mode"

In engineering, a "failure mode" is a vulnerability inherent in the machinery.

"Failure modes"



35W Bridge Collapse, 2007. startribune.com

UNIVERSITY OF MINNESOTA

## Review

We started with a discussion of failure modes...

and did a deep dive into the non-compensable dopamine model.

What if dopamine **is** delta?



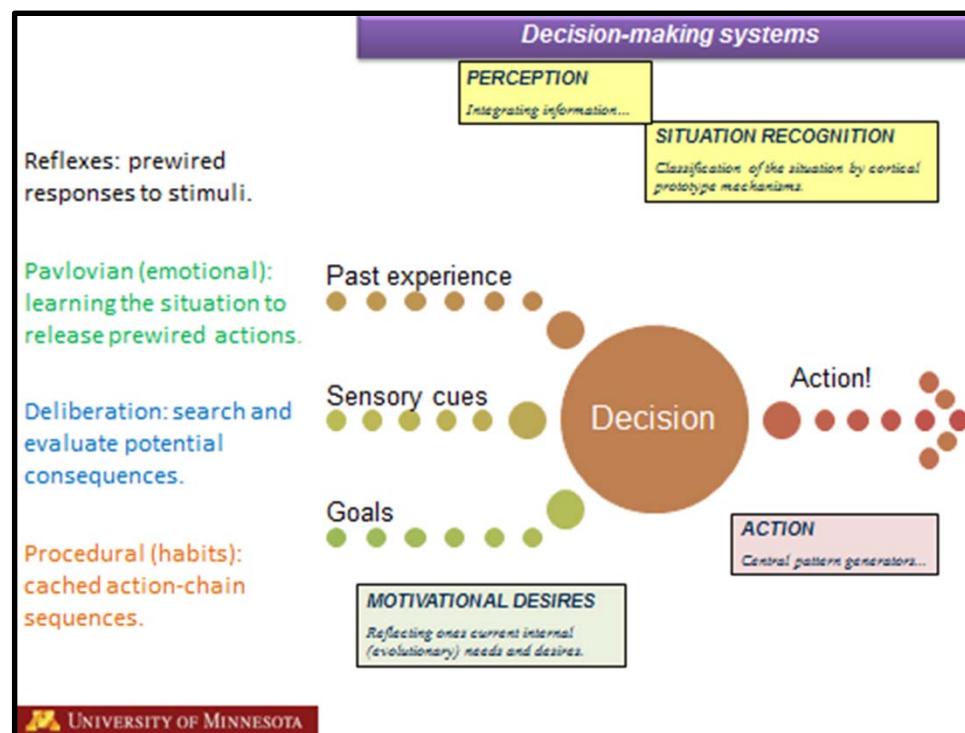
$$\delta = \max(\gamma^d(R(S_l) + V(S_l)) - V(S_k) + D(S_l), D(S_l))$$

# Review

We started with a discussion of failure modes...

and did a deep dive into the non-compensable dopamine model.

We looked at failure modes across the decision-making system



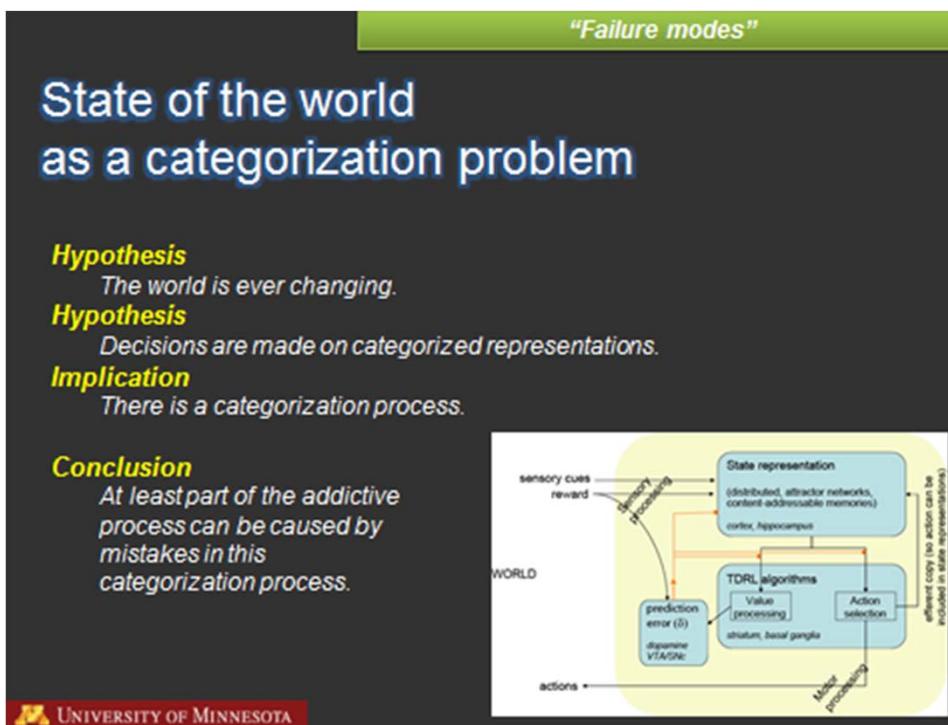
# Review

We started with a discussion of failure modes...

and did a deep dive into the non-compensable dopamine model.

We looked at failure modes across the decision-making system

and did a deep dive into situation recognition.



## Review

We started with a discussion of failure modes...

and did a deep dive into the non-compensable dopamine model.

We looked at failure modes across the decision-making system

and did a deep dive into situation recognition.

And ended with implications for treatment

# Review

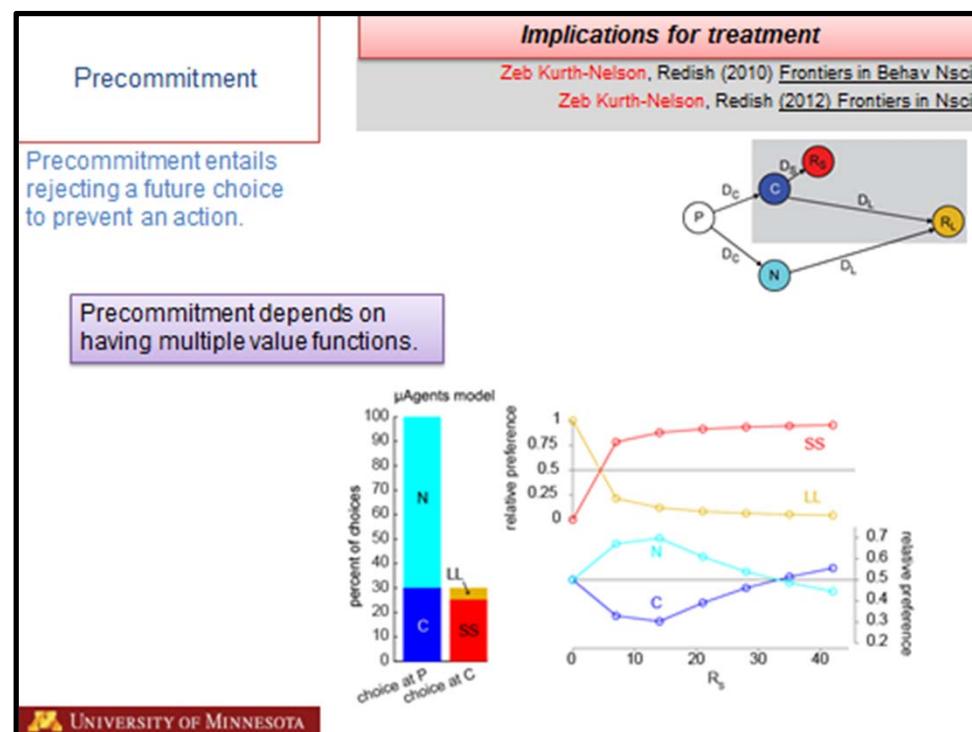
We started with a discussion of failure modes...

and did a deep dive into the non-compensable dopamine model.

We looked at failure modes across the decision-making system

and did a deep dive into situation recognition.

And ended with implications for treatment including a dive into precommitment



## Review

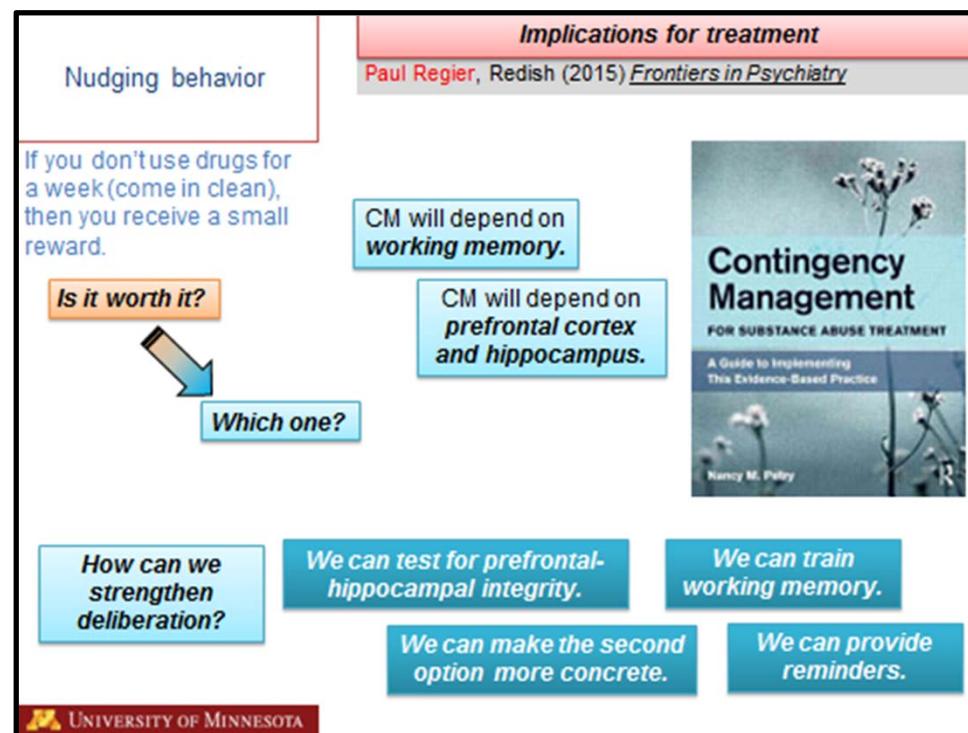
We started with a discussion of failure modes...

and did a deep dive into the non-compensable dopamine model.

We looked at failure modes across the decision-making system

and did a deep dive into situation recognition.

And ended with implications for treatment including a dive into precommitment and into contingency management.

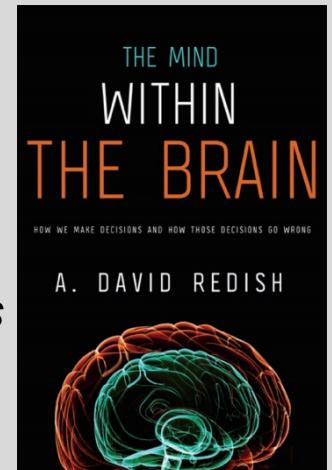


# Computational psychiatry

*taking an engineer's view*

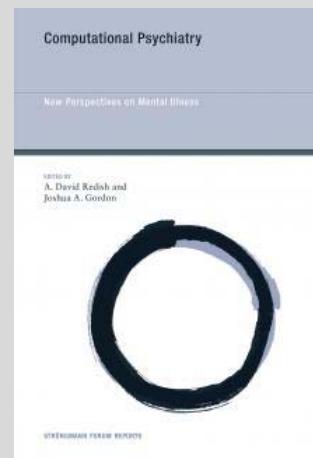
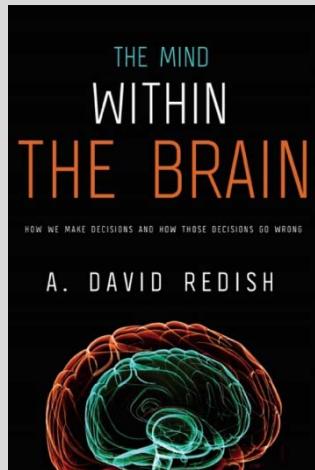
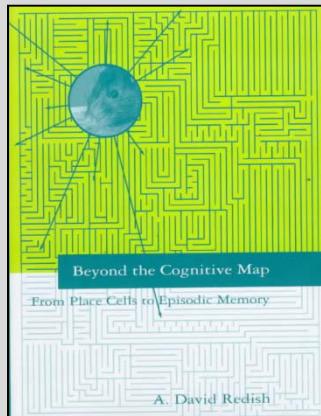
**To all students in the program:**

1. All of our papers are available on the laboratory website  
[redishlab.neuroscience.umn.edu](http://redishlab.neuroscience.umn.edu)
2. Those of you interested in the “decision making” story, the story is told with extensive citations in the Mind within the Brain book.
3. We didn’t talk about it, but I also recommend the report from the Strungmann Forum with excellent papers by a number of researchers, **including both our hosts**.
4. Finally, I am currently **looking for postdocs**, particularly people who want to work at the boundary between computation and experiment.
5. Feel free to send me email with any questions or discussions  
[redish@umn.edu](mailto:redish@umn.edu)



# Computational models of addiction

A. David Redish  
Department of Neuroscience  
University of Minnesota  
[redish@umn.edu](mailto:redish@umn.edu)  
<http://umn.edu/~redish>



RedishLab at the University of Minnesota  
<http://redishlab.neuroscience.umn.edu>

