

Neuroeconomics: **From The Failures of Expected Utility to the Neurobiology of Choice**

Paul Glimcher PhD





Julius Silver Professor of Neural Science, Economics and Psychology
Director, Institute of the Study of Decision Making
New York University

Blaise Pascal



Genius:

Expected Value Theory

	Probability x Value =		Expected Value
#1	0.5	100 	50 
#2	1.0	45 	45 

Pascal's Wager				
	If God Exists		If God Doesn't Exist	
	(Prob x Value)	+	(Prob x Value)	= Exp. Value
Believe in God	$>0 \times \infty$	+	$\geq 0 \times 0$	= ∞
Do not Believe in God	$>0 \times -\infty$	+	$\geq 0 \times 0$	= $-\infty$

Problem:

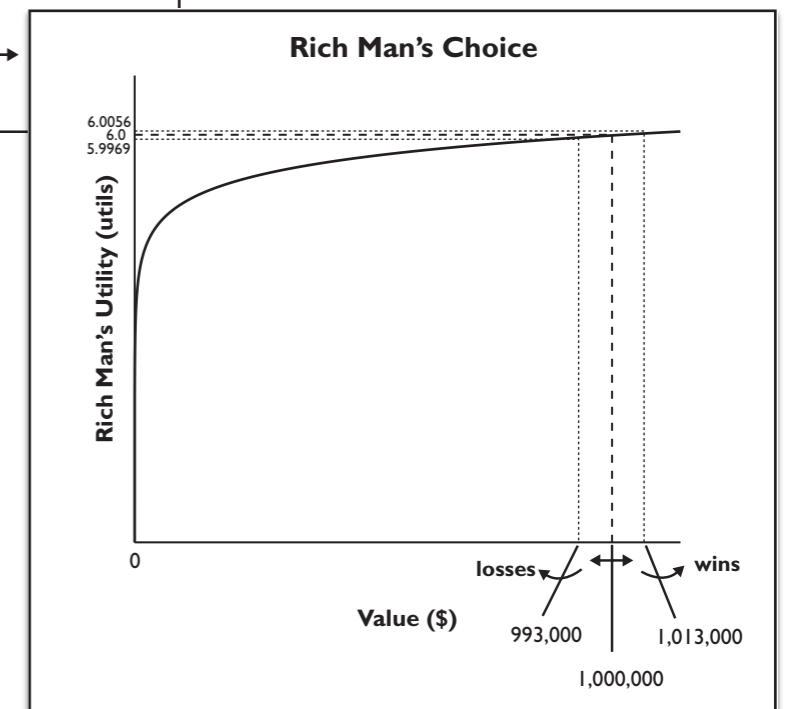
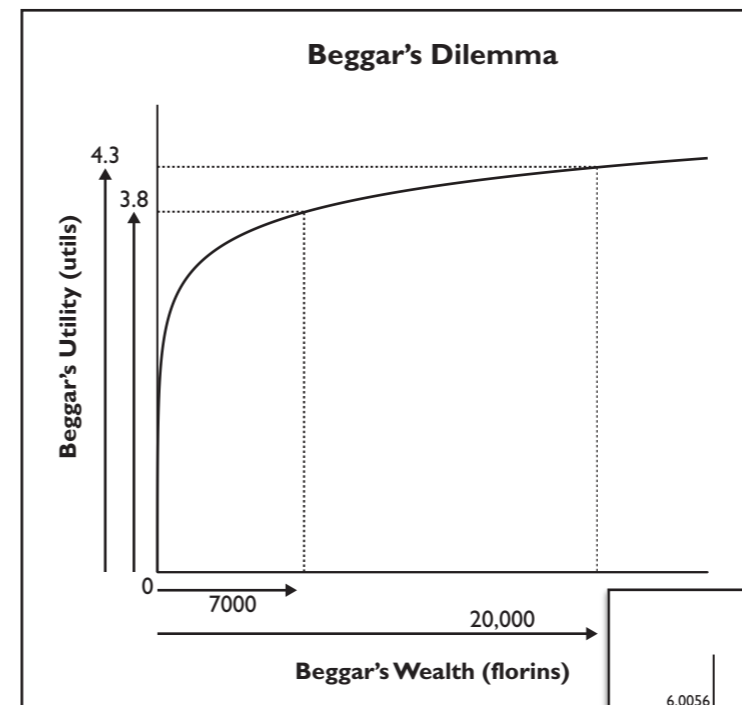
The Beggar's Dilemma

Daniel Bernoulli



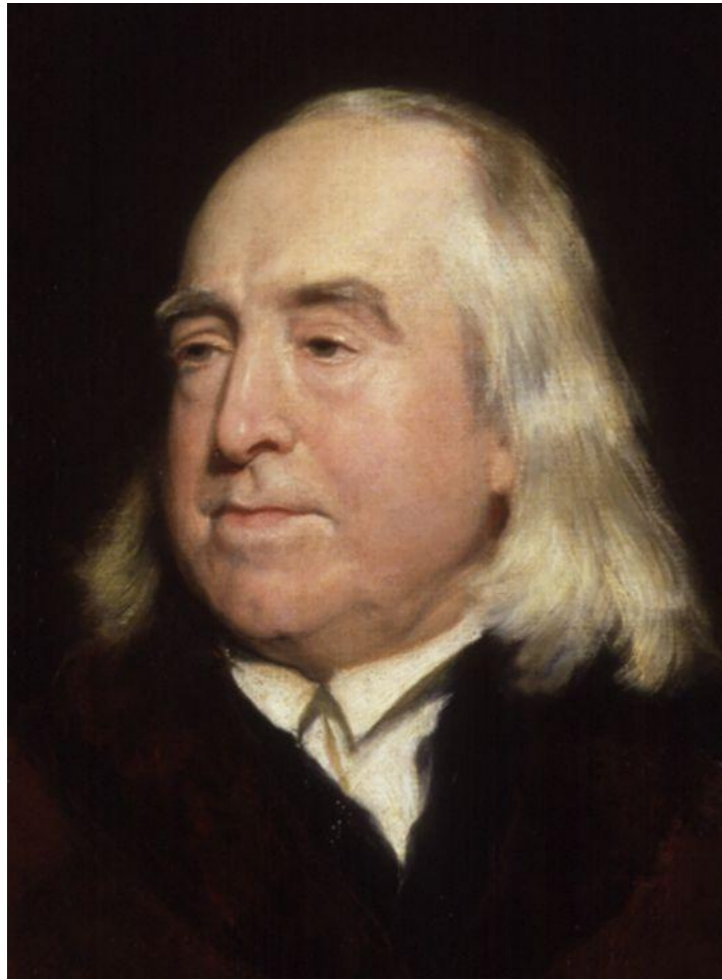
Genius:

Expected *Utility* Theory



Problem: Bentham, Pareto, Samuelson

Jeremy Bentham



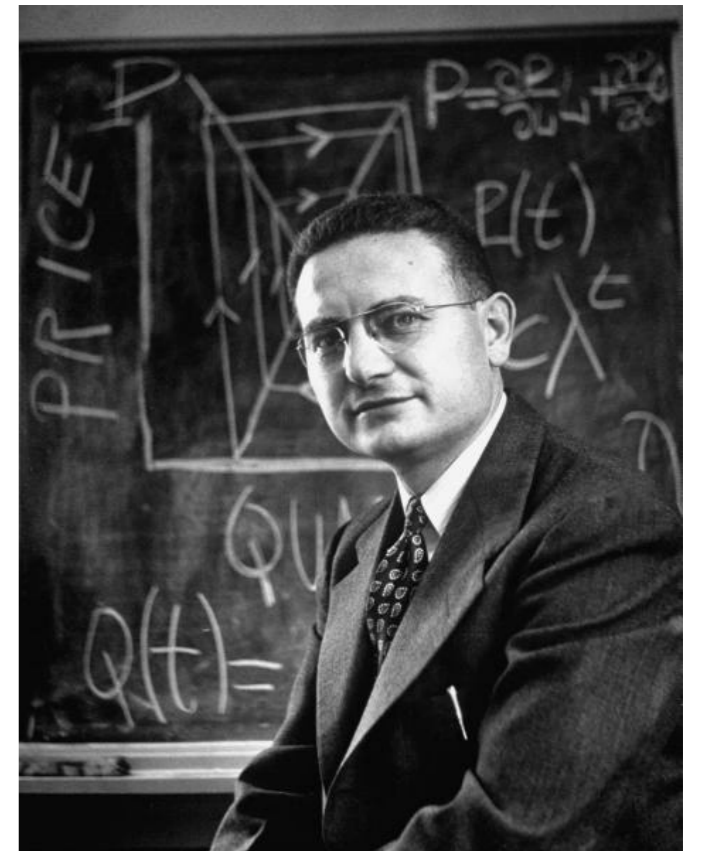
The Calculus of Utility

Vilfredo Pareto



The Intrinsic
Arbitrariness of Utility

Paul Samuelson



Ordinal Objective
Utility

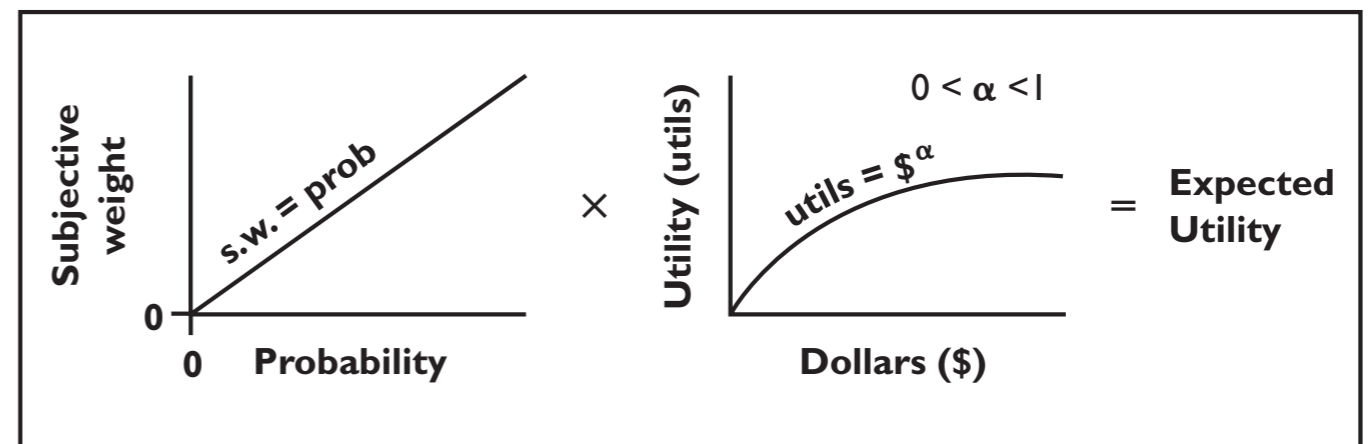
John von Neumann Oskar Morgenstern



Genius:

Modern

Expected ***Utility*** Theory



Critical Advantages:

- Precise
- Compact
- **Normative** (people make sense)

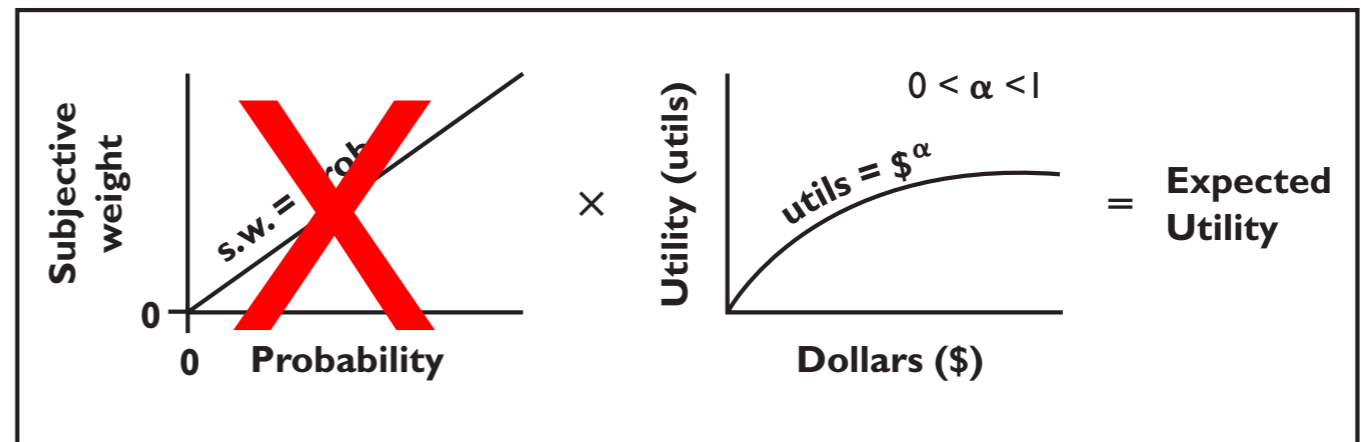
Problem:

Maurice Allais

Maurice Allais



People Do Not Obey EU
all the time



Critical Questions:

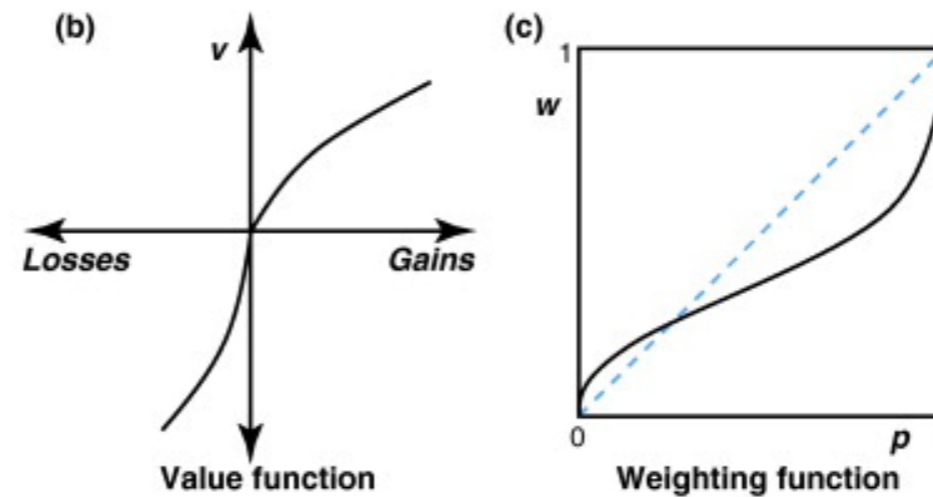
- How to Predict People?
- Are People Dumb?
- **Why?**

Amos Tversky

Danny Kahneman



Prospect Theory



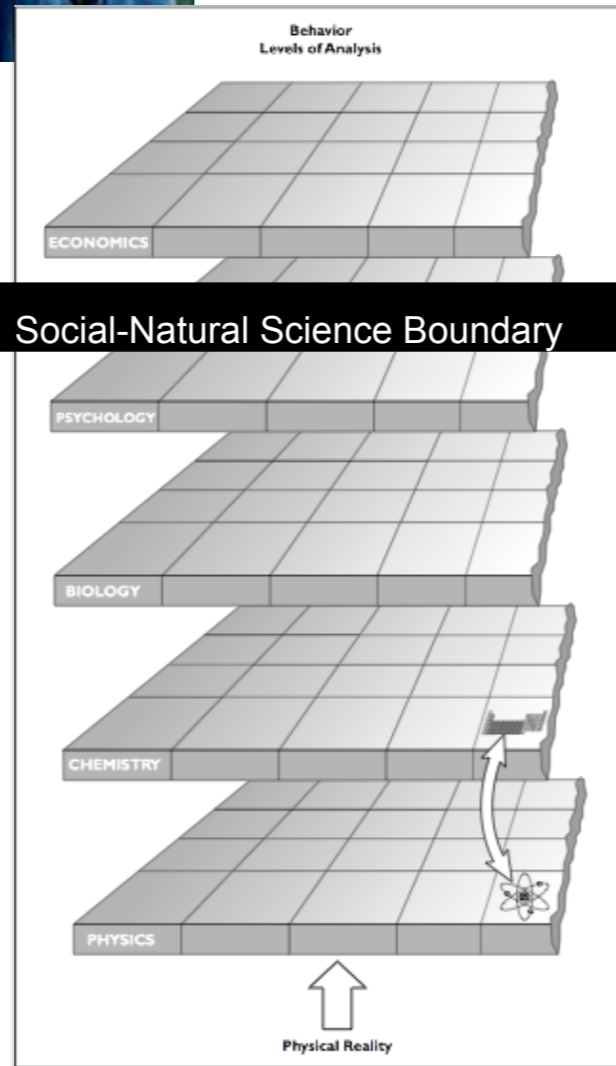
Critical Advantages:

- Predictive

Critical Disadvantages:

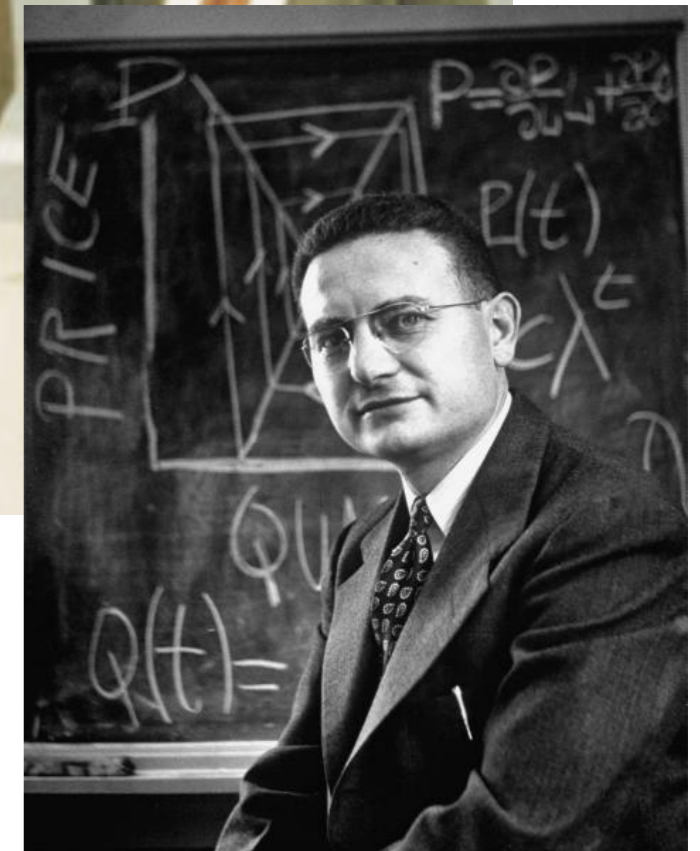
- Bulky
- No Why

Behavioral



**Economics
vs.
Psychology**

Traditional



Behavioral



Traditional



The Reductive Levels Of Decision Science

Behavior
Levels of Analysis

Samuelson

ECONOMICS

Social-Natural Science Boundary

Kahneman

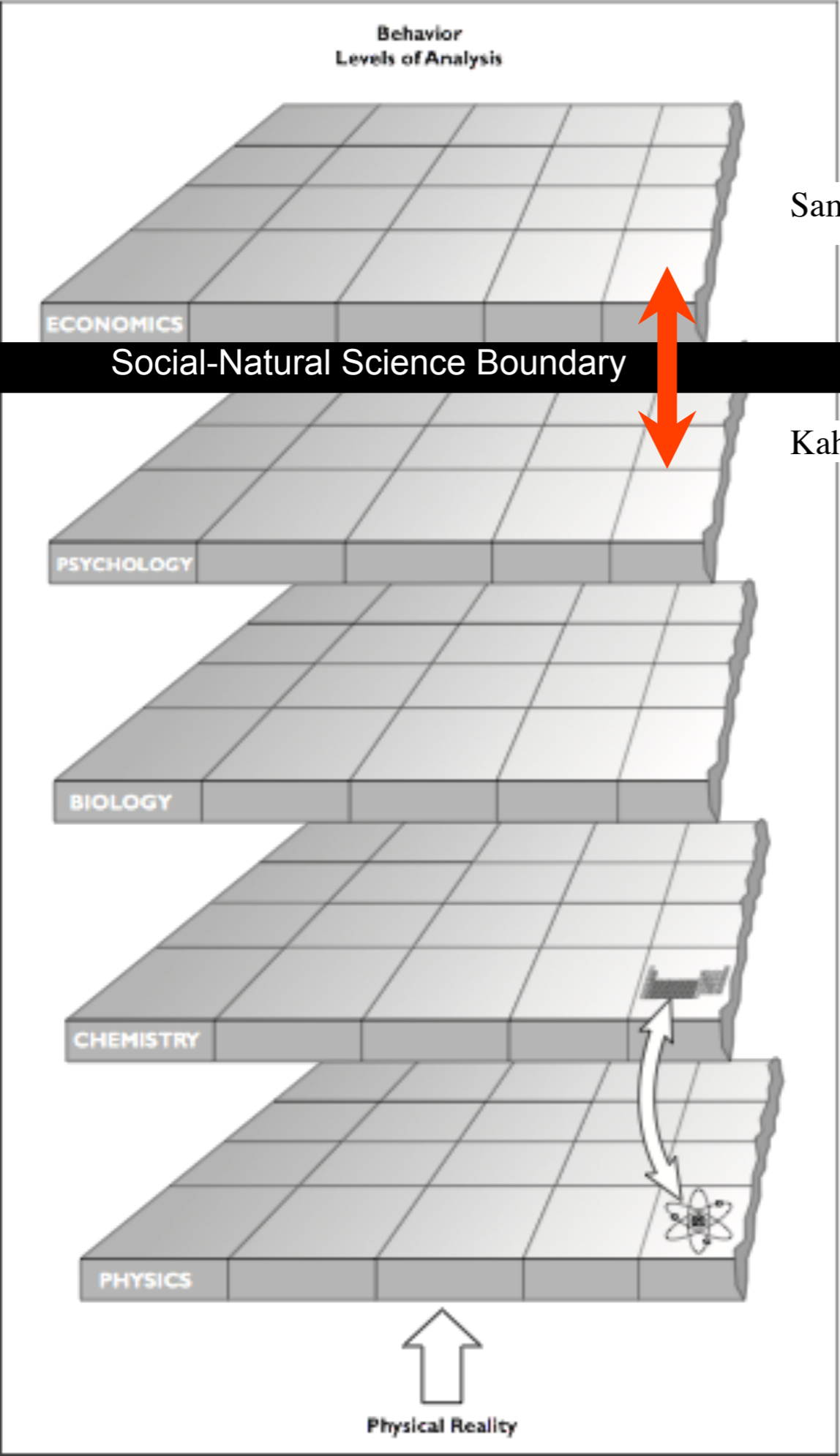
PSYCHOLOGY

BIOLOGY

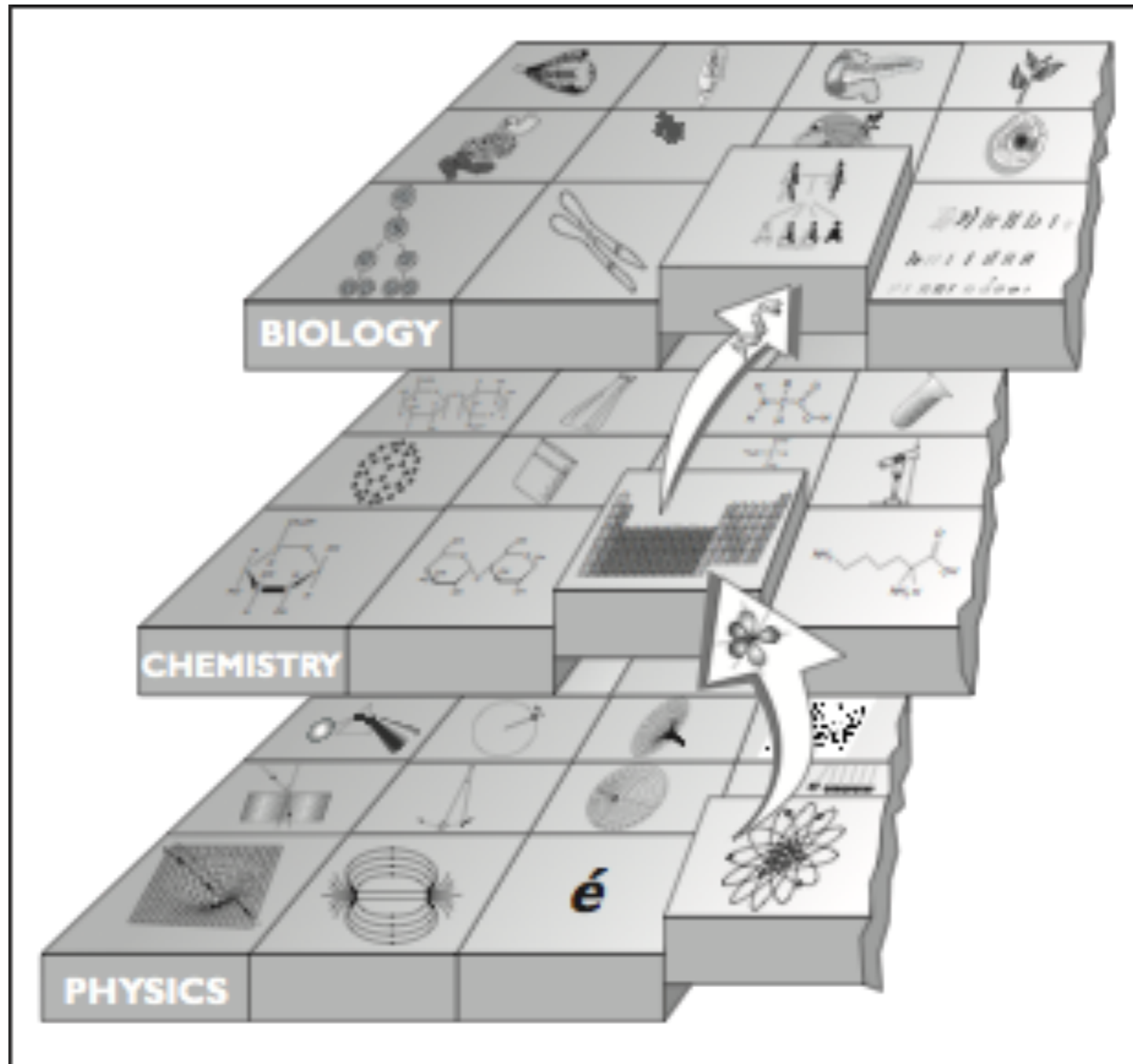
CHEMISTRY

PHYSICS

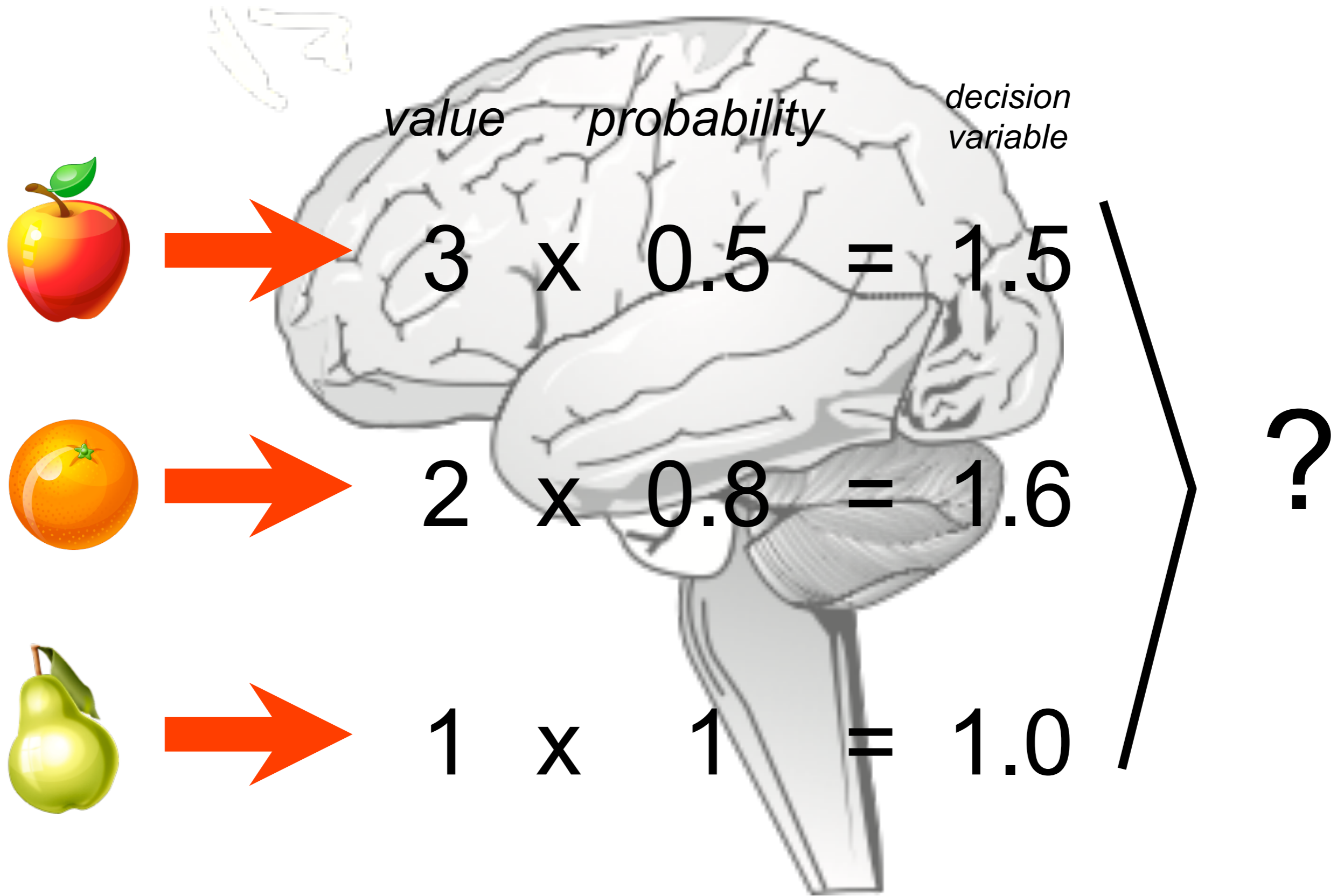
Physical Reality



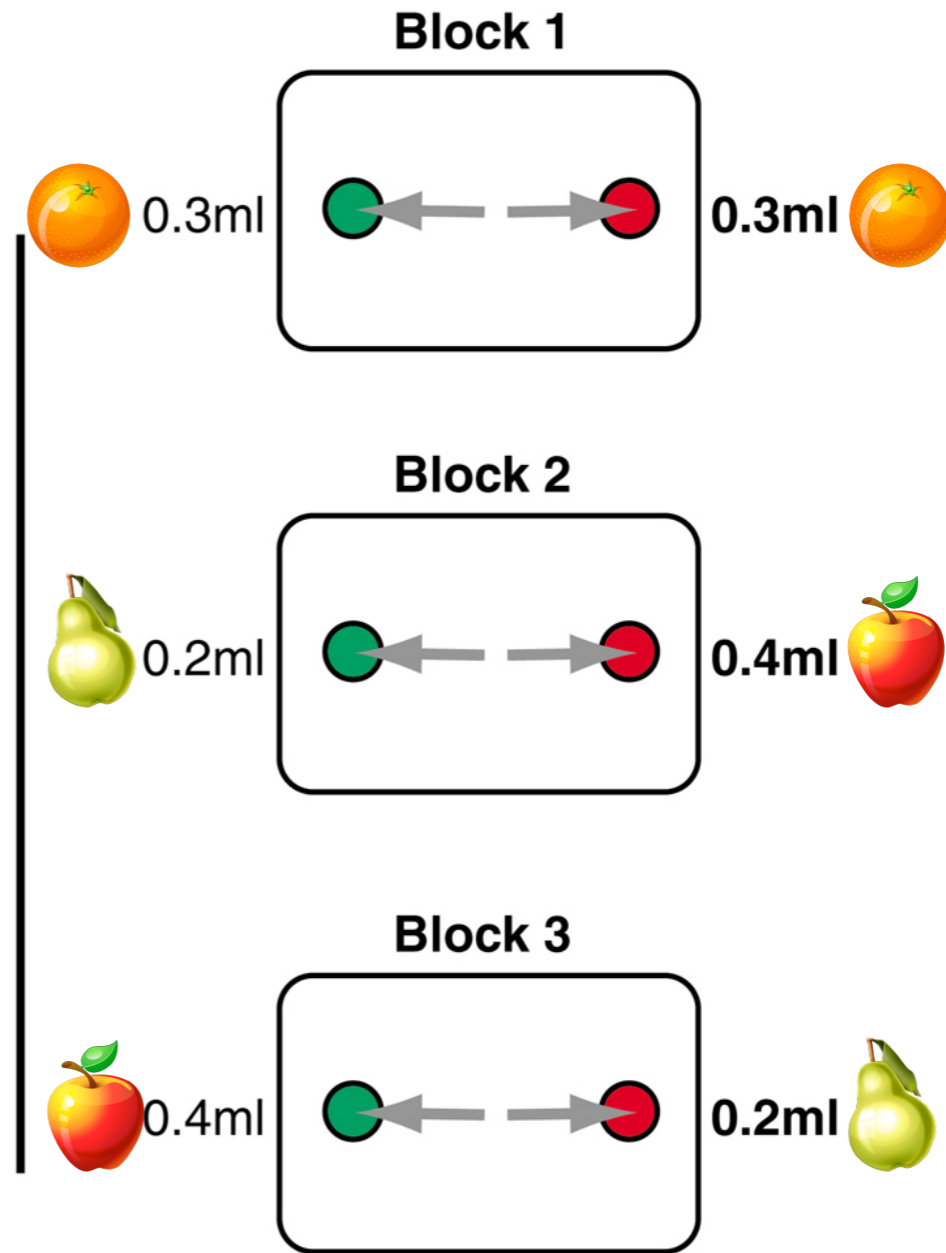
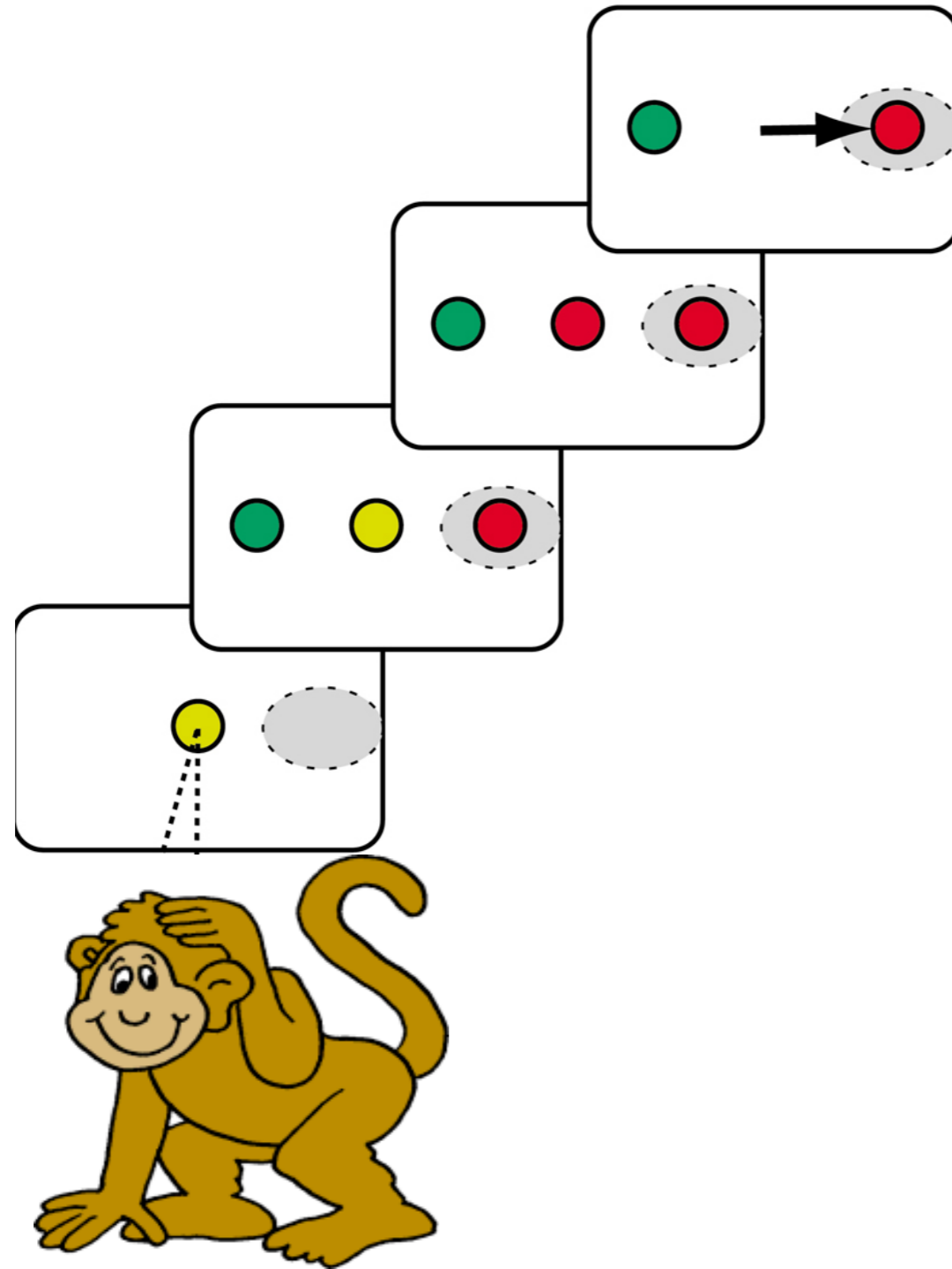
The Central Goal of Neuroeconomics *(for me)* was/is Interdisciplinary Fusion



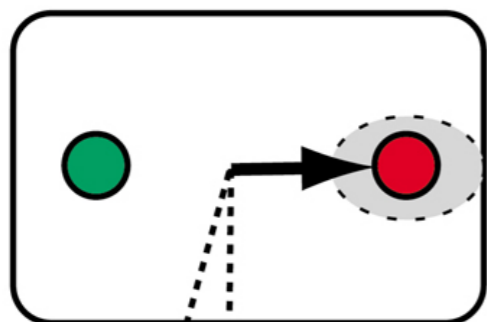
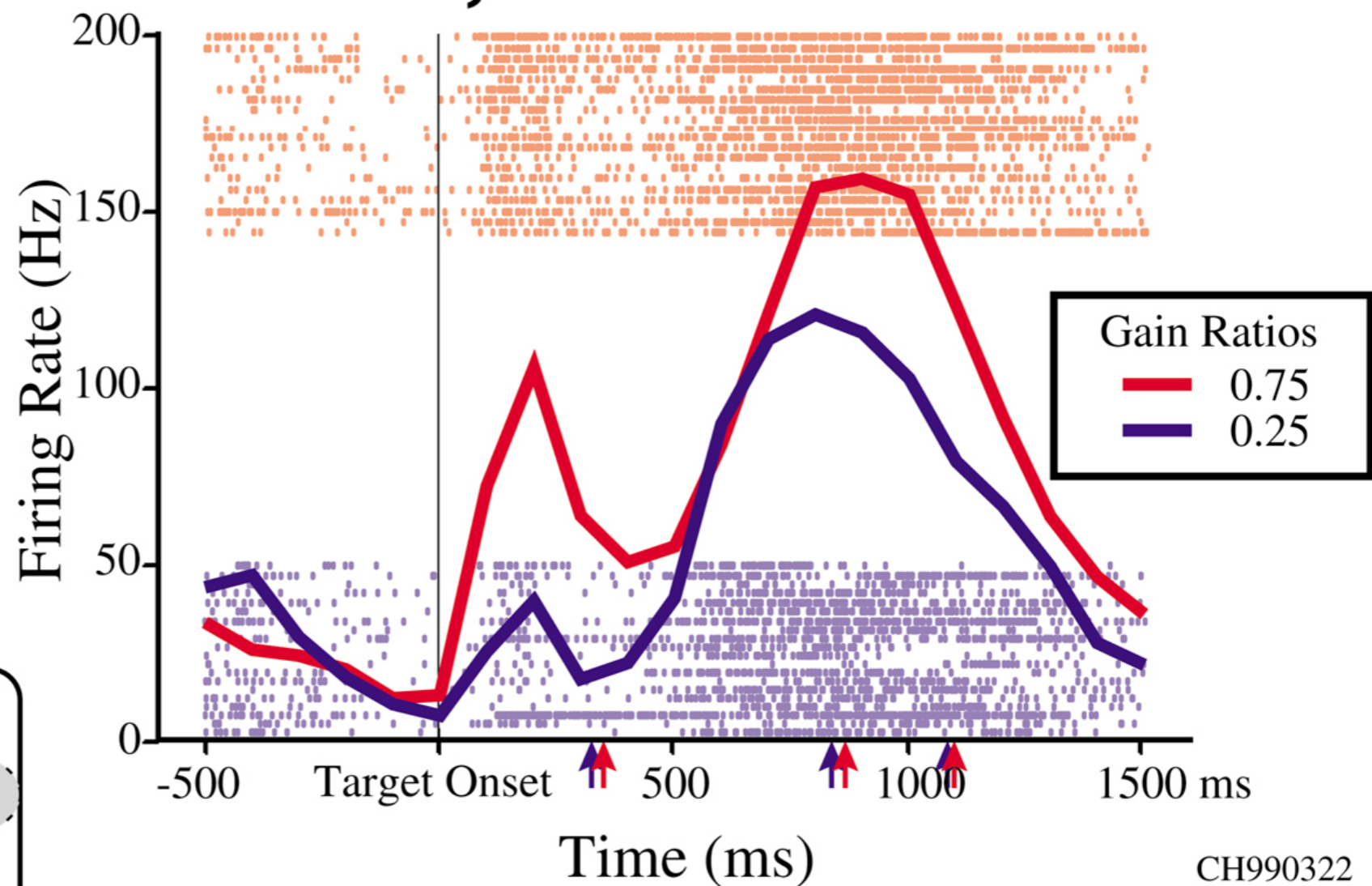
How Can We Combine These?



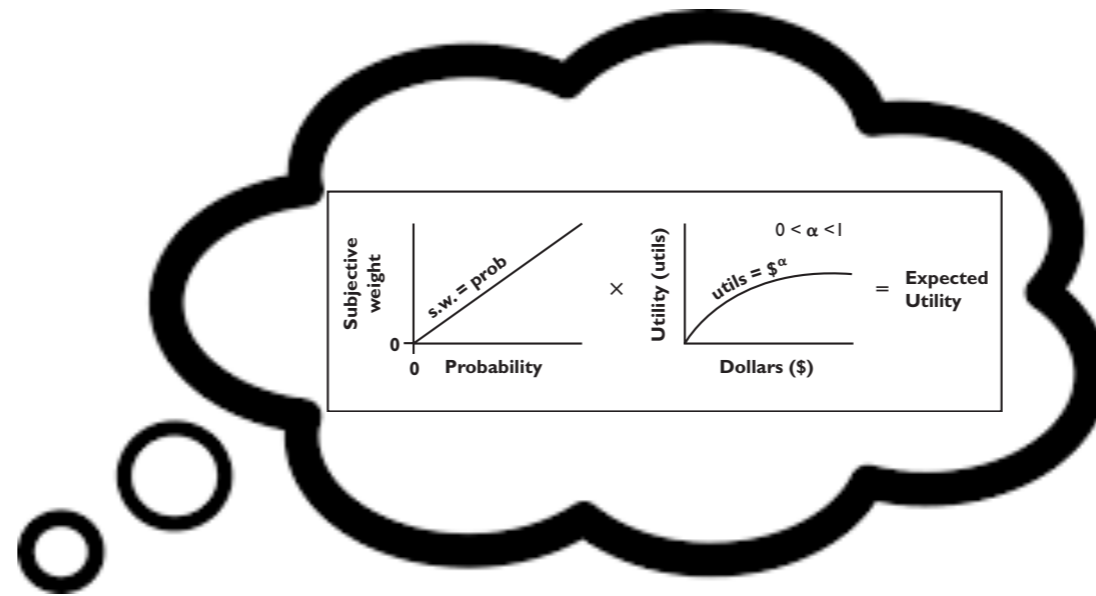
So What Happens When We Vary Value?

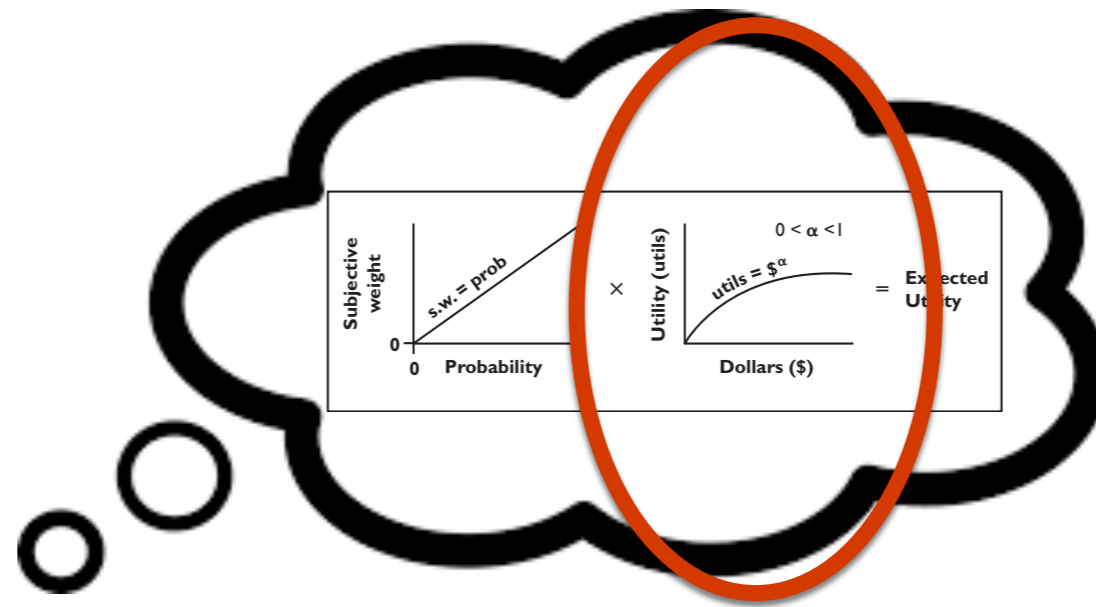


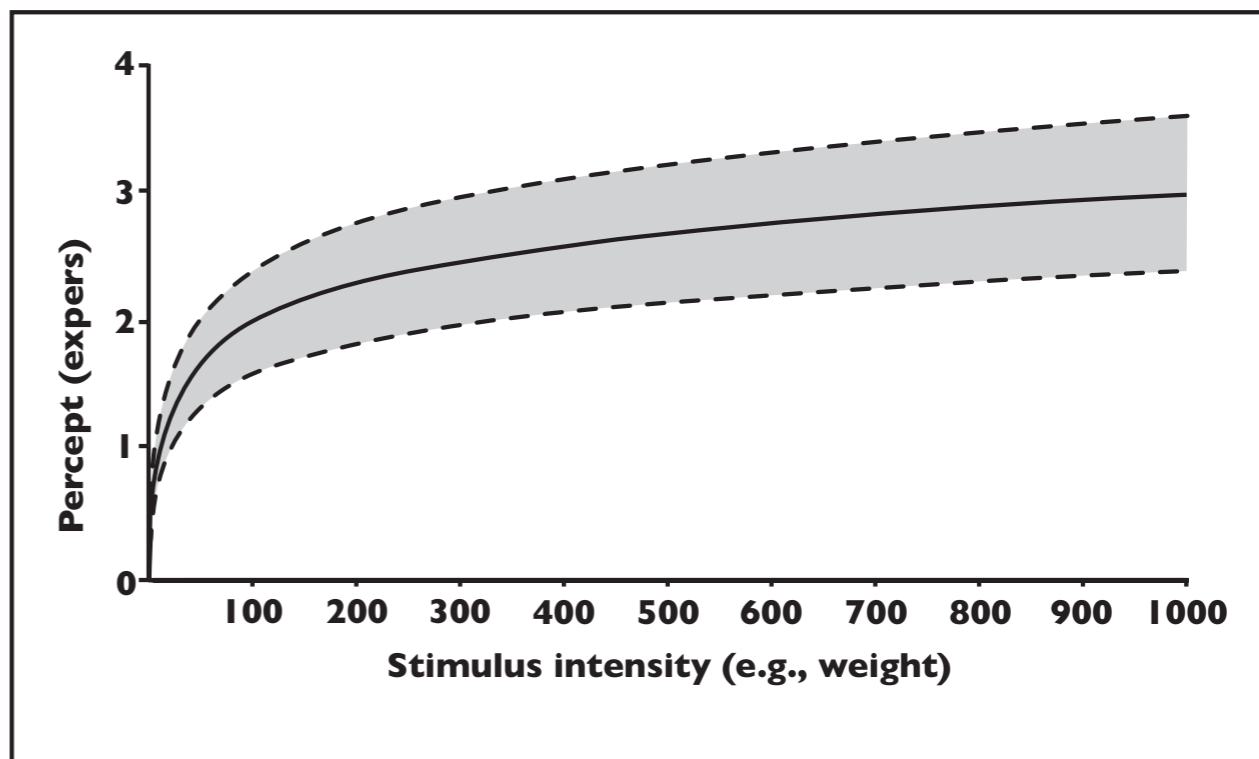
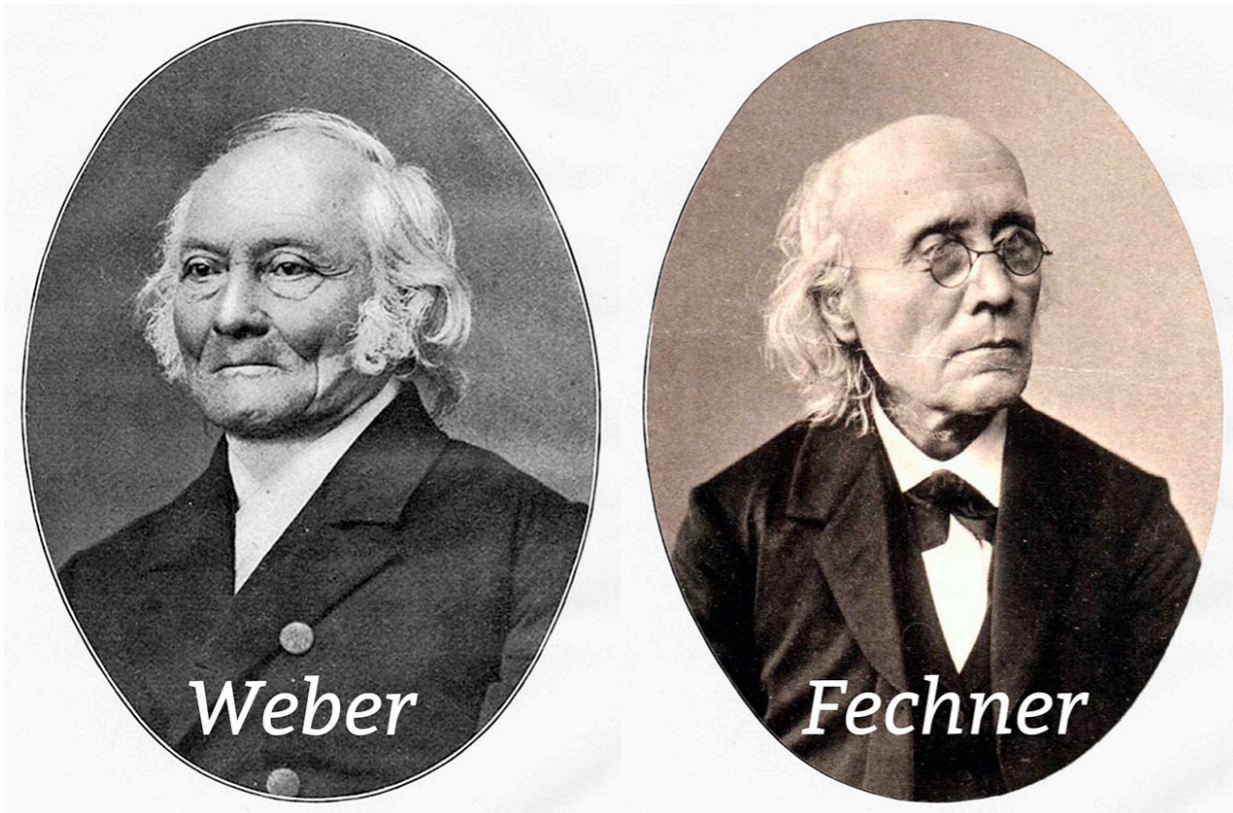
Same Movement, Different Values



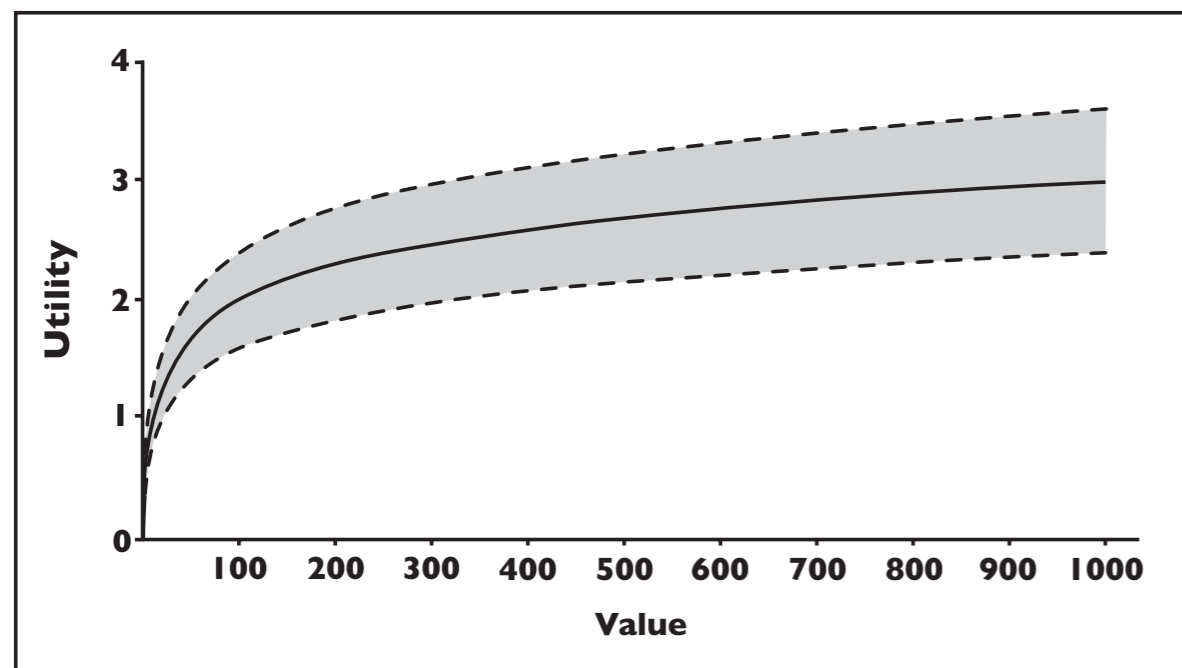
Large Reward Expectation
Small Reward Expectation







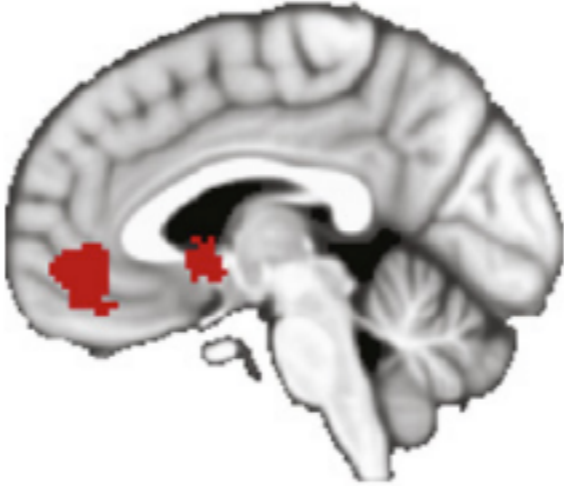
Daniel McFadden



Random Utility Theory



Knutson, Delgado and Elliott



$x = -3$



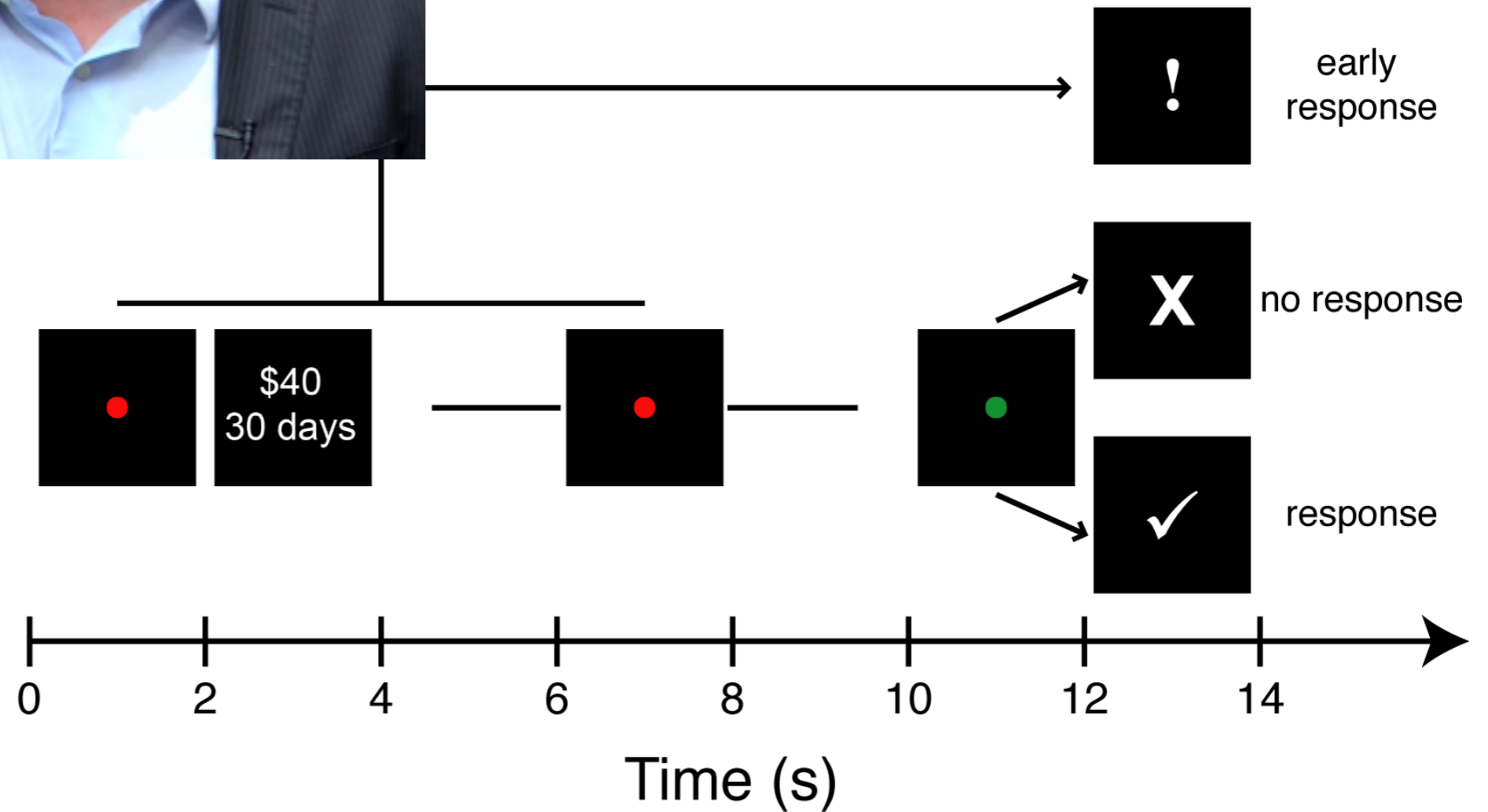
$y = 10$



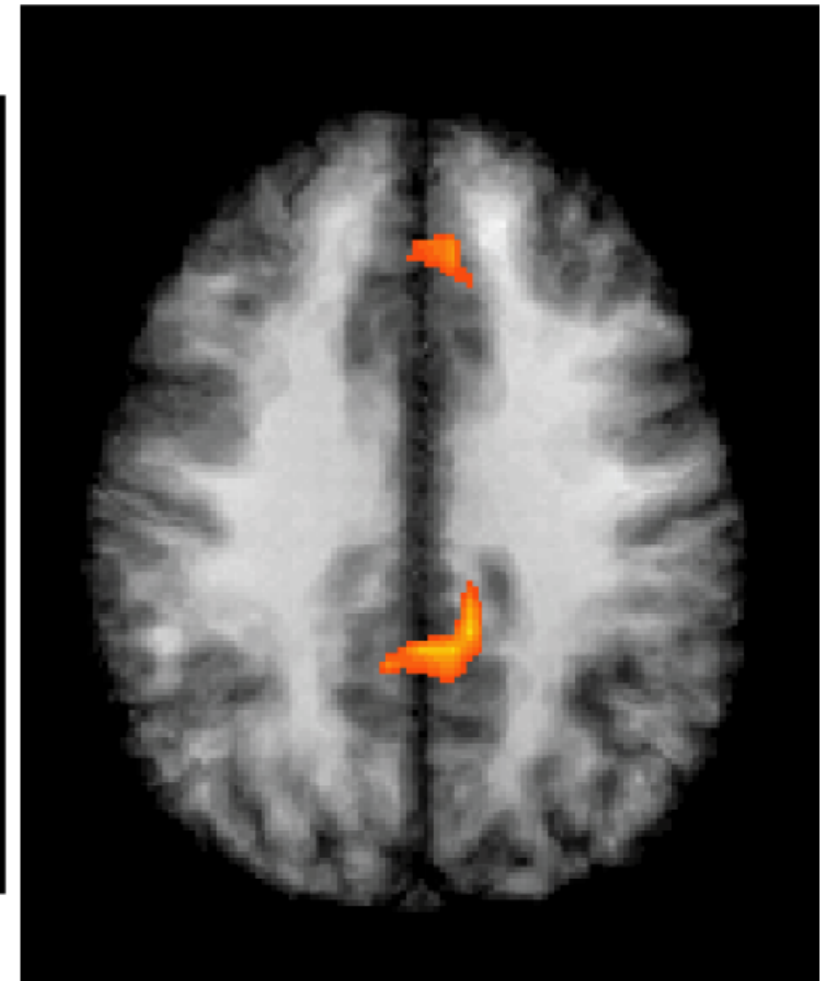
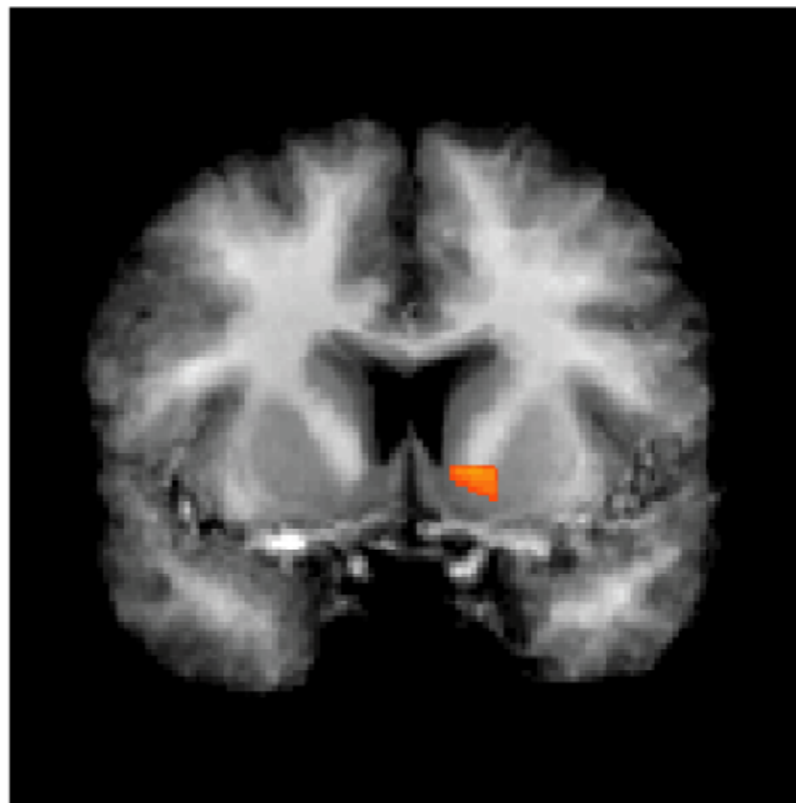
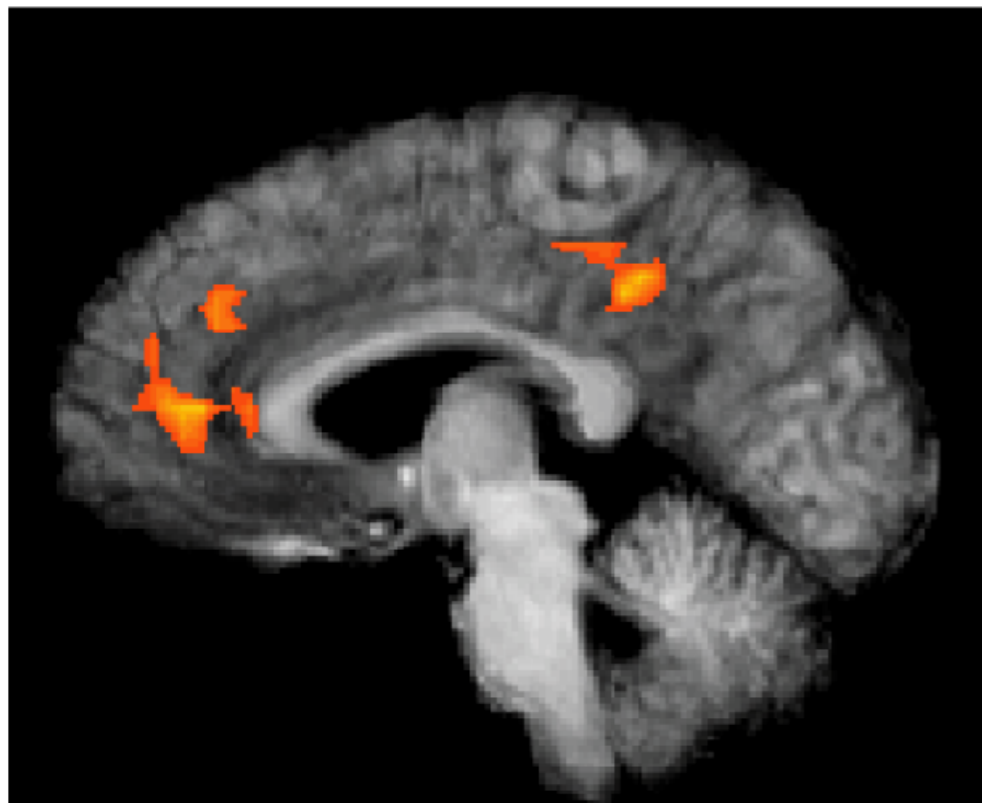
$z = -4$



Joe Kable

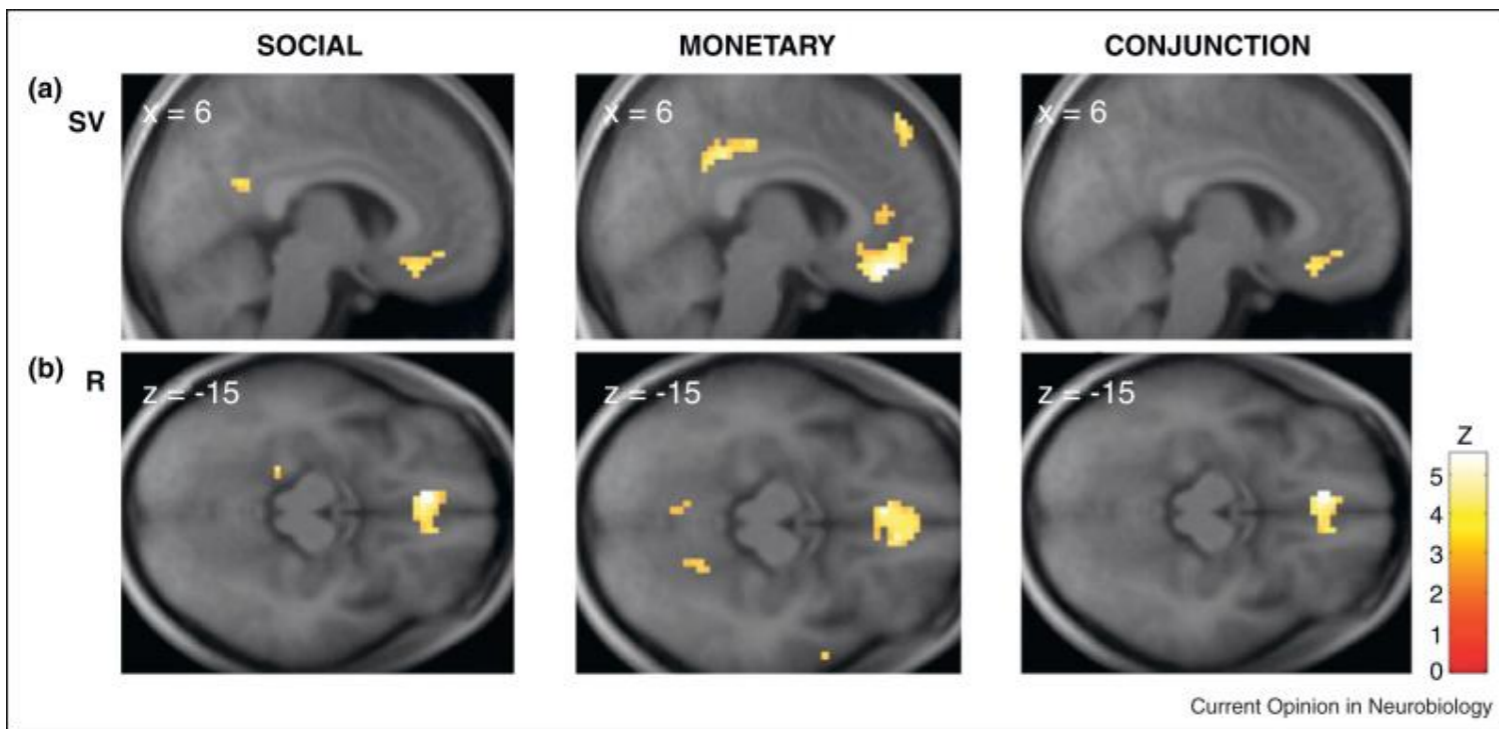


Correlation with Subjective Value



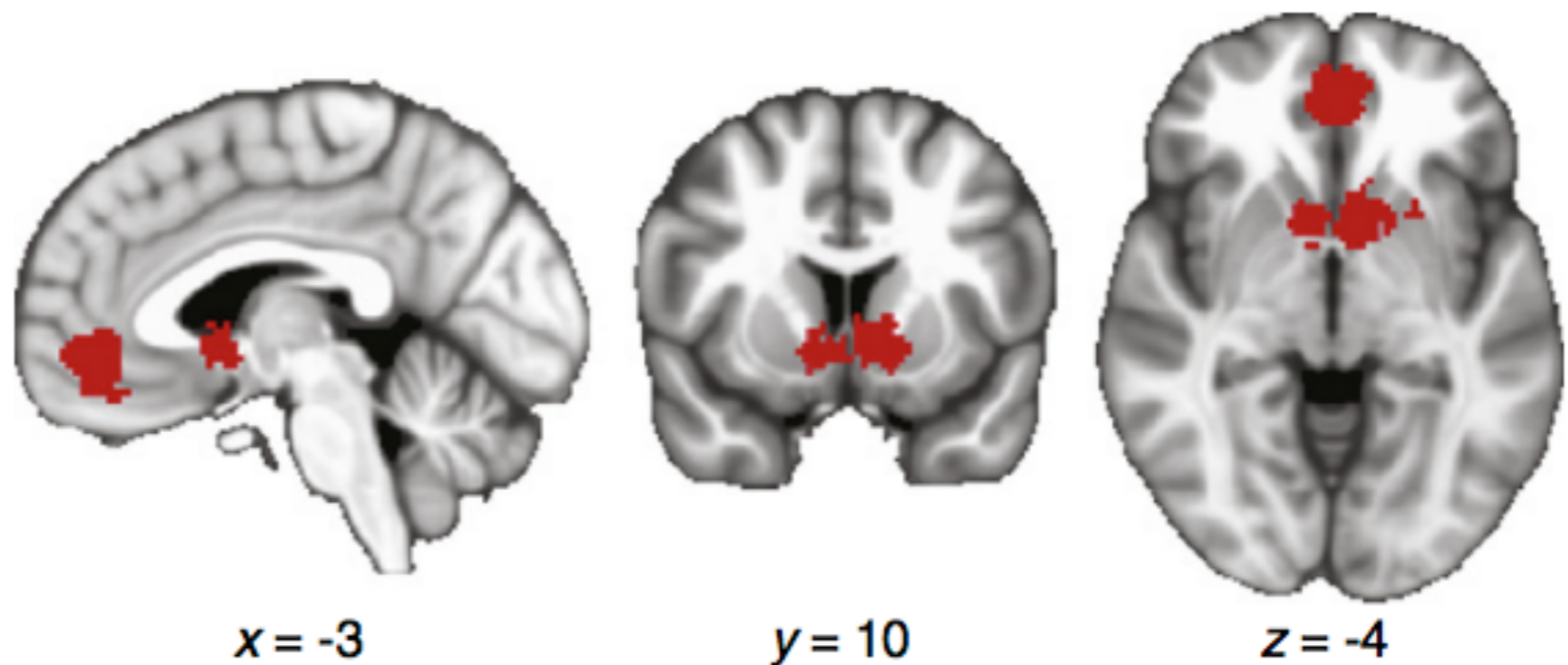
Random effect (n=10), TRs 4-6



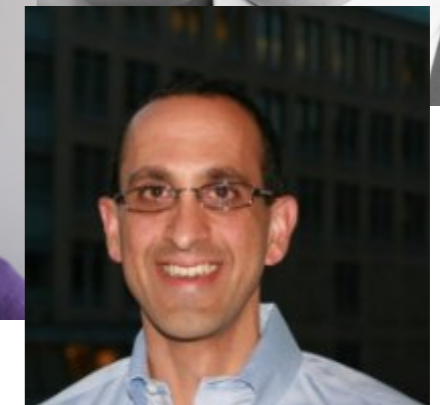
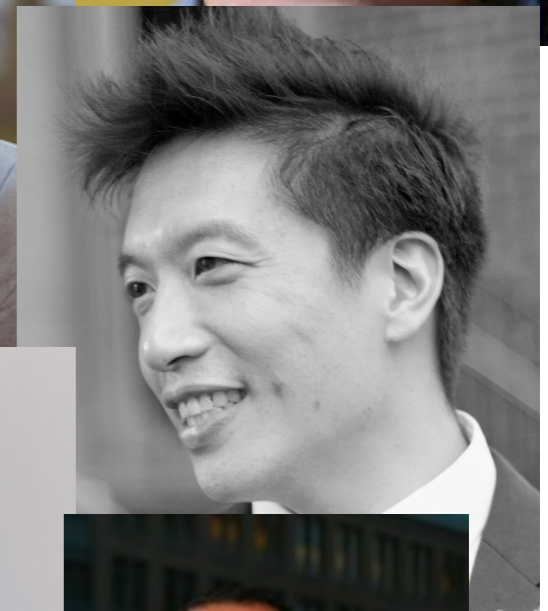
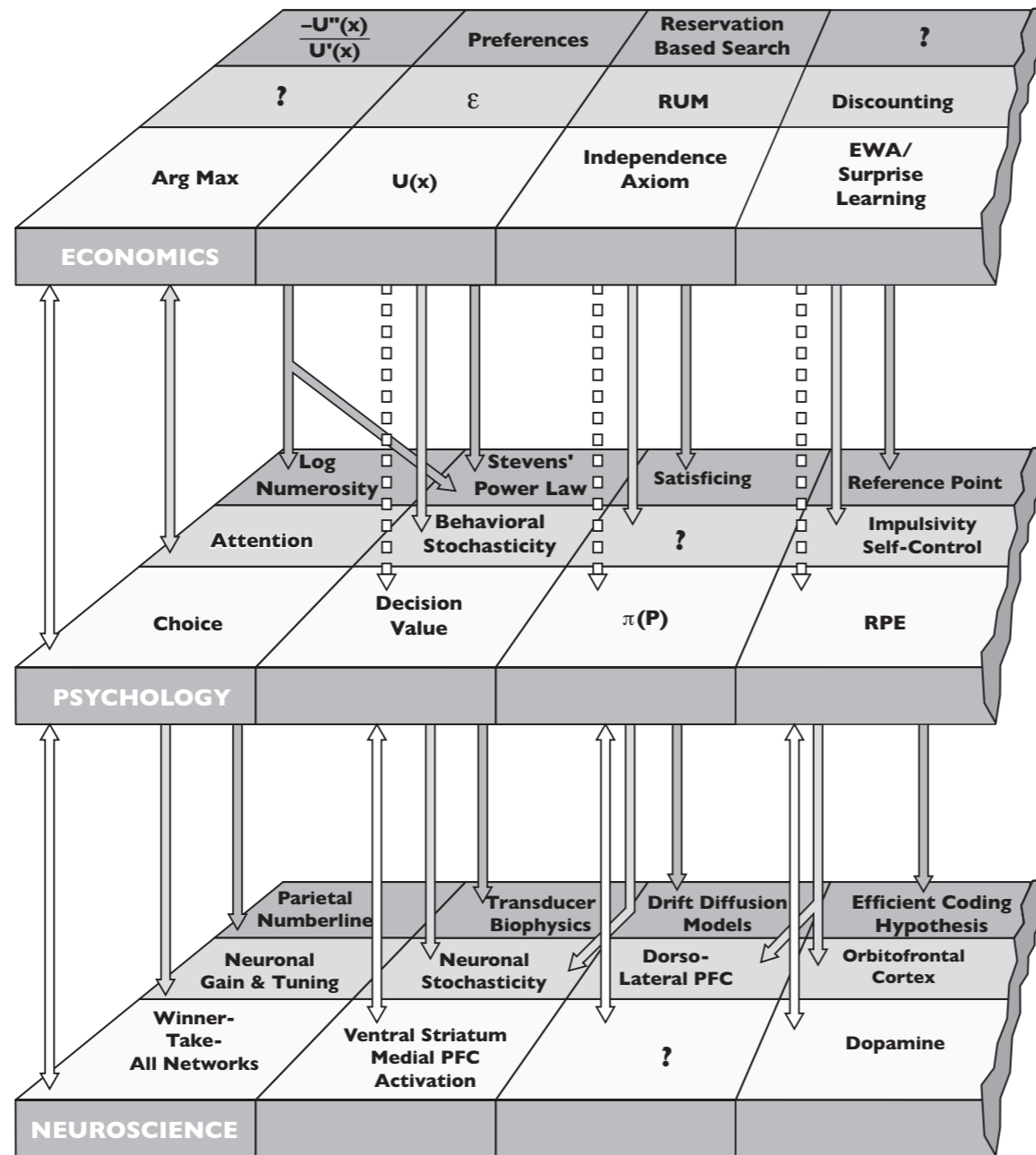


Levy and Glimcher

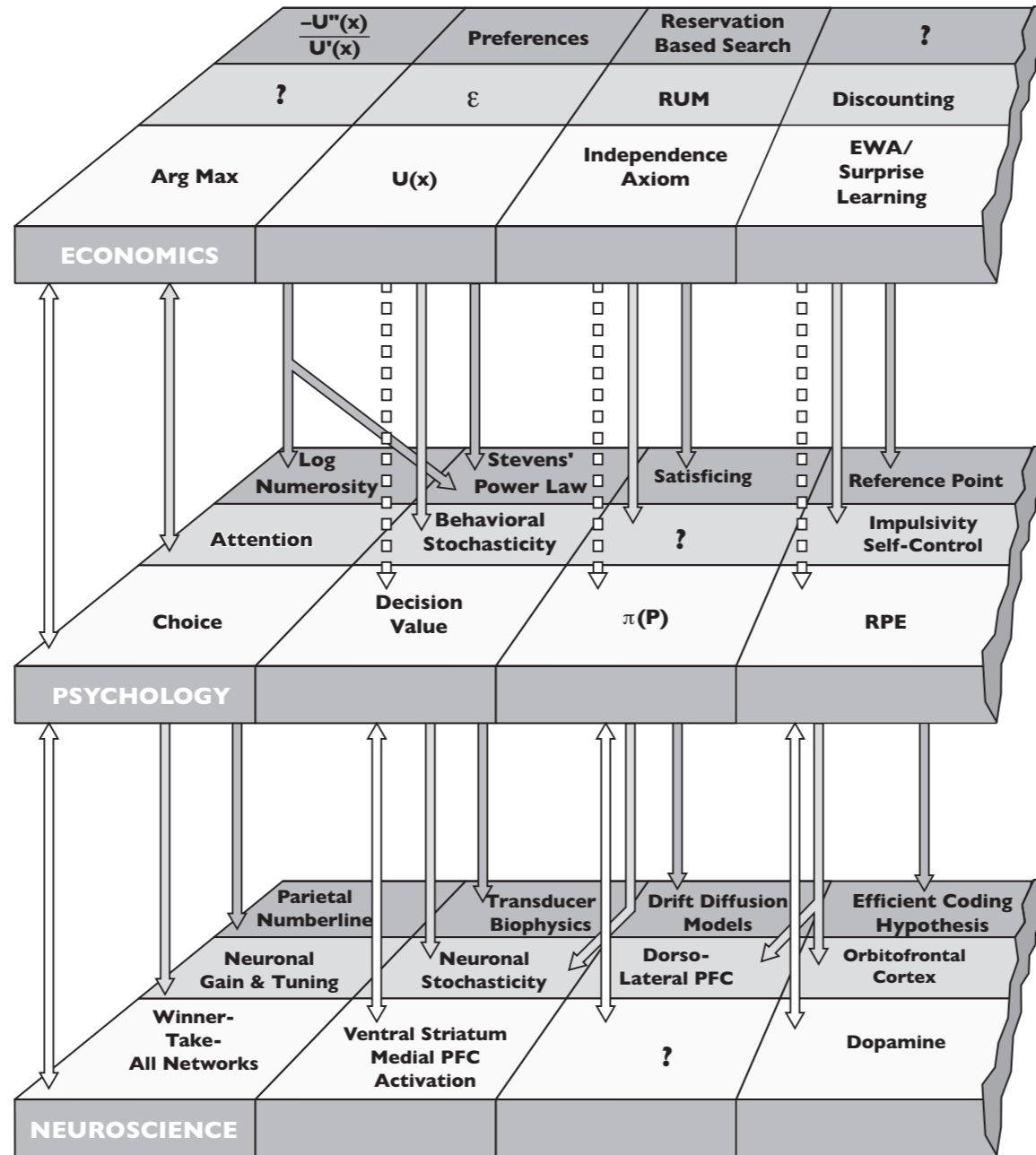
Bartra, McGuire and Kable



Up to this point:



When?



An Image



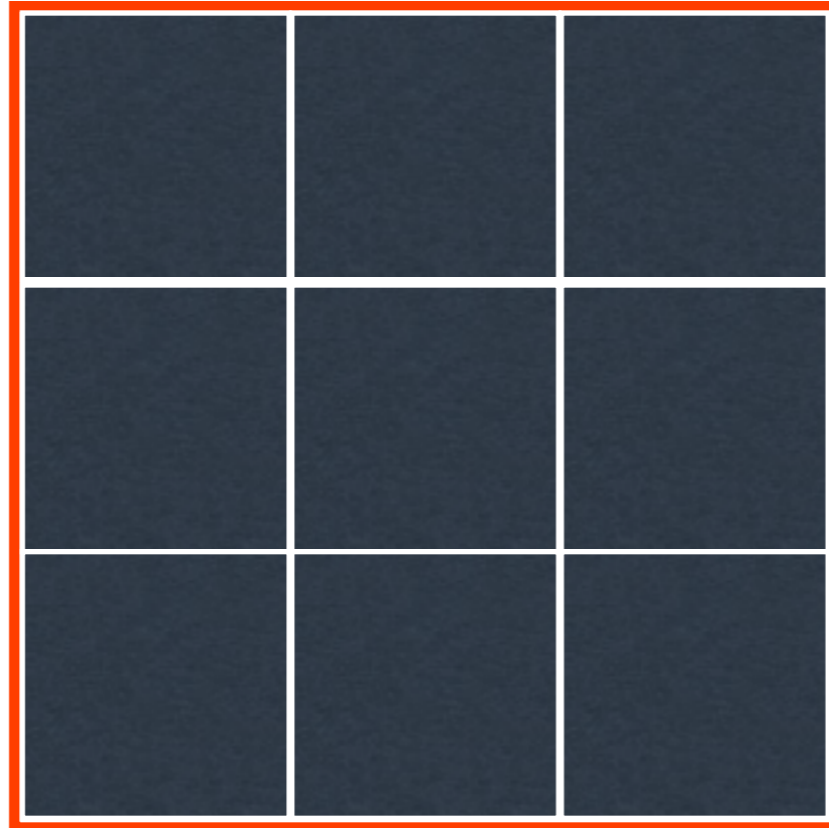
Horace Barlow

An Image



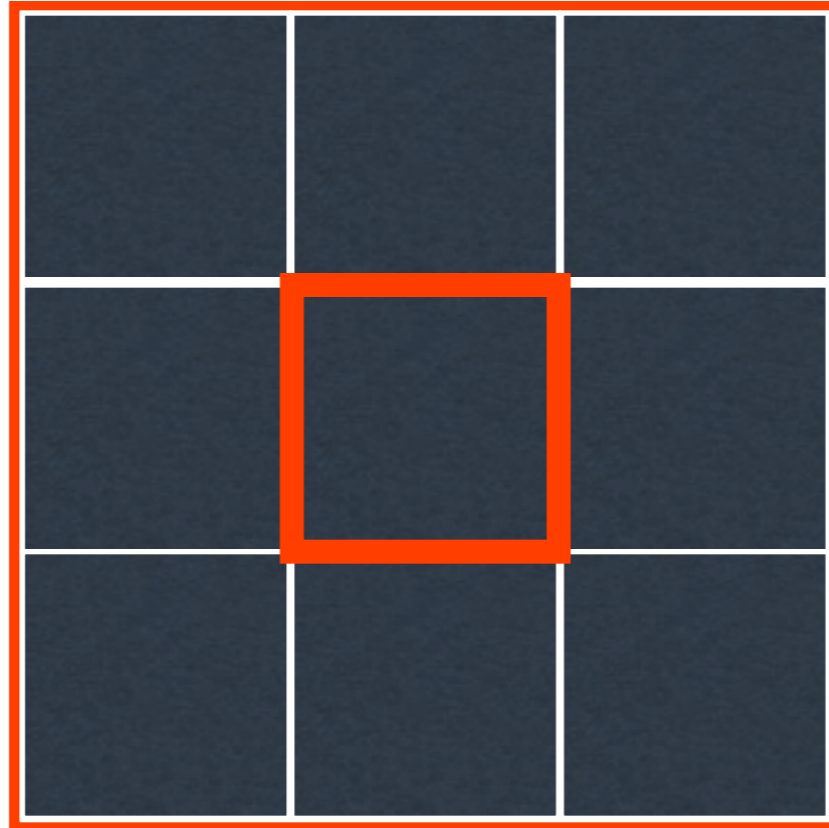
Horace Barlow

9 Pixels



Probability Pixel is Black: Conditional on
Adjacent Pixel Being Black: 0.75

9 Pixels

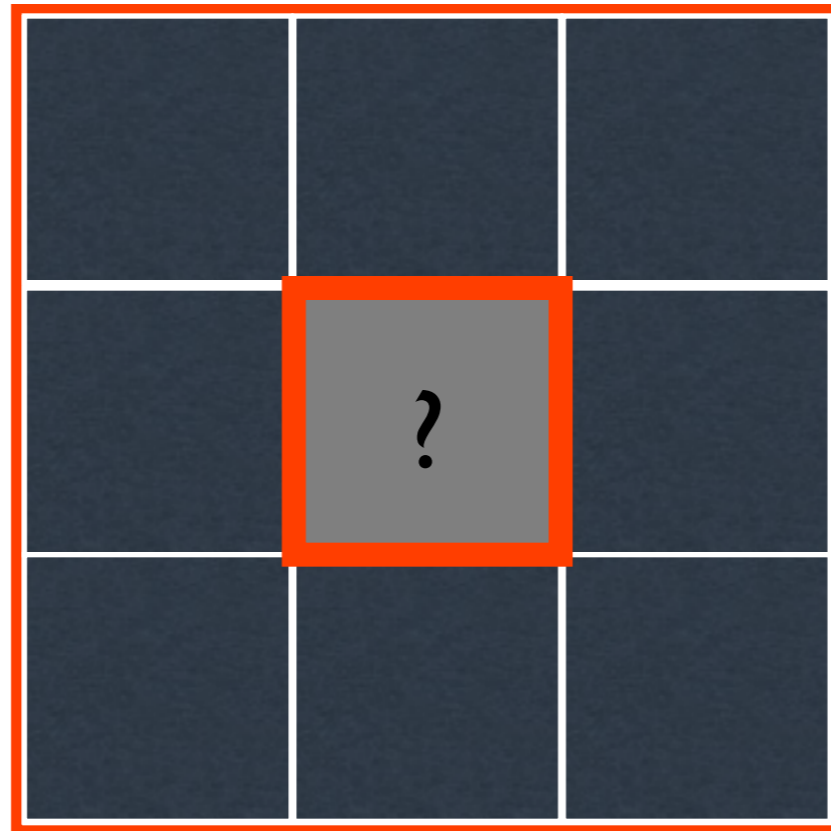


Black = 10 spikes

White = 0 Spikes

90 Spikes Total

9 Pixels

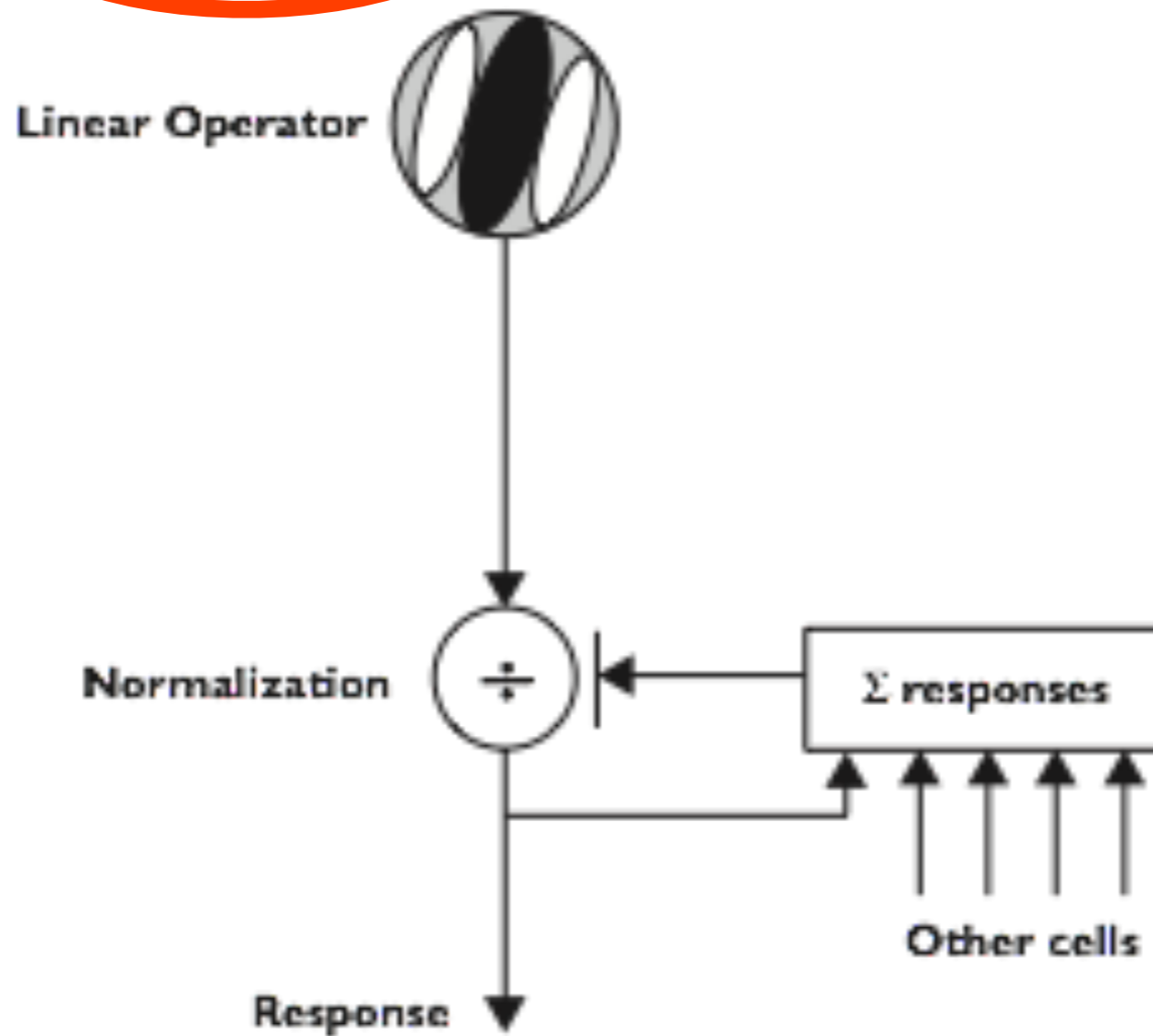


Conditional on these 8 pixels being black the ex ante probability of the center being black is ~ 1 .

So Why Waste Spikes

Heeger Normalization

$$R_i(t) = \frac{A_i(t) + \beta}{\sigma^2 + \sum_j A_j(t)}$$



Adapted from Heeger, 1993



Eero Simoncelli and Co. Showed That:

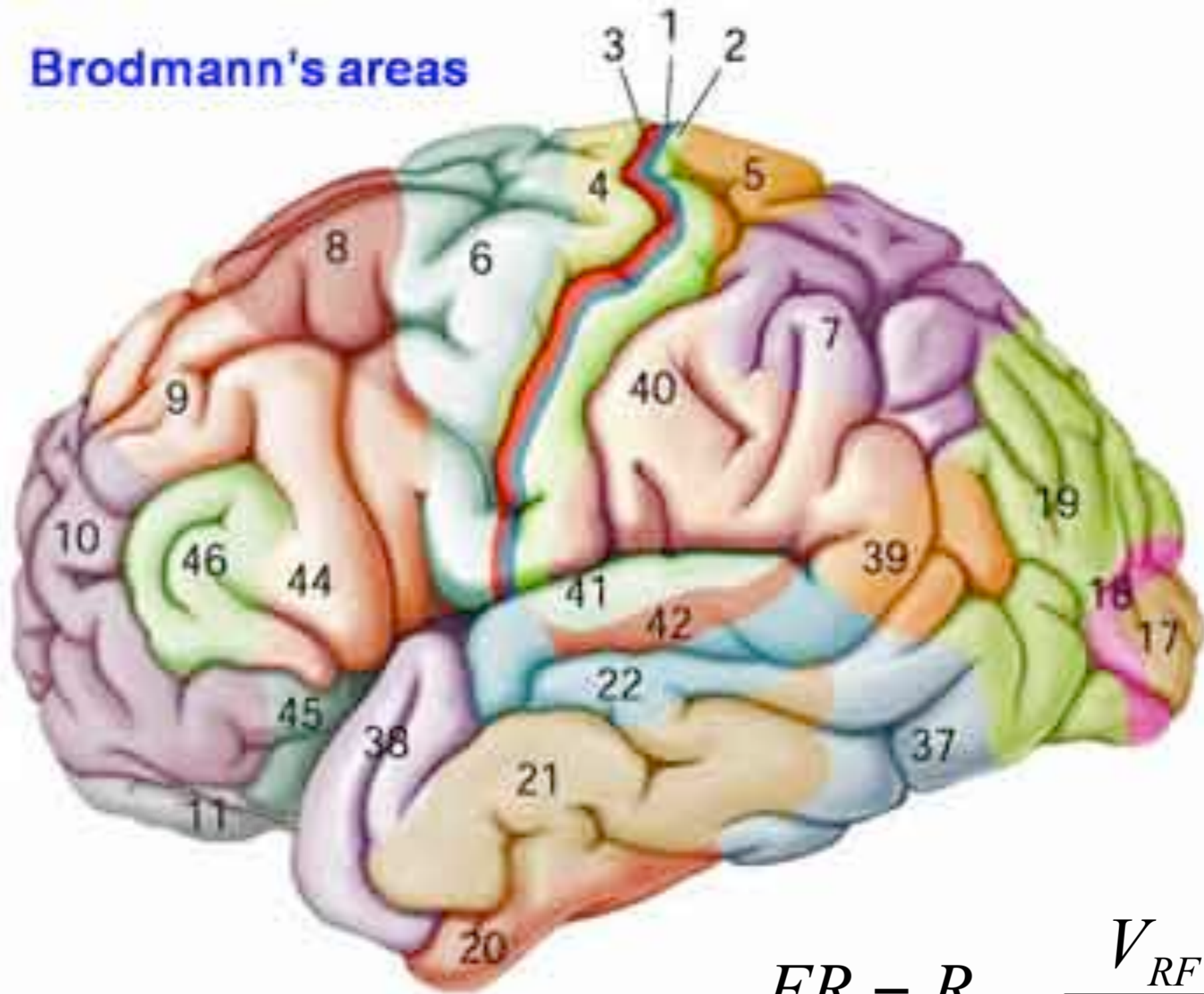
For any given pixel to pixel correlation structure
there is a set of w_i 's such that the minimum
number of spikes is used per bit of information

$$FR = R_{\max} \frac{V_{RF} + \beta}{\sigma^2 + \sum_j V_j w_j}$$

This is a form of decorrelation



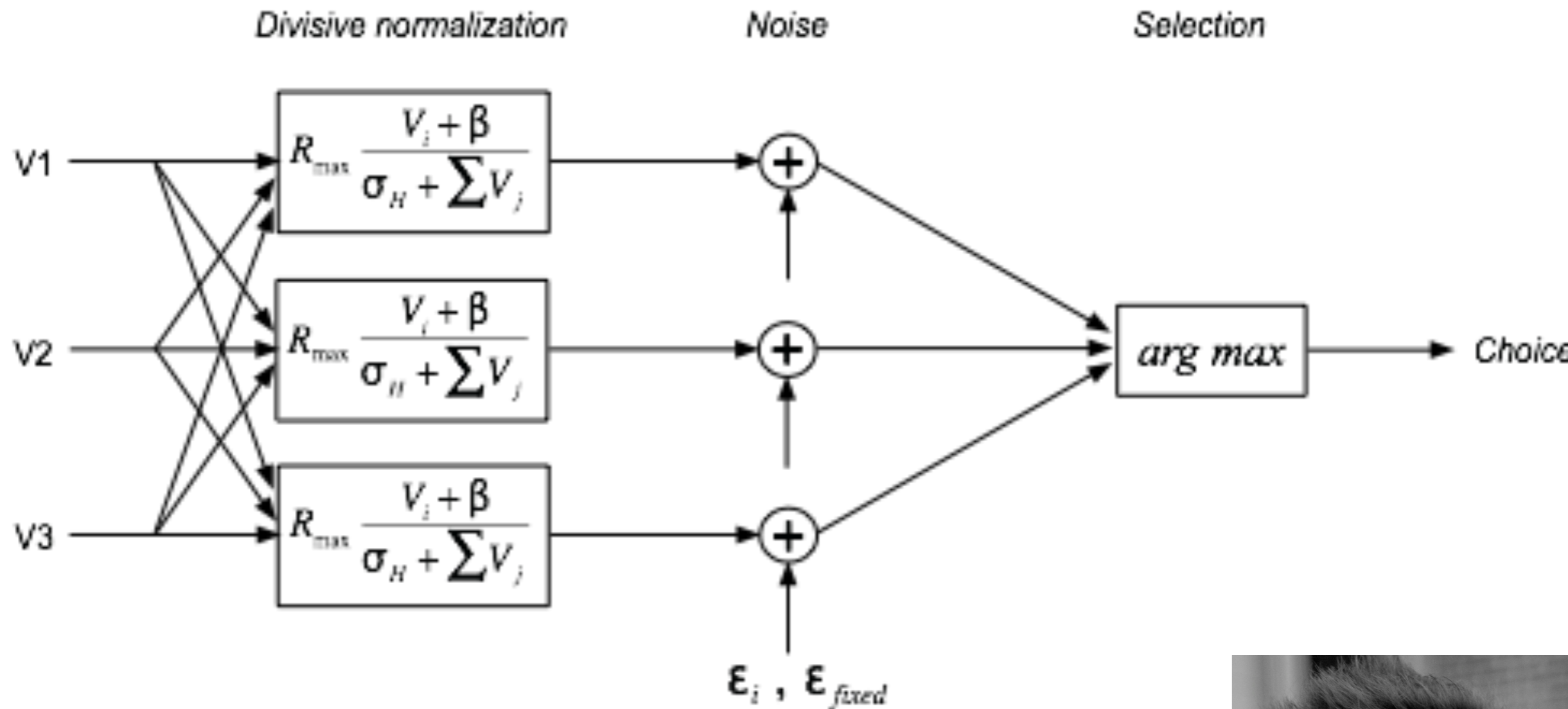
Brodmann's areas



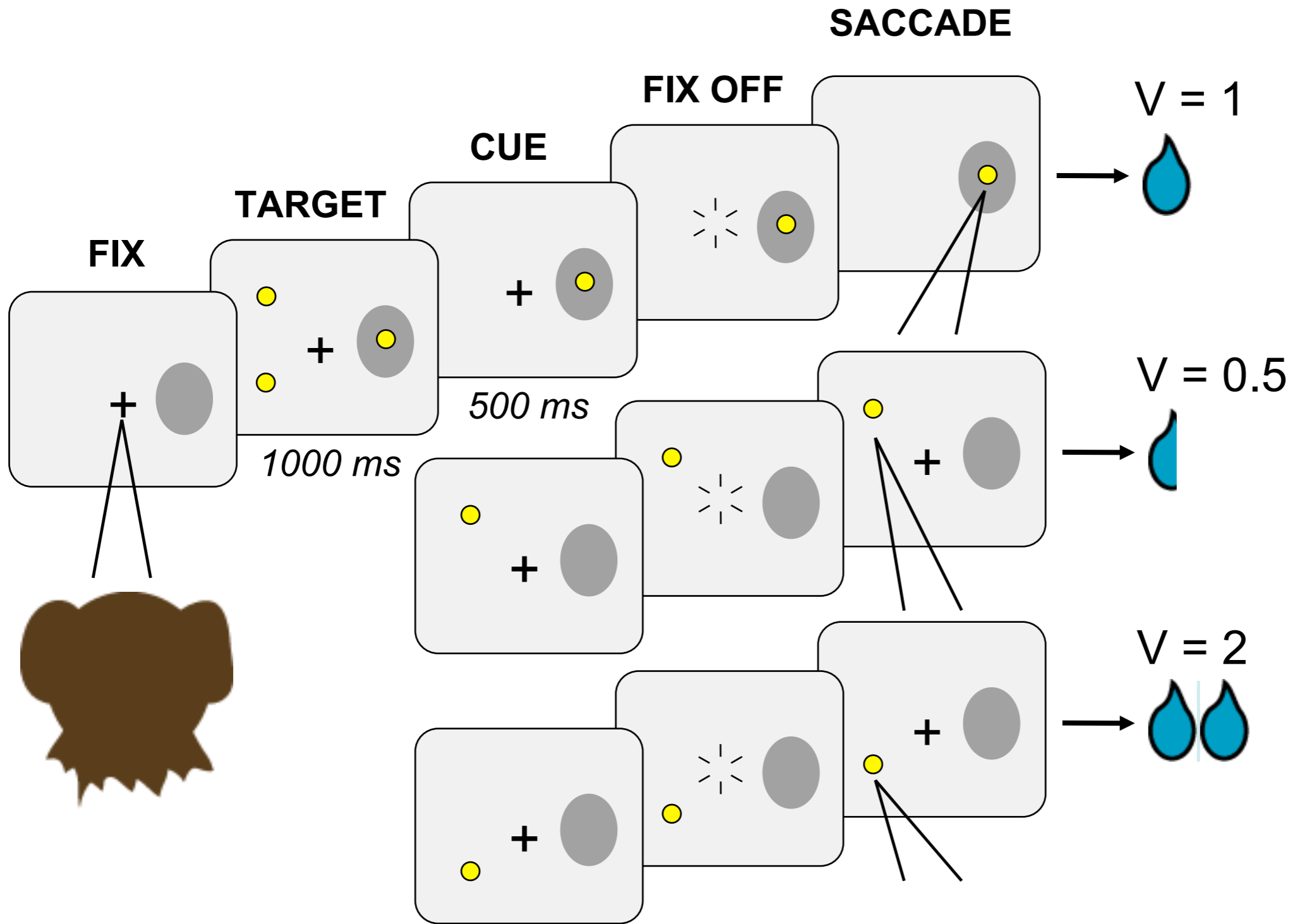
$$FR = R_{\max} \frac{V_{RF} + \beta}{\sigma^2 + \sum_j V_j}$$

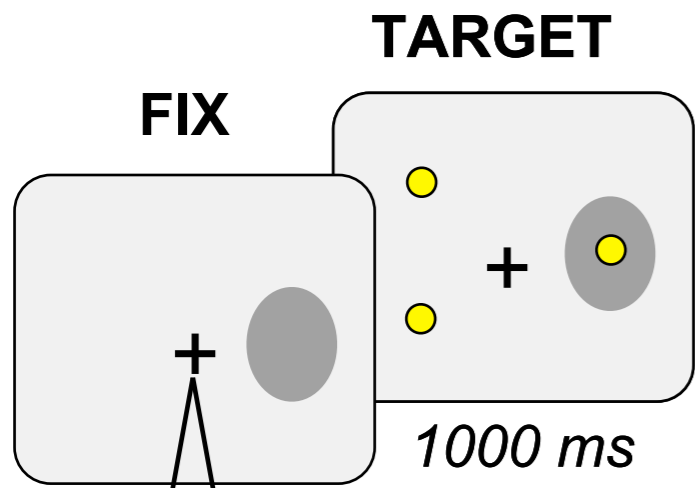
These Kinds of Networks Yield Strong 'Outside the Classical RF Effects'

A



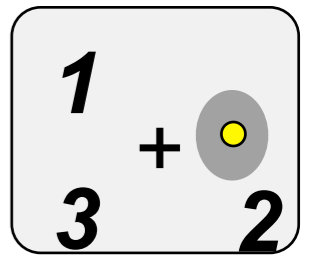
**Is There Evidence of Normalization in
Choice-Related Circuits?**



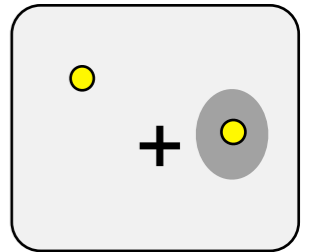


Randomized target array presentation

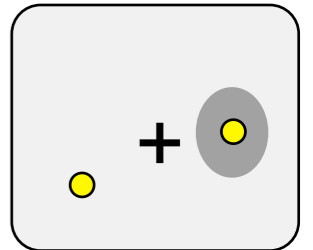
$$V_{RF} : \sum V_j \quad 1:1$$



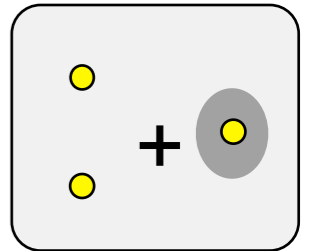
1:1.5



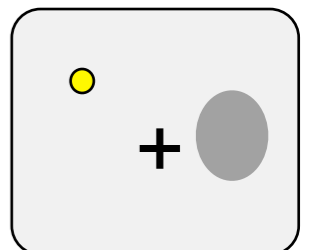
1:3



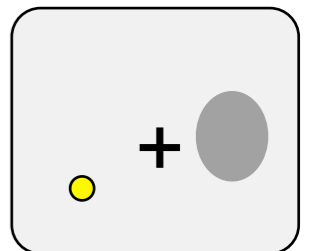
1:3.5



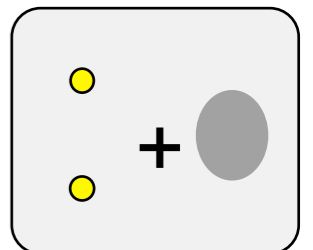
0:0.5

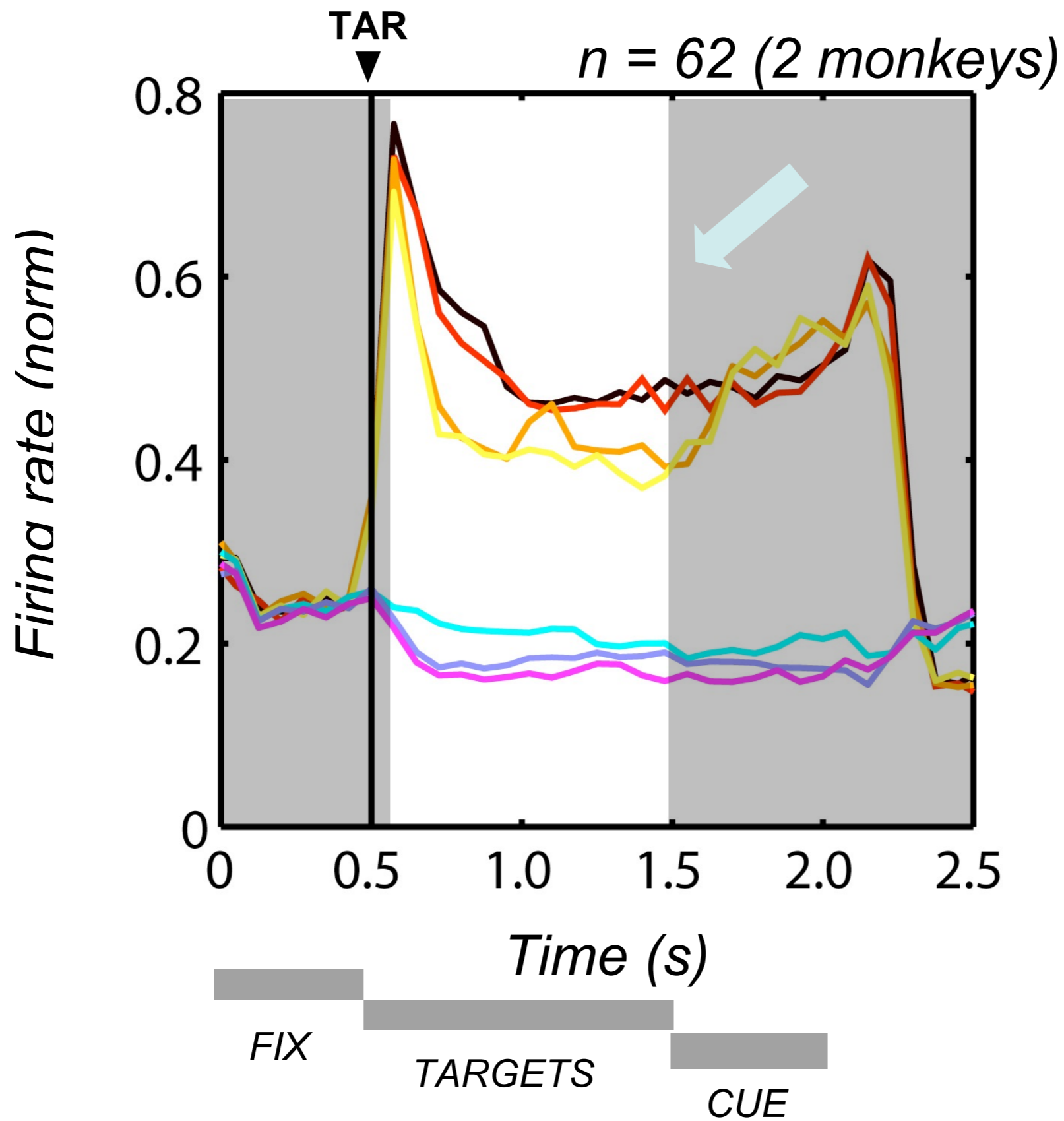


0:2

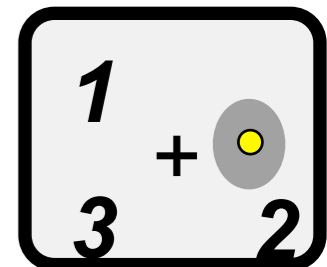


0:2.5

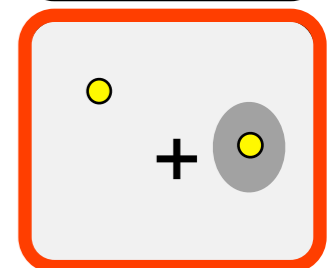




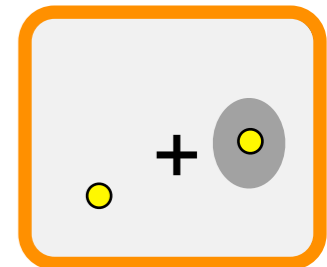
1:1



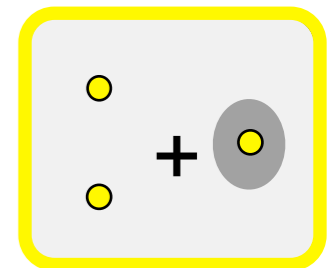
1:1.5



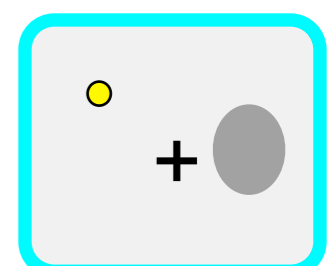
1:3



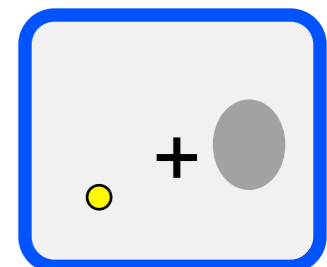
1:3.5



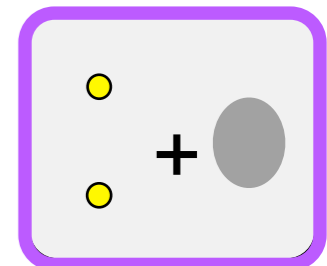
0:0.5



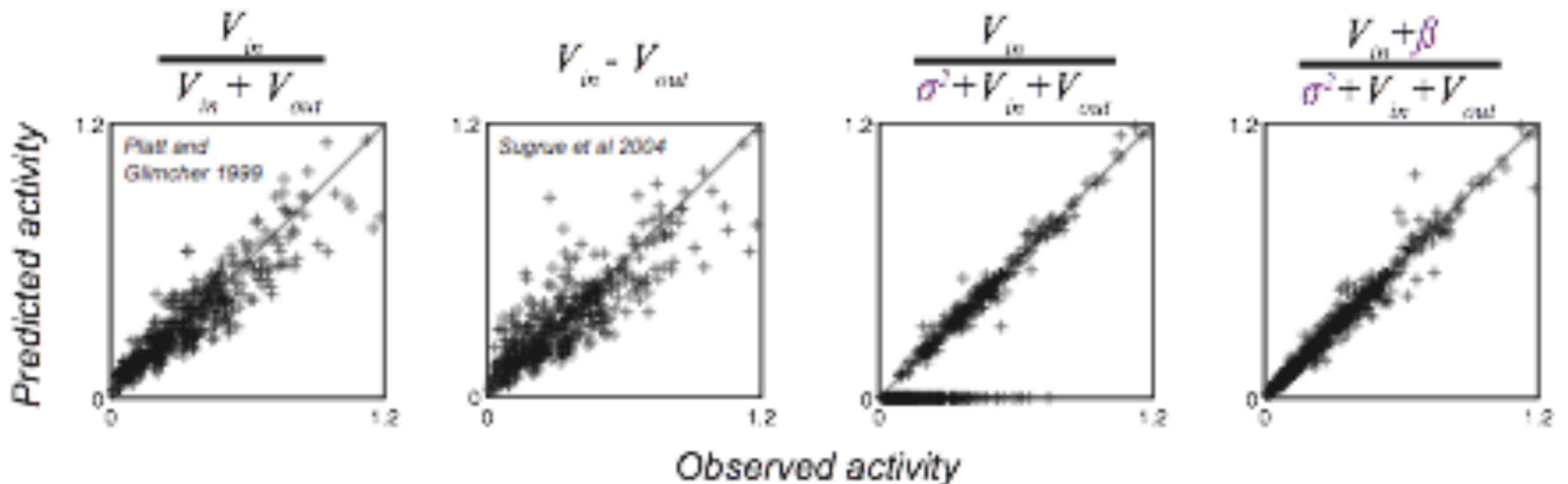
0:2



0:2.5



Comparing Models Across the Dataset



Platt &
Glimcher

Sugrue,
Corrado &
Newsome

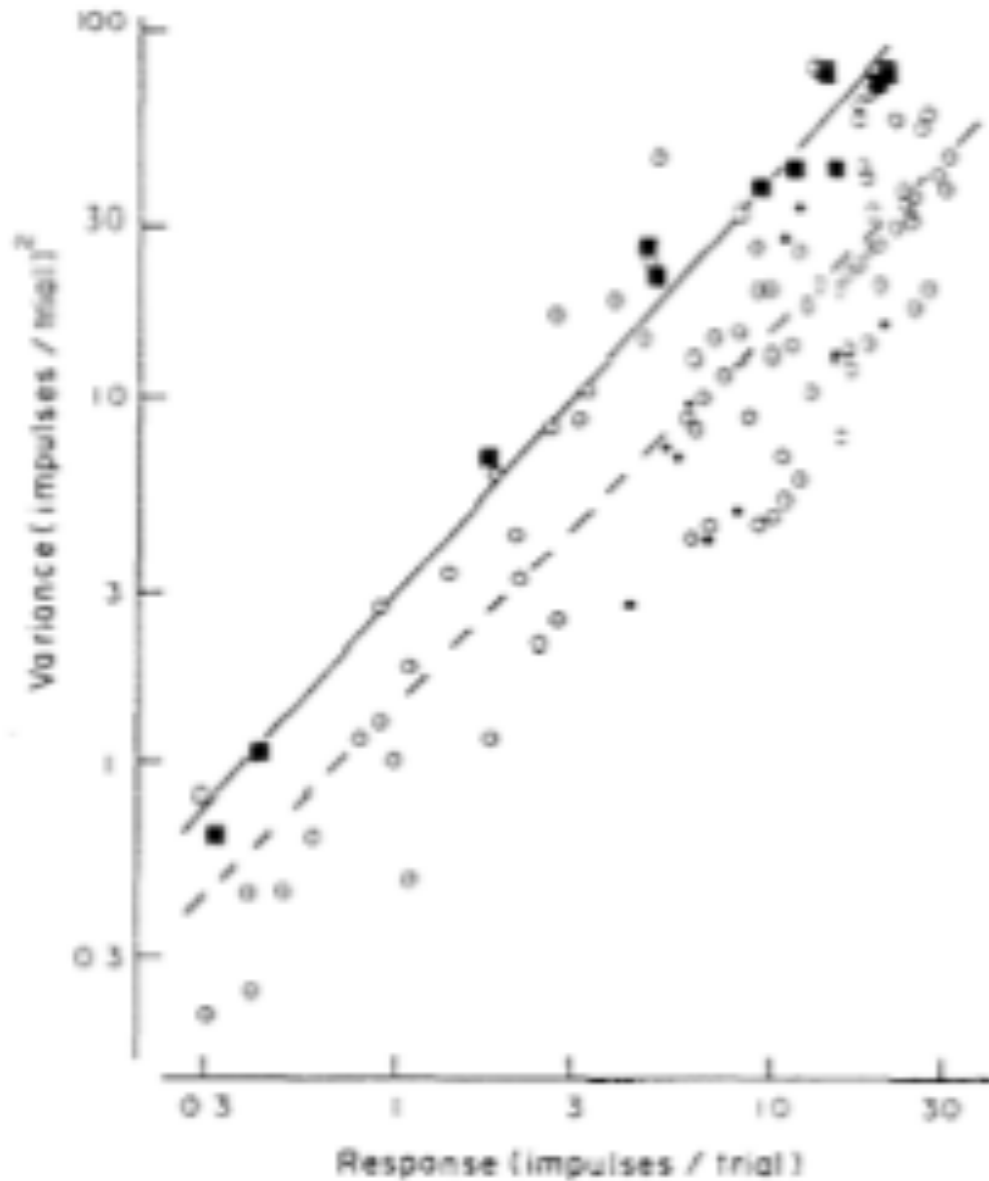
Heeger

Heeger &
Reynolds

If There is Normalization, Would it Influence Choice Behavior?

The Three Option Problem

Variance



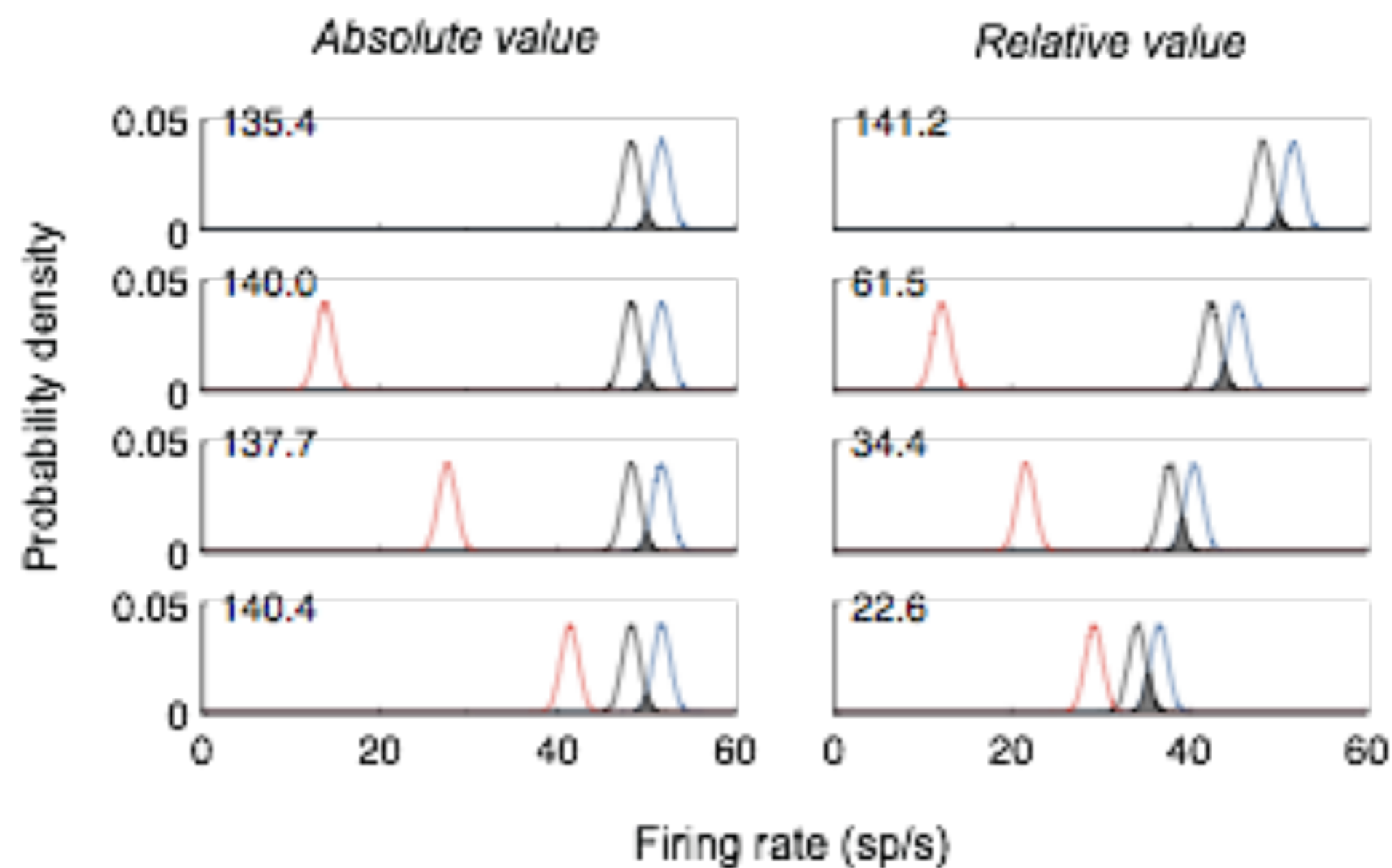
Neuronal Rate

Tolhurst, Movshon and Dean, 1983

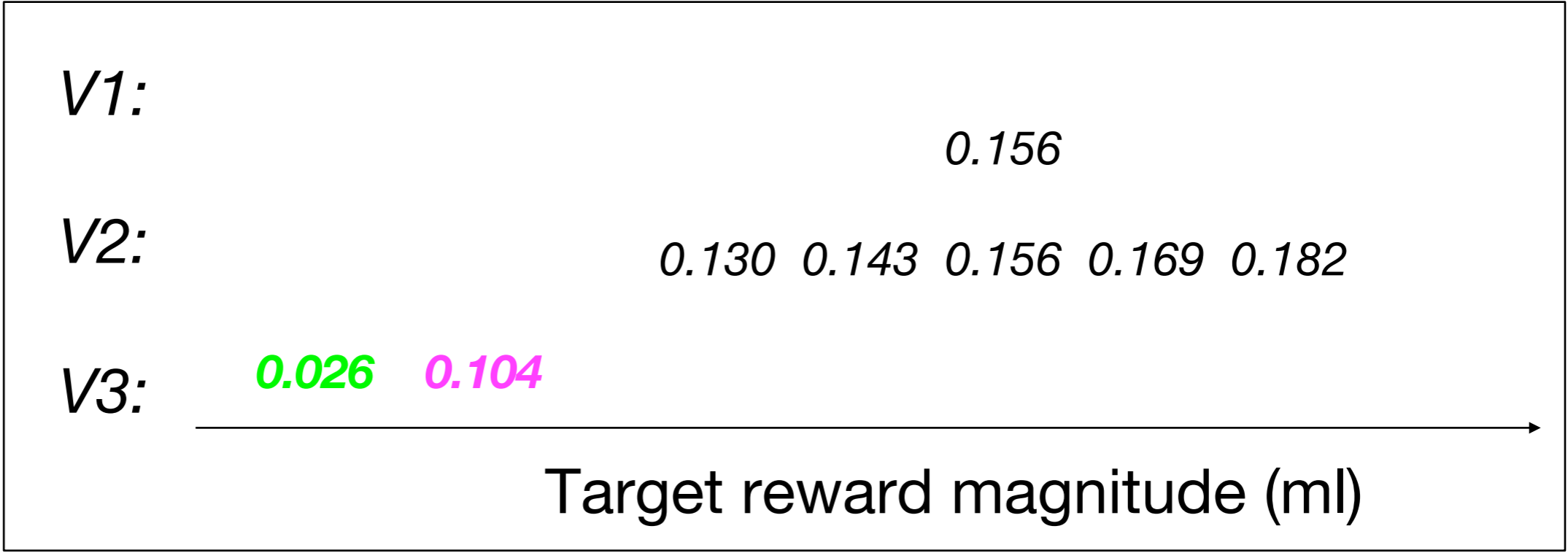
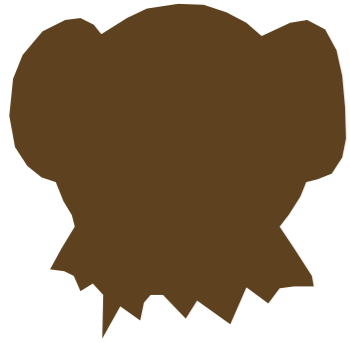
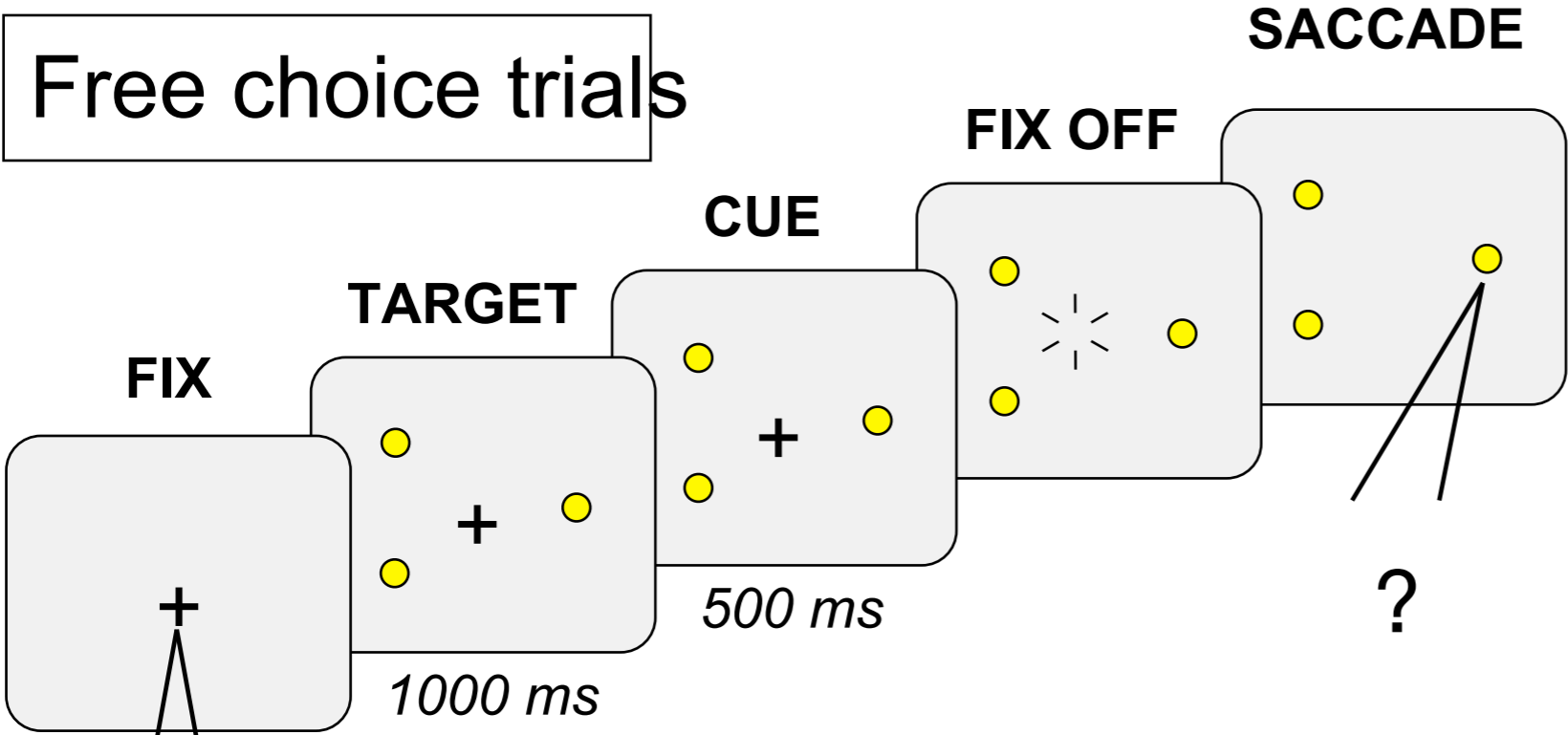
For Neurons:

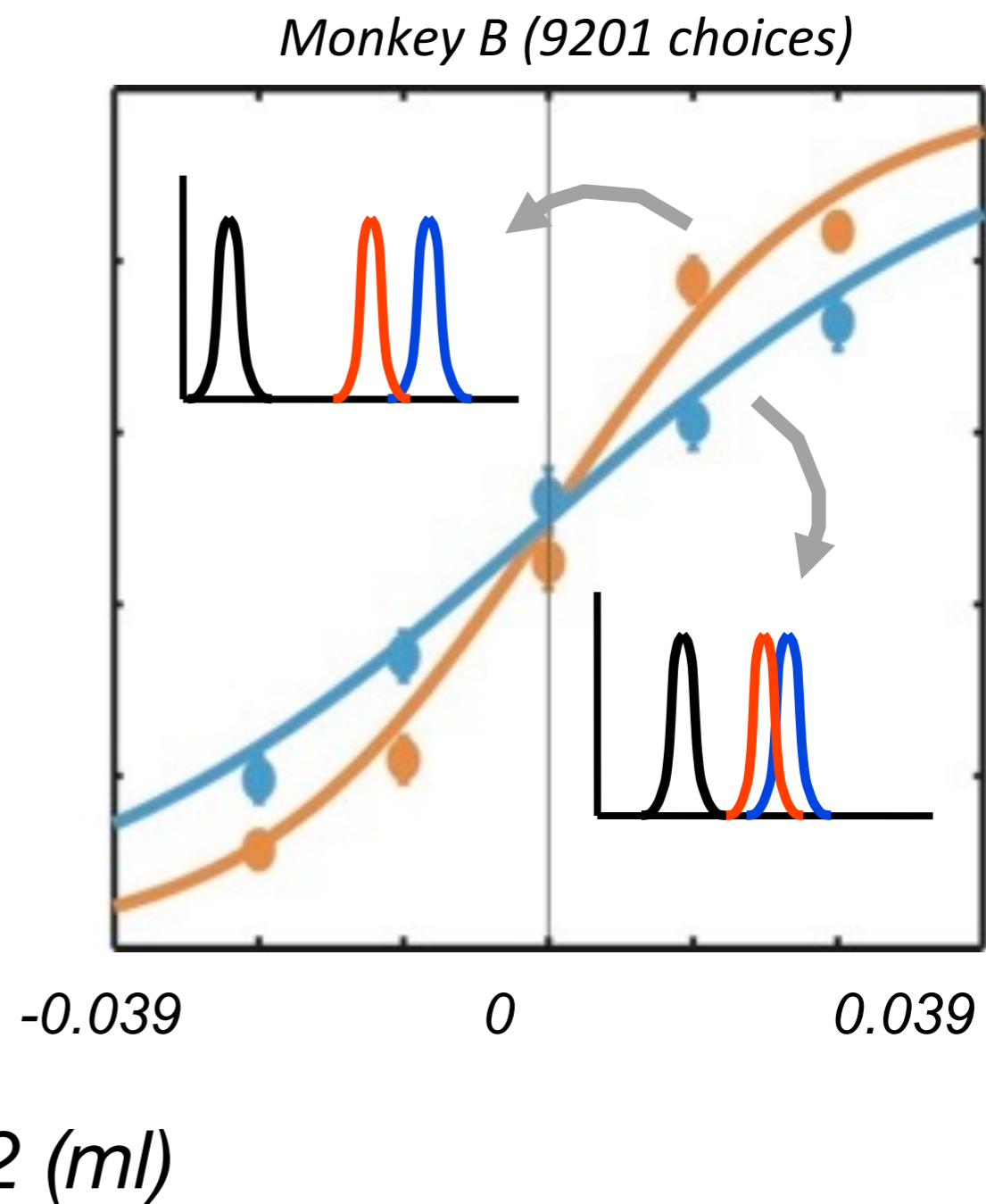
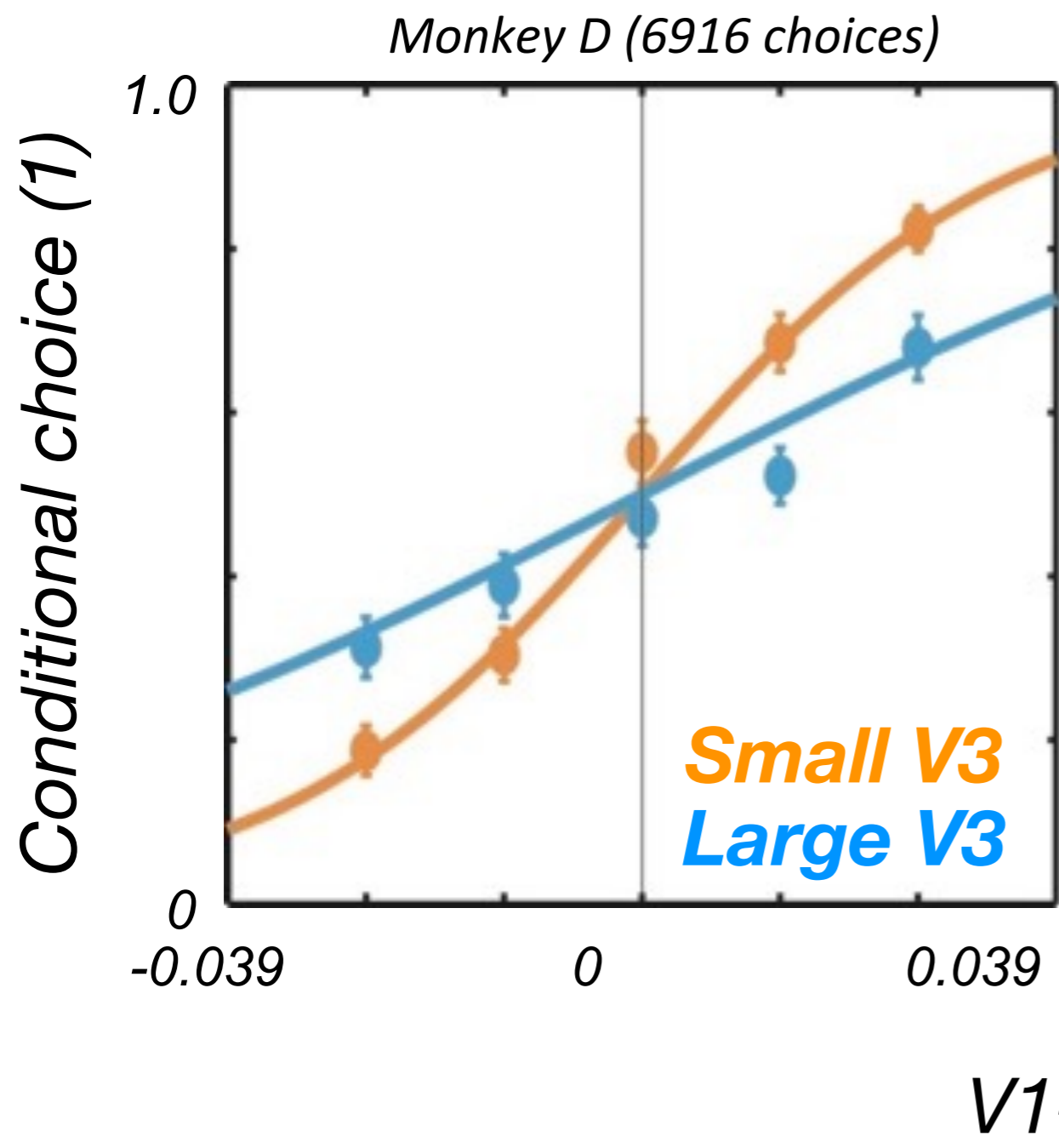
- Variance Scales with Mean
- Variance is the average **Squared** distance from the mean
- Thus as means reduce, discriminability goes down

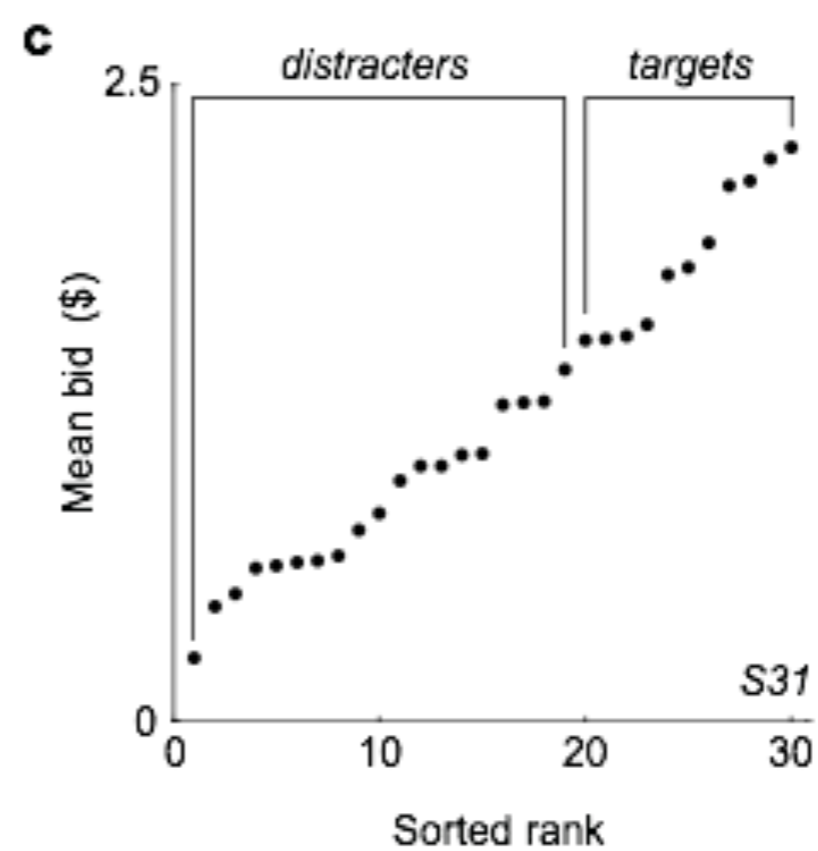
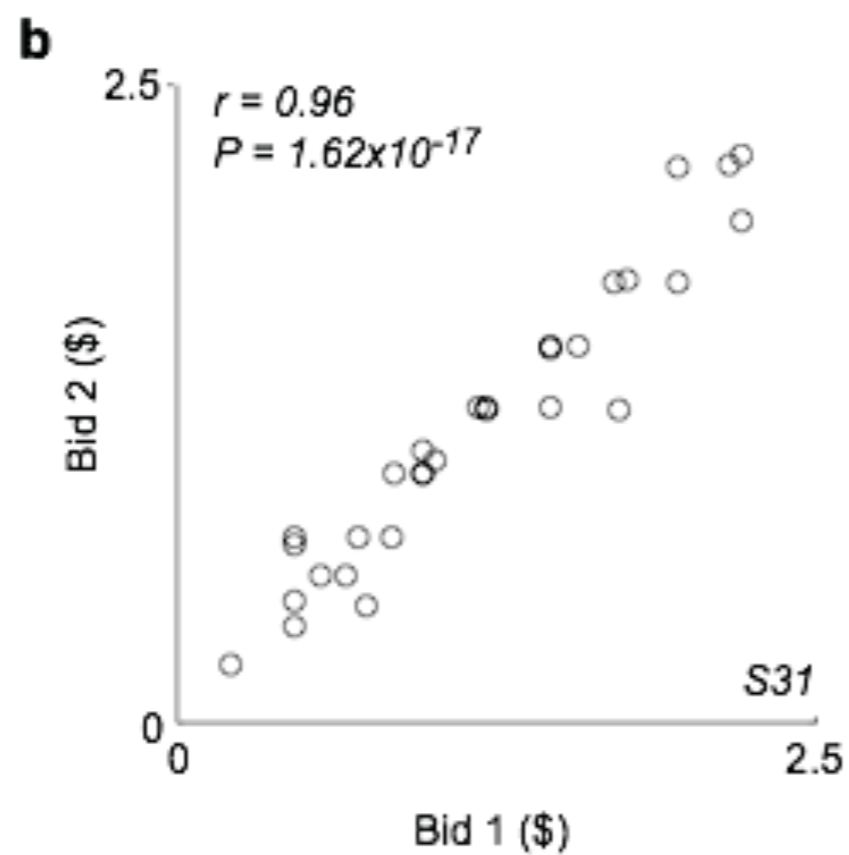
variance ~1.0-1.5 mean



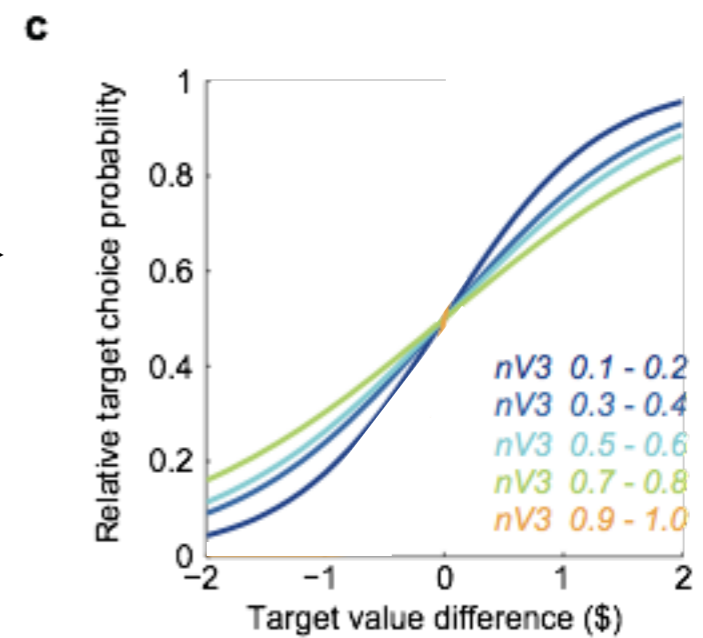
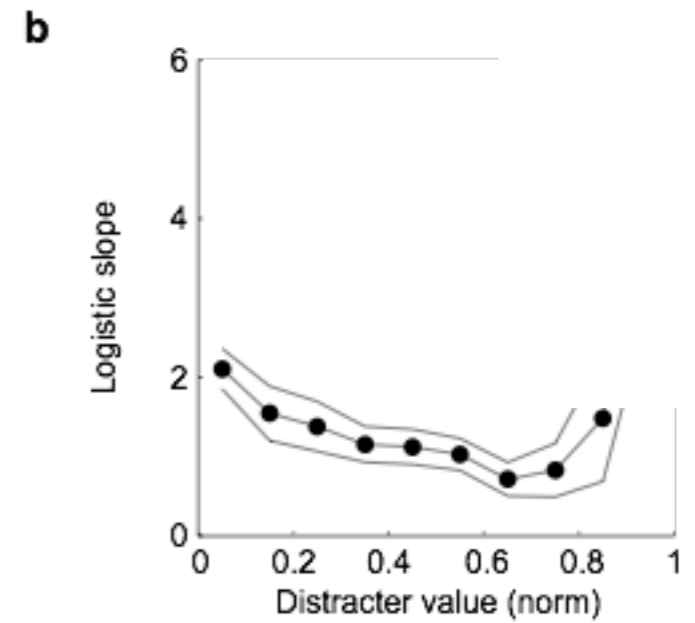
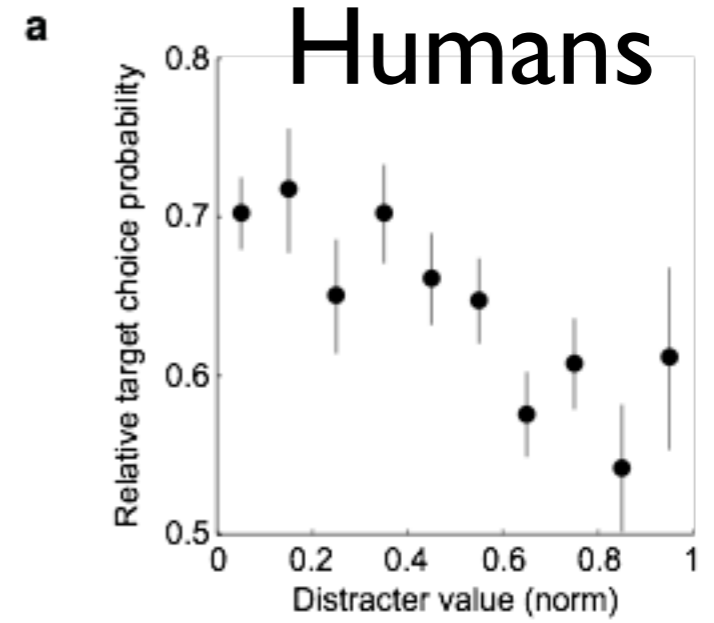
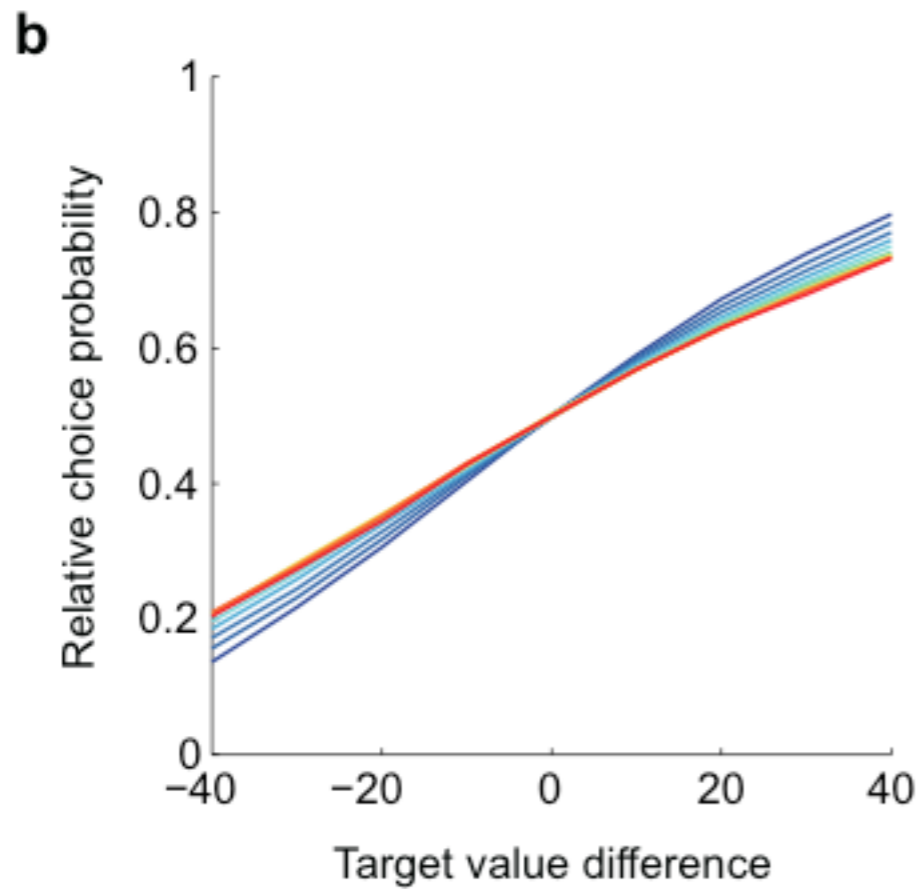
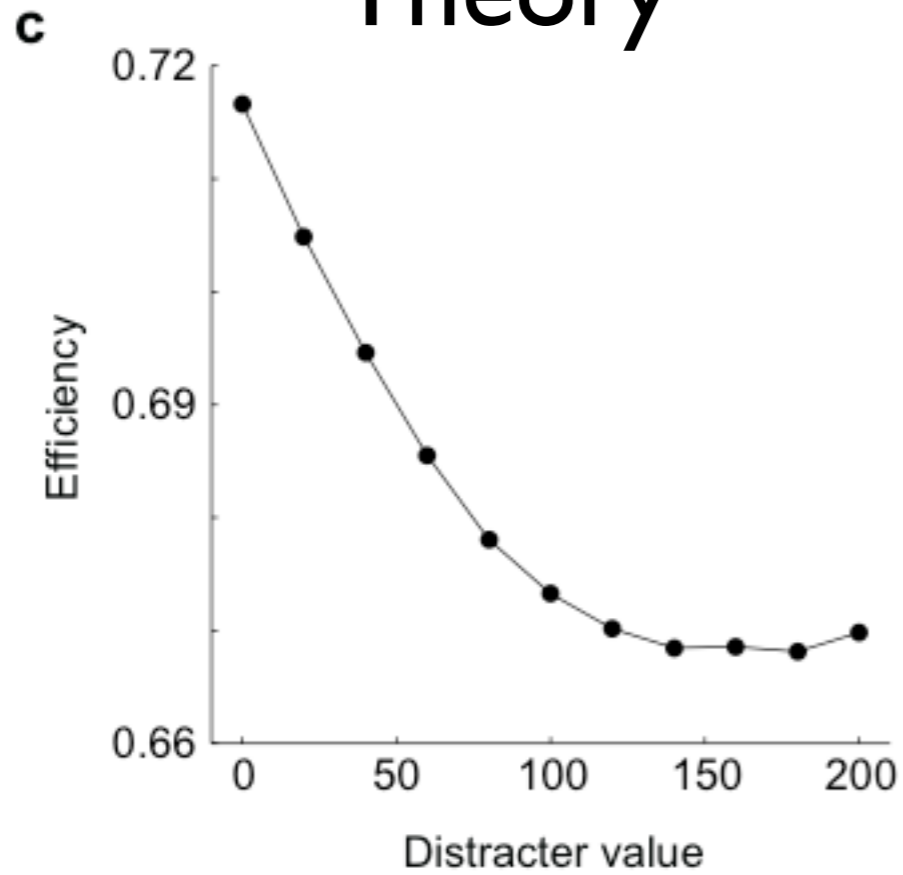
Free choice trials







Theory



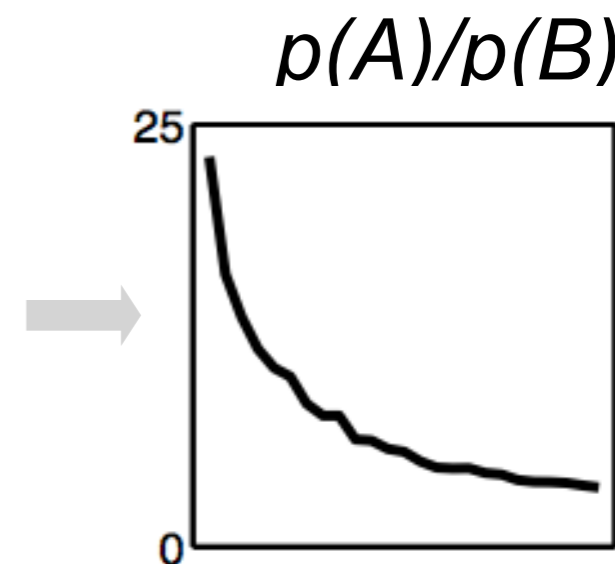
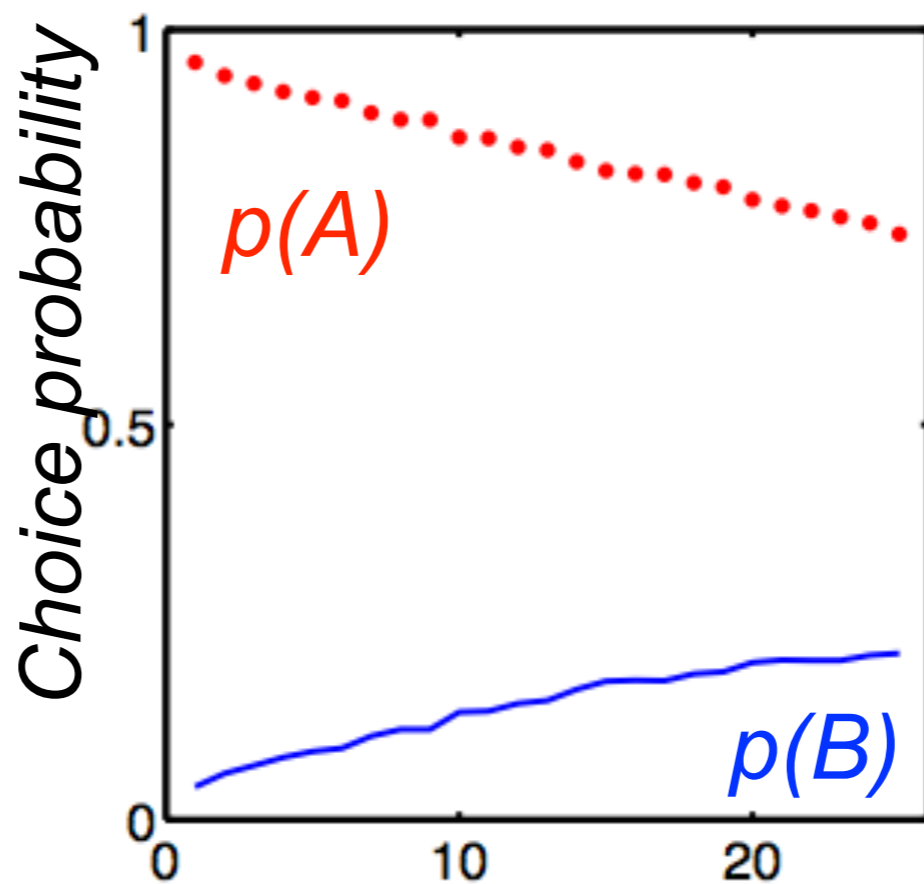
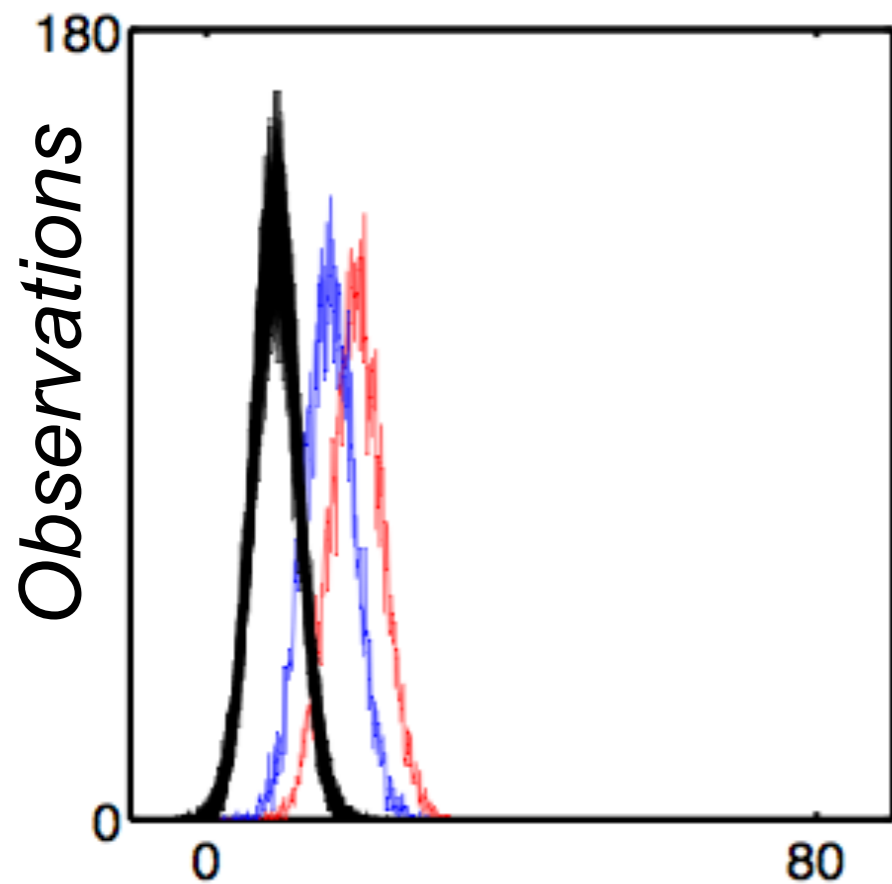
If There is Normalization, Would it Influence Choice Behavior?

The Three Option Problem

The Multi-Option Problem

The Curse of Choice

var/mean = 1.0



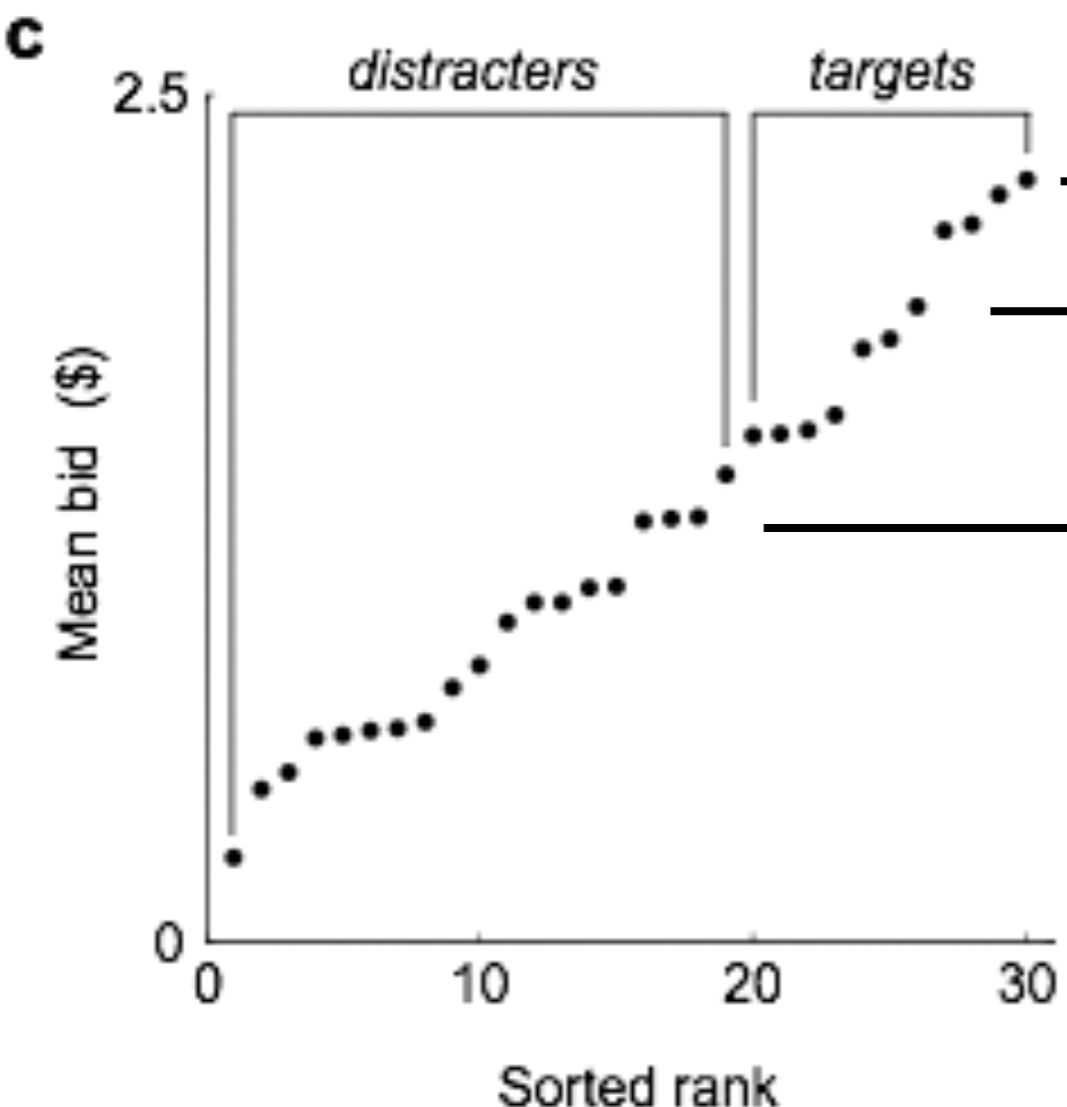
Simulated activity



Additional alternatives

Divisive normalization + cortical variability → *choice-set effect*

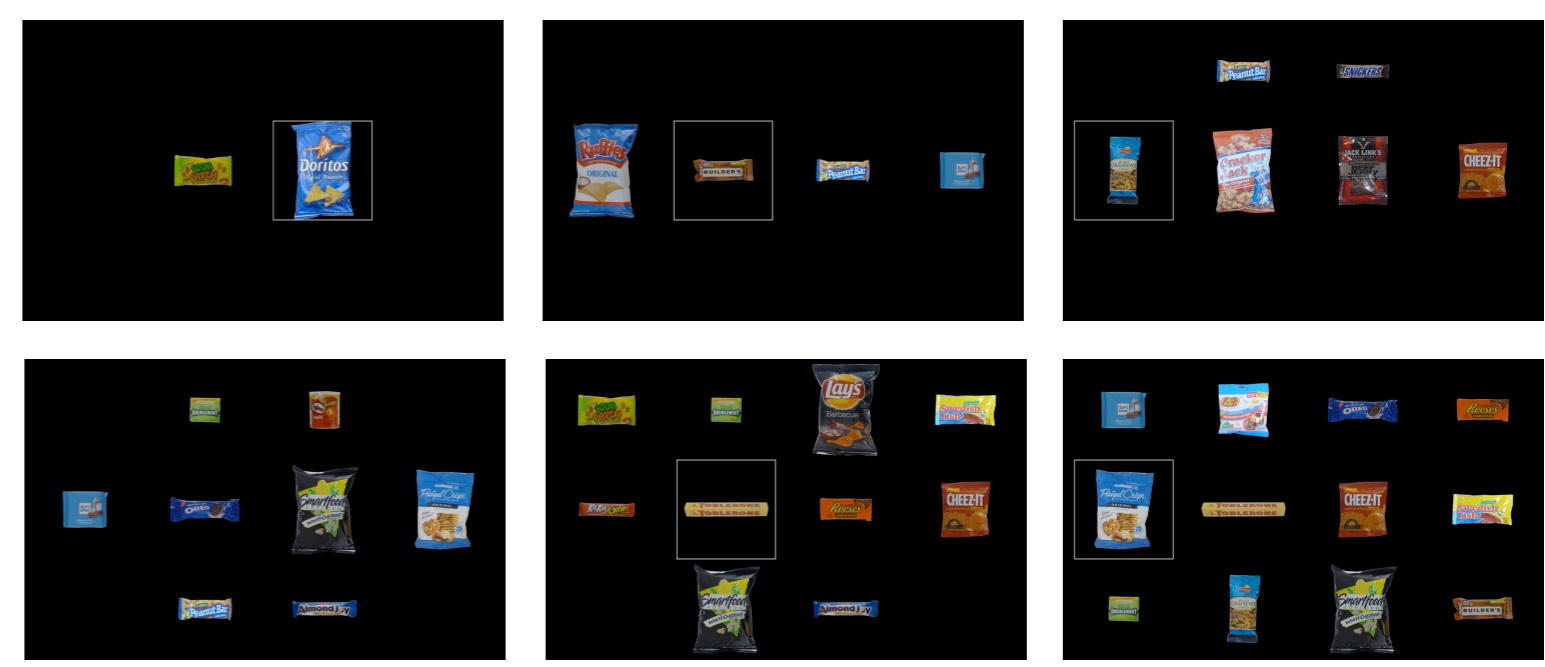
etc.



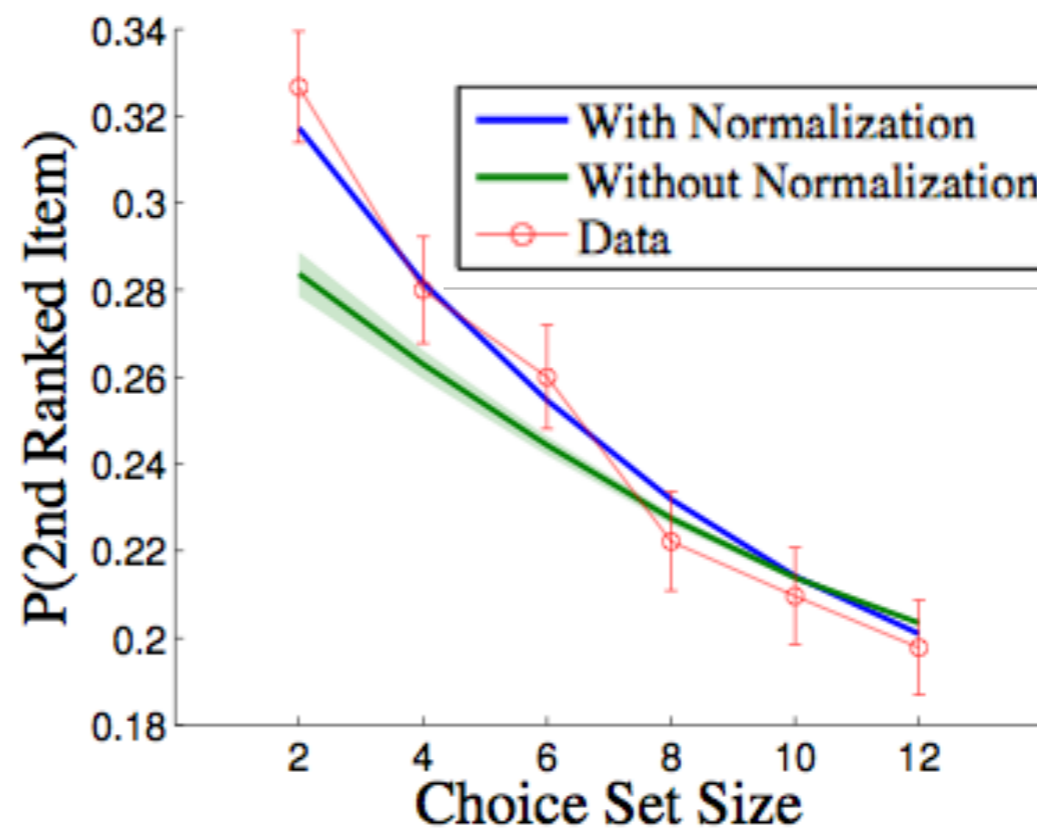
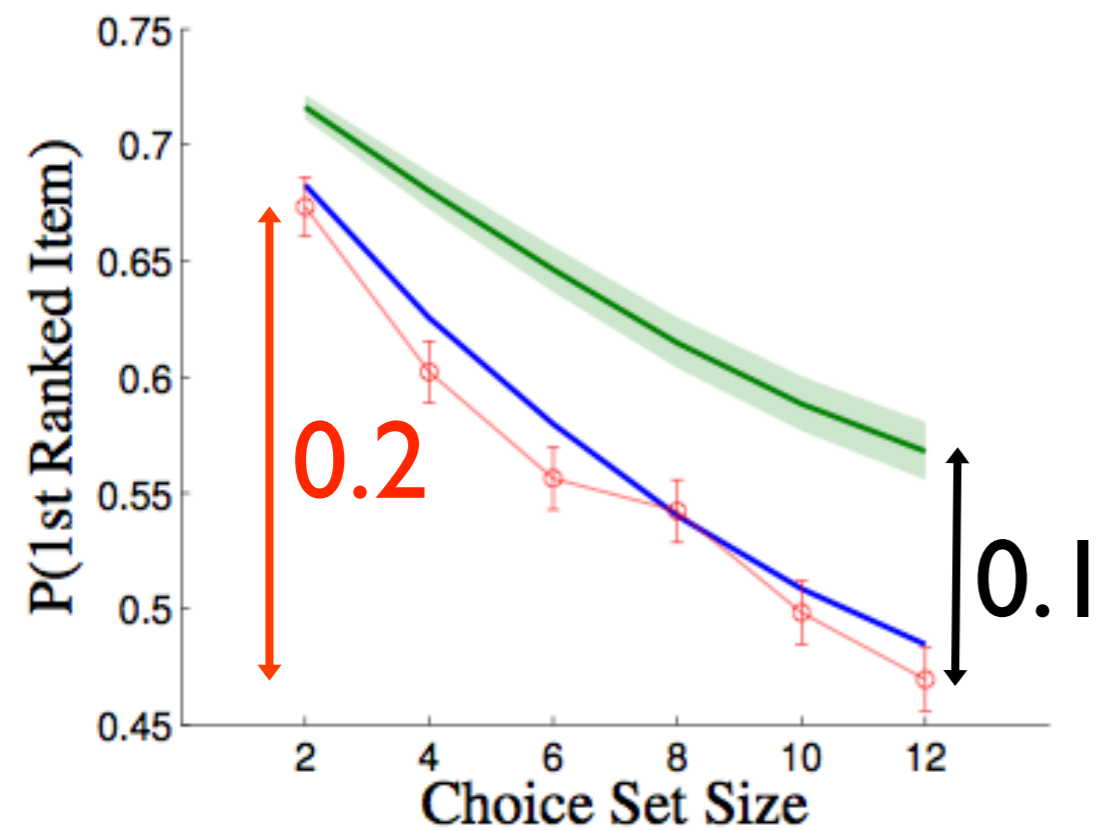
Target pairs

Distractors

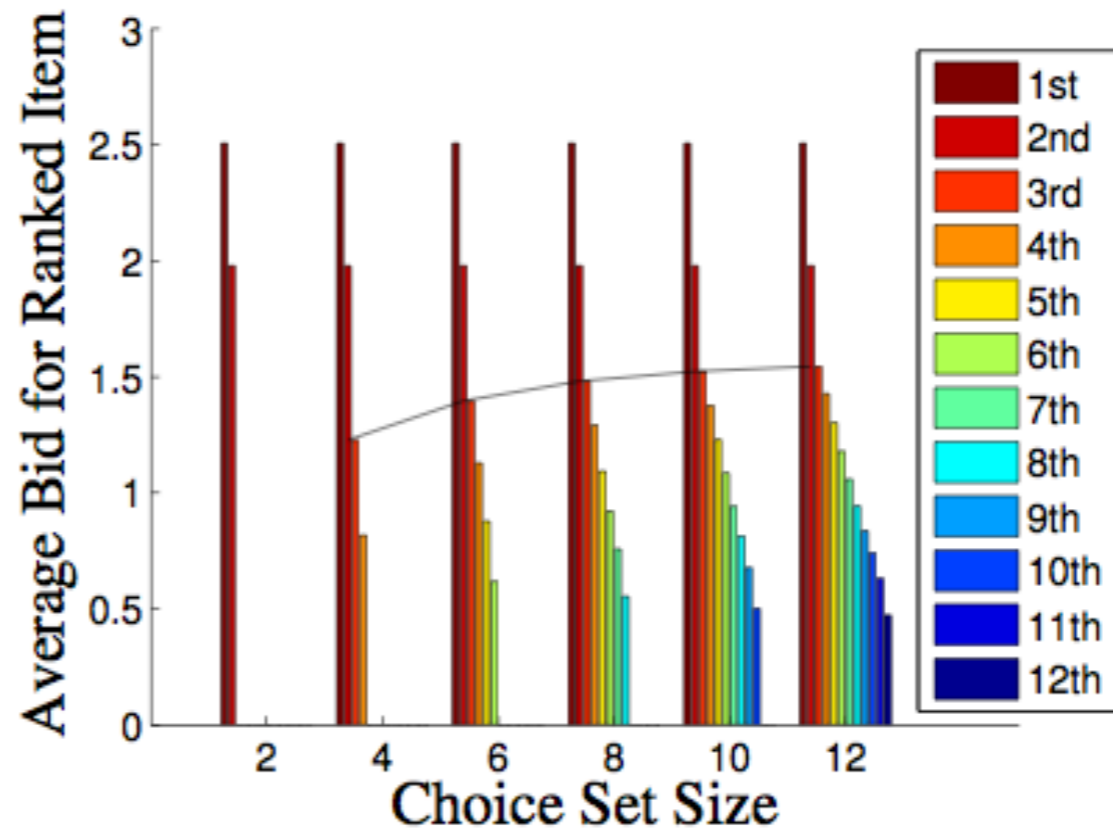
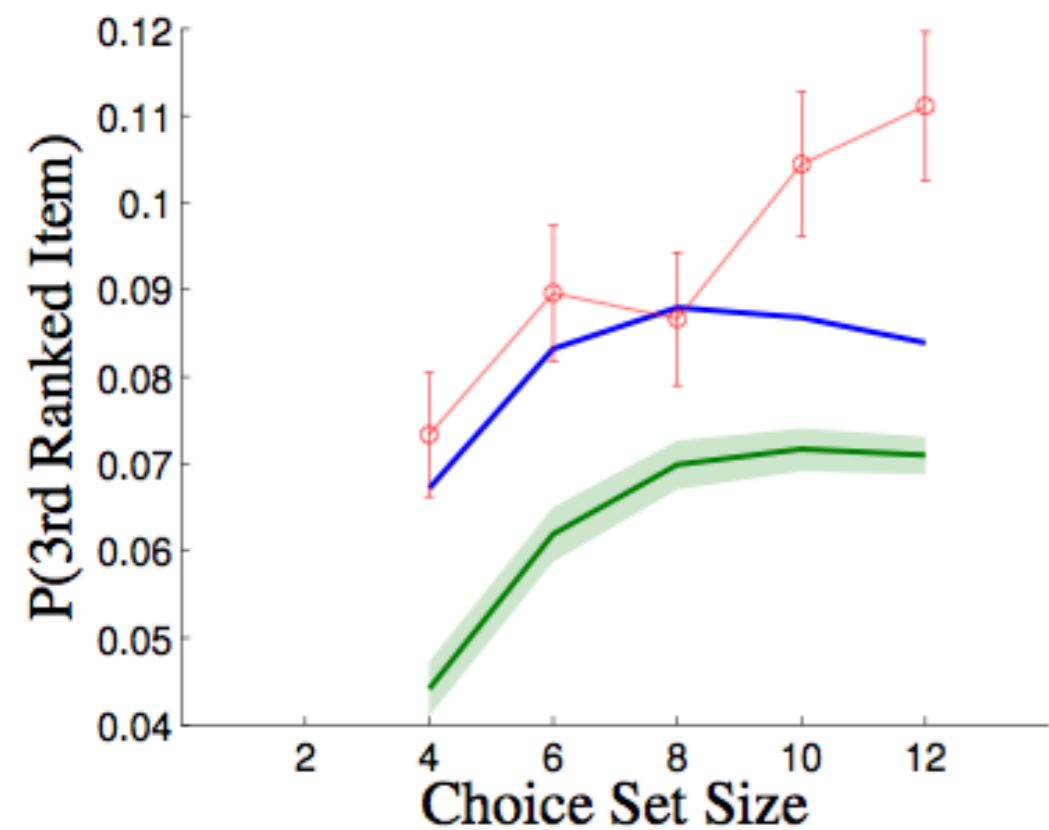
$$N \in \{2, 4, 6, 8, 10, 12\}$$



30 Subjects
 6x45 Trials Per
 8100 Trials Total

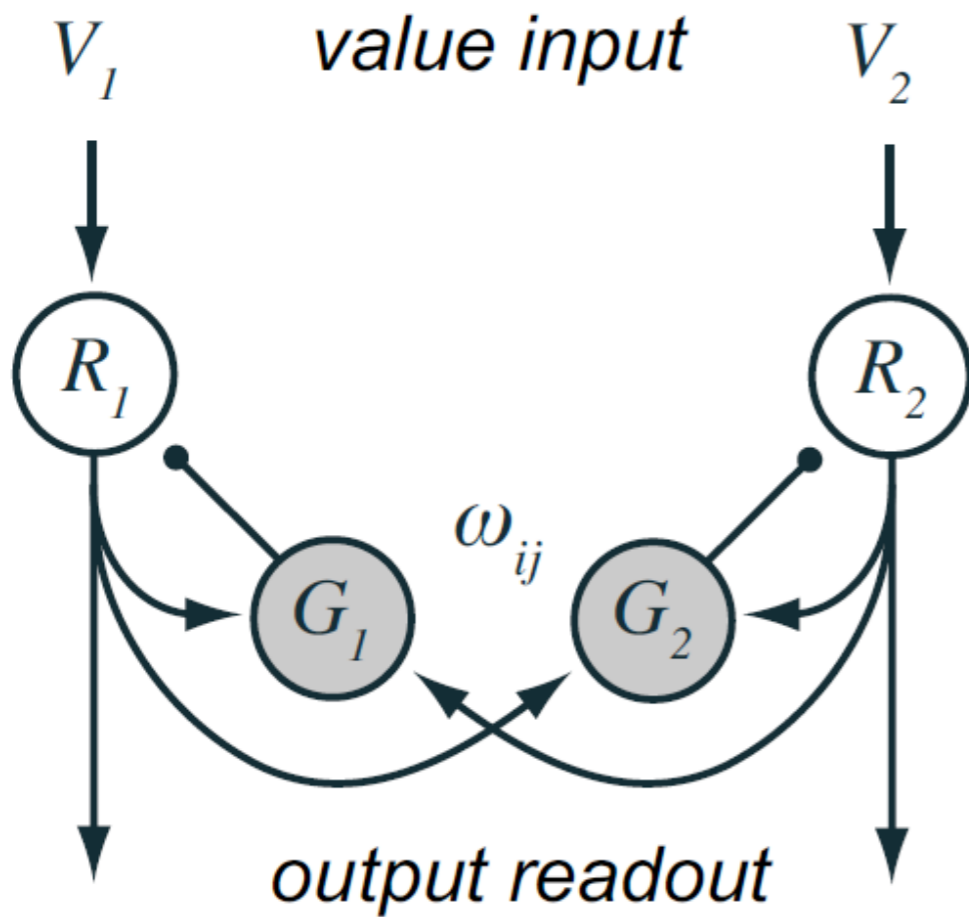


	$\hat{\Gamma}$	$\hat{\omega}$	$\hat{\beta}$
	0.00	0.44	18.8
	0.92	0	1



**But How Would A Network
Implement Normalization?**

**What Would the Network Dynamics
Be Like?**

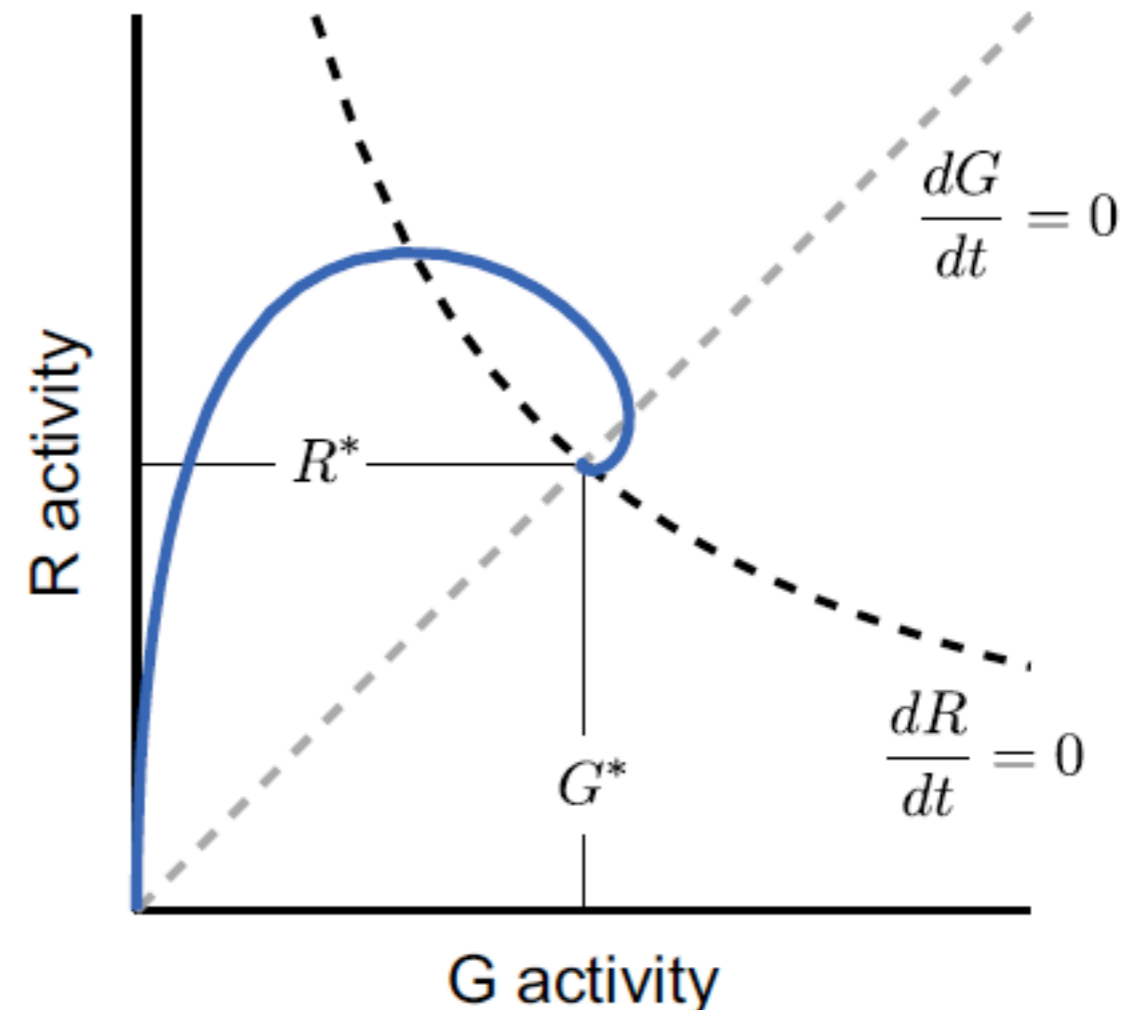
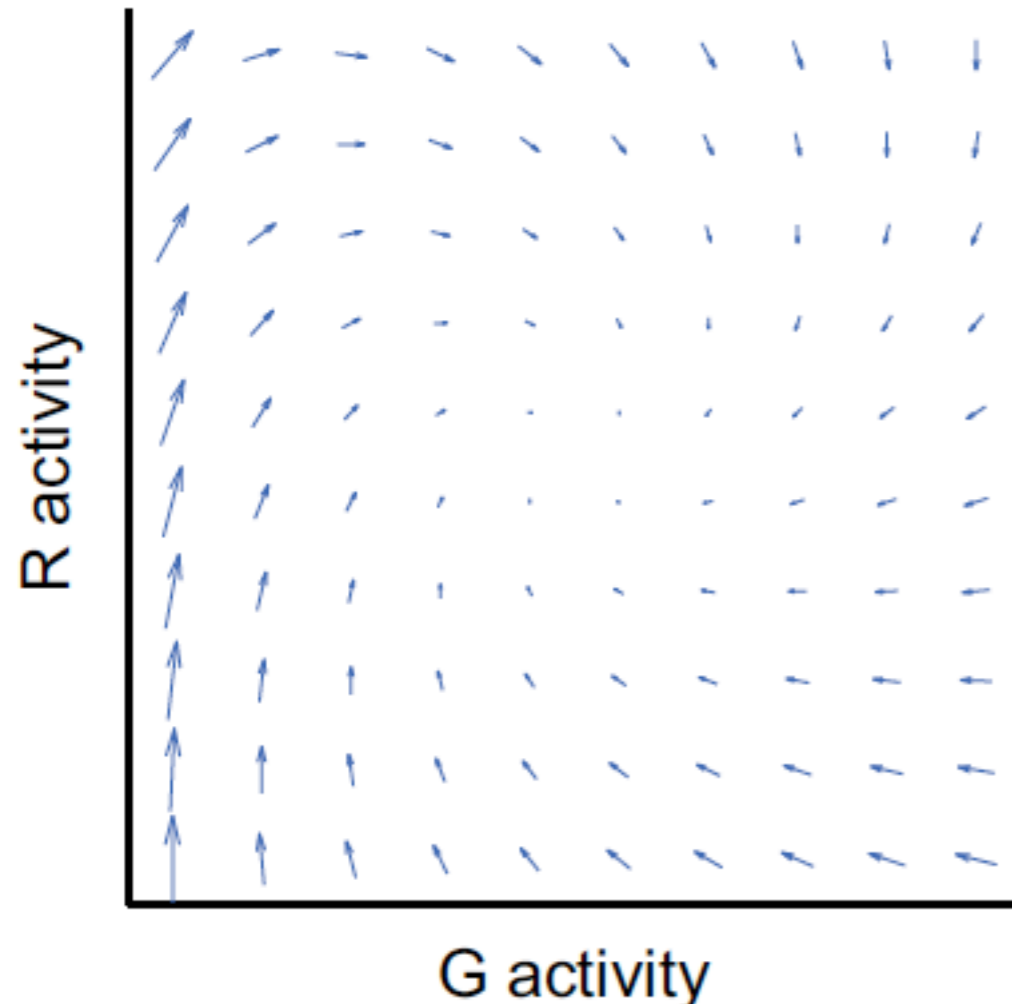


Excitatory neurons

$$\tau \frac{dR_i}{dt} = -R_i + \frac{V_i}{1 + G_i}$$

Inhibitory neurons

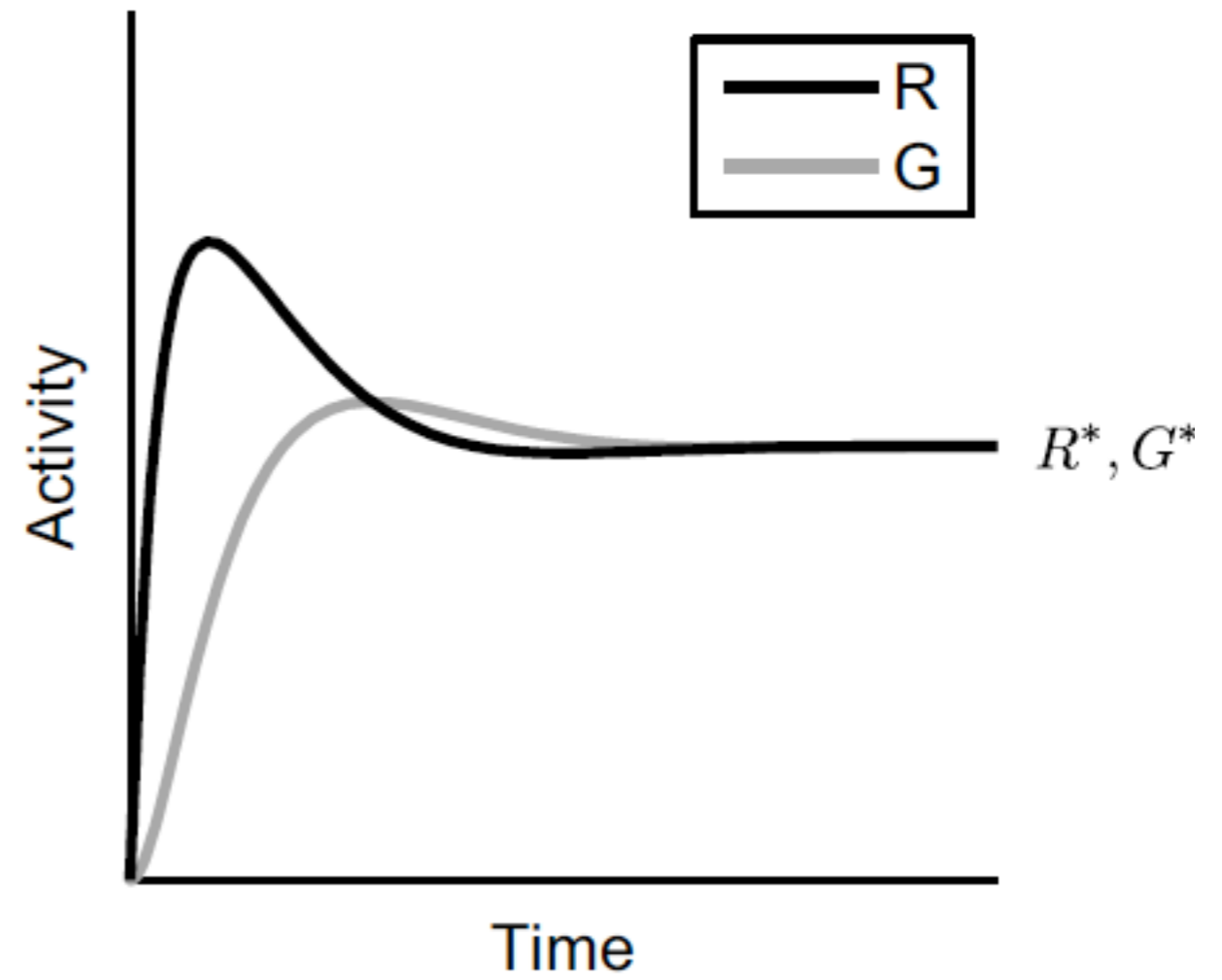
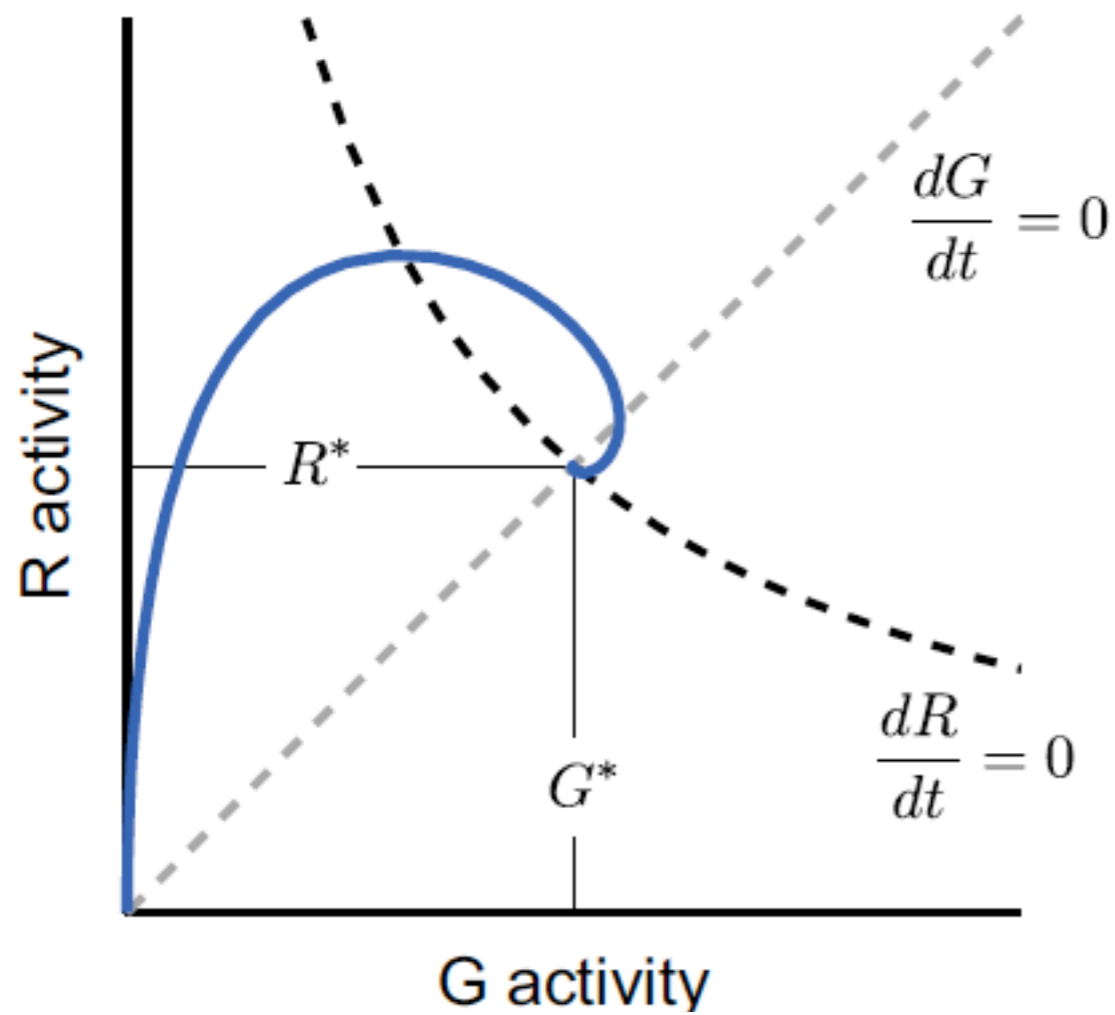
$$\tau \frac{dG_i}{dt} = -G_i + \sum_{j=1}^N \omega_{ij} R_j$$



**Normalization is the
Unique
Equilibrium State for
Networks
of this Kind**

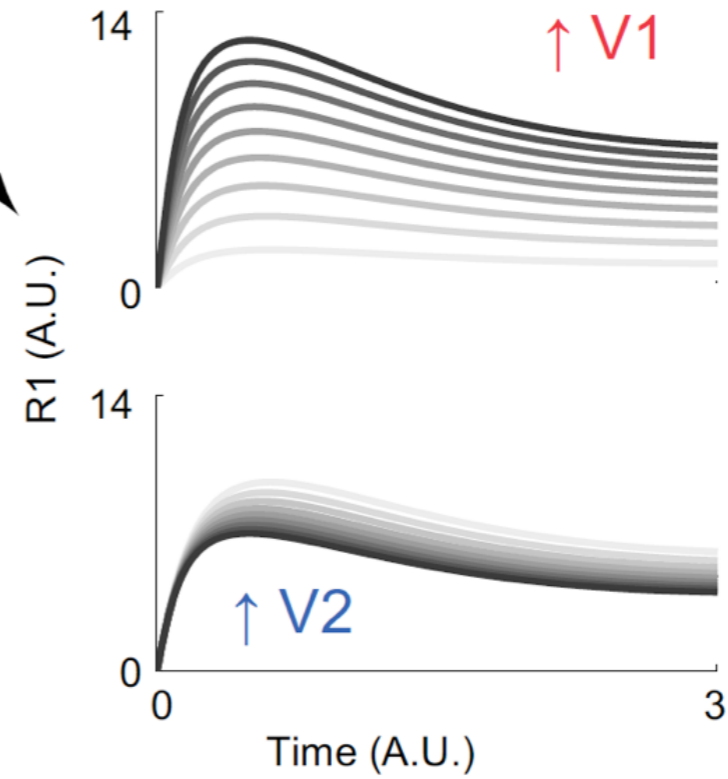
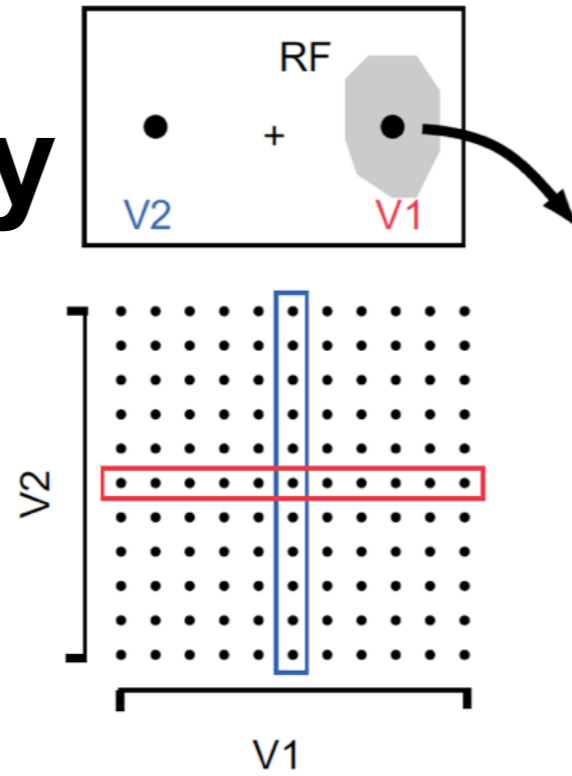


These Networks Have Specific Dynamics

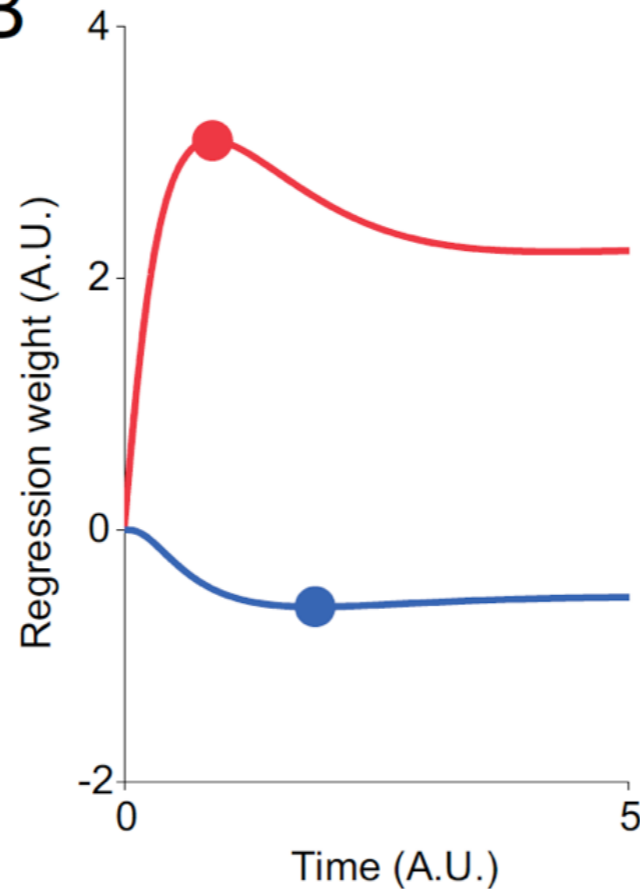


Which Actually Are Observed

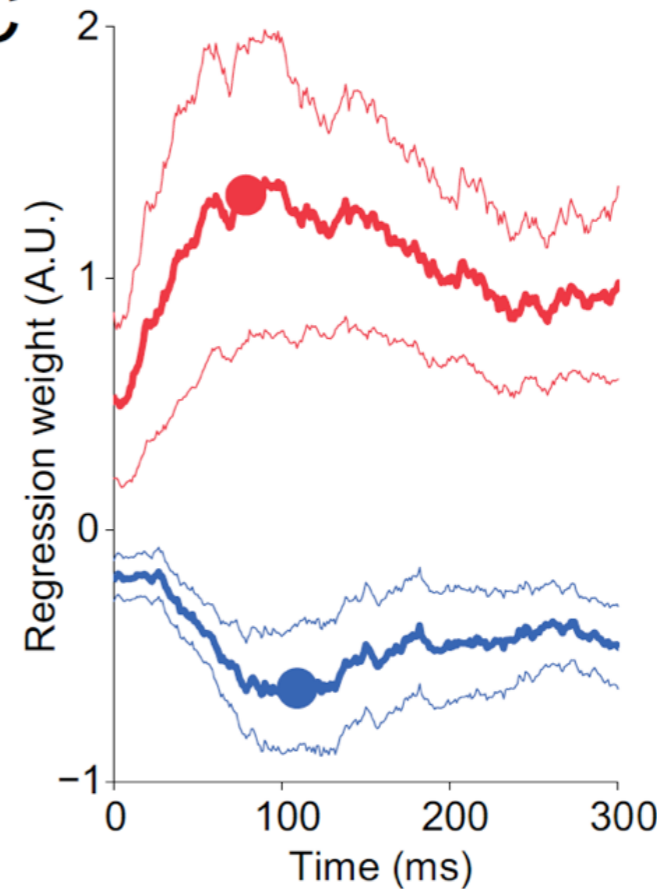
A



B



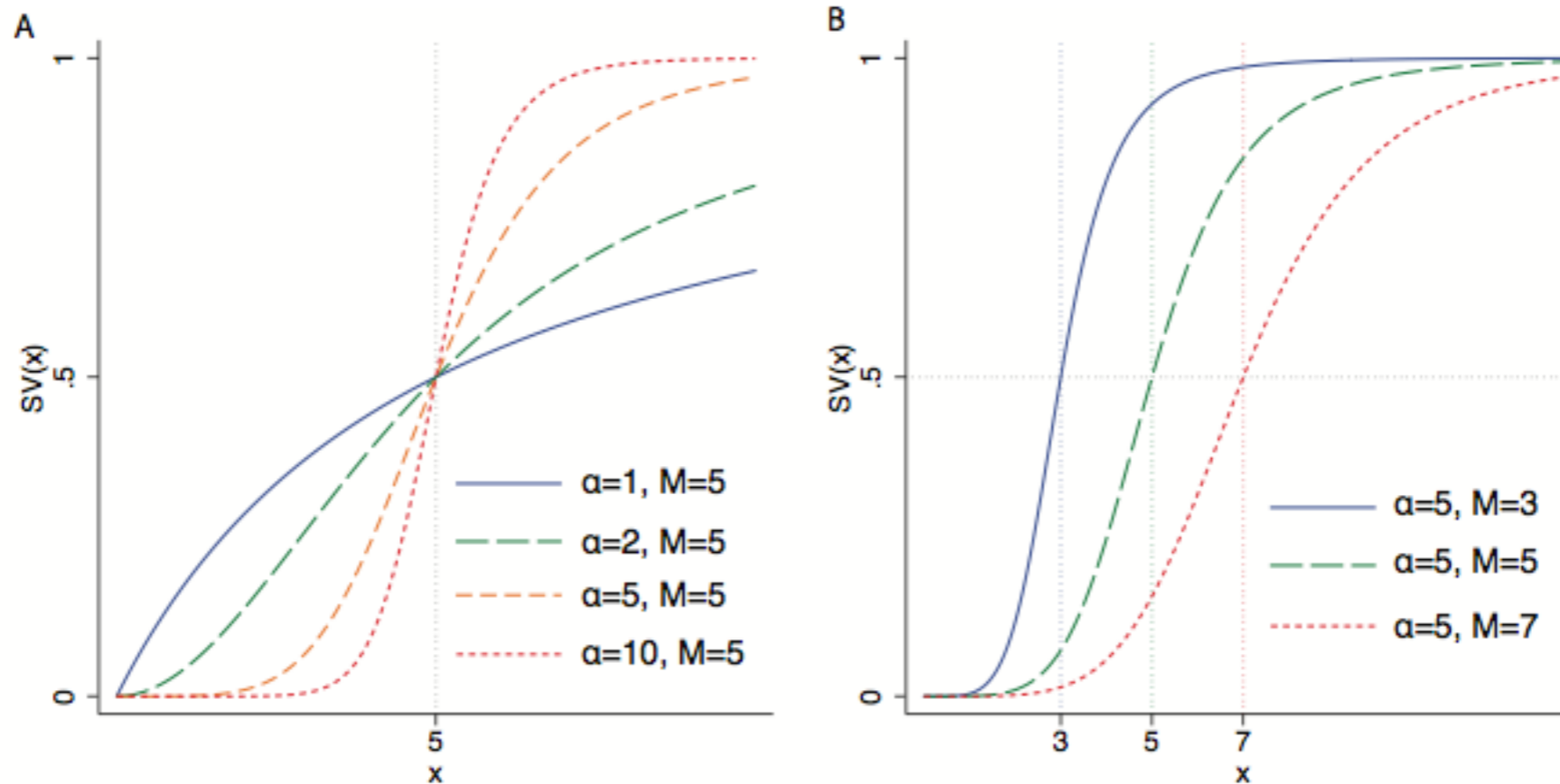
C



If a Choice Network Used These
Dynamics, How Would It Choose?

The ESVT “Value Function”

$$V(x) = \frac{x^\alpha}{x^\alpha + M_t^\alpha}$$

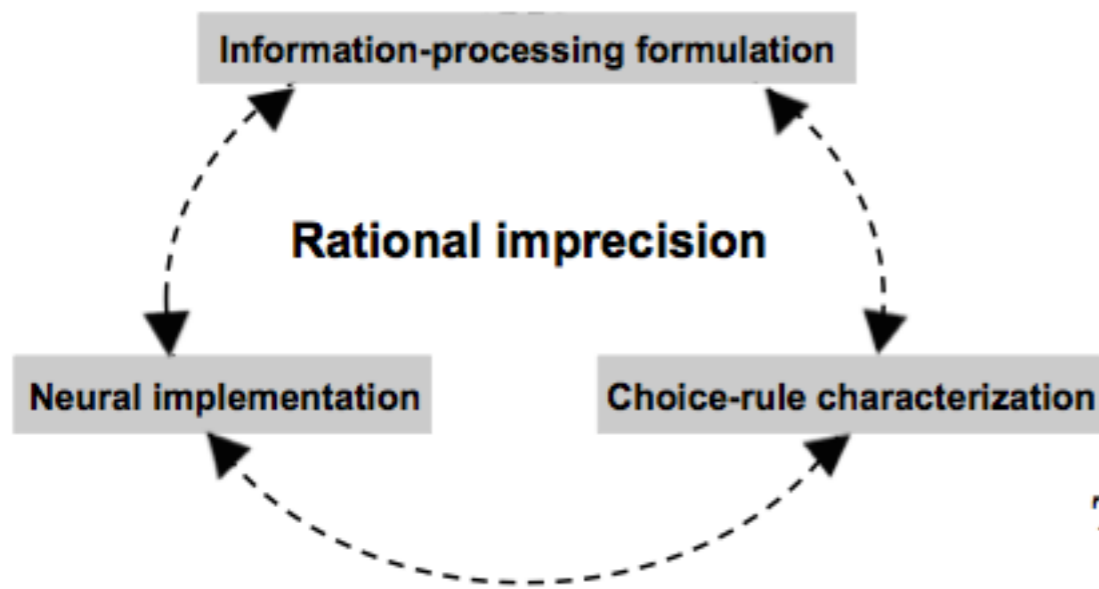


Captures:

- Reflection Effect
- Probability Distortion
- Loss Aversion
- Endowment Effect



Would It Be Normative?



Theorem 1. *For any random choice rule ρ the following are equivalent:*

1. ρ has an information-processing formulation.
2. ρ has a neural-normalization implementation.
3. ρ has a relaxed IIA characterization.

Definition 1. A random choice rule ρ has an **information-processing formulation** if there exist functions $v : X \rightarrow (0, \infty)$ and $F : \mathcal{A} \rightarrow (0, \infty)$ such that for all $A \in \mathcal{A}$:

$$\rho \in \arg \max_{\hat{\rho} \in \mathcal{P}} \left\{ \sum_{x \in A} \hat{\rho}(x, A) v(x) - F(A) (H_{\max}(A) - H(\hat{\rho}, A)) \right\}.$$

Definition 2. A random choice rule ρ has a **neural-normalization implementation** if there exist functions $v : X \rightarrow (0, \infty)$ and $F : \mathcal{A} \rightarrow (0, \infty)$ such that for all A in \mathcal{A} and x in A :

$$\rho(x, A) = \Pr \left(x = \arg \max_{y \in A} \frac{v(y)}{F(A)} + \varepsilon_y \right),$$

where the ε_y 's are i.i.d. and Gumbel with location 0 and scale 1.

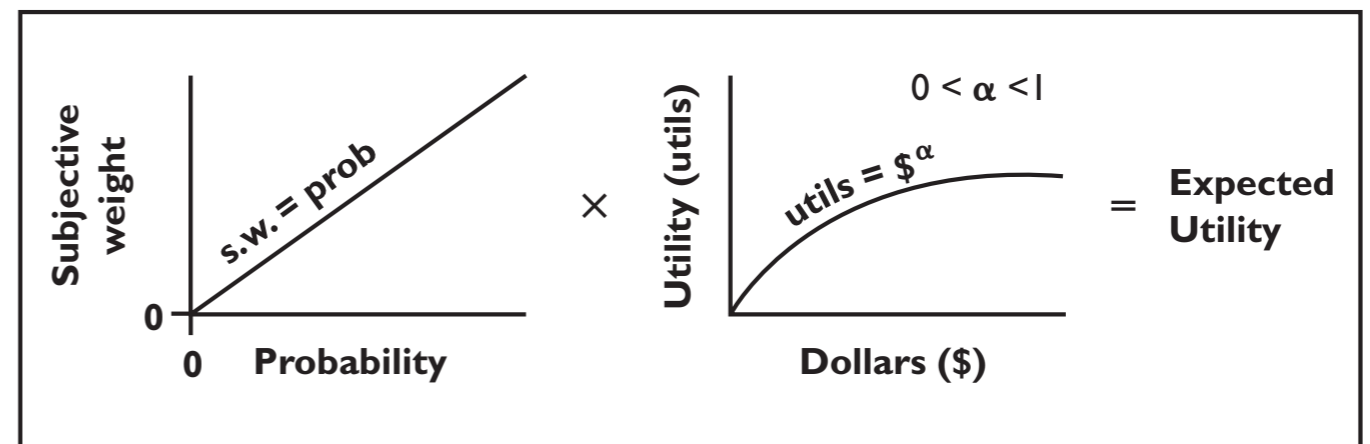
Definition 5. A random choice rule ρ is a **relaxed IIA** rule if there exists an admissible family of functions $\{G_A\}_{A \in \mathcal{A}}$ with respect to which ρ obeys Equation (1).



John von Neumann Oskar Morgenstern



Modern Expected *Utility* Theory



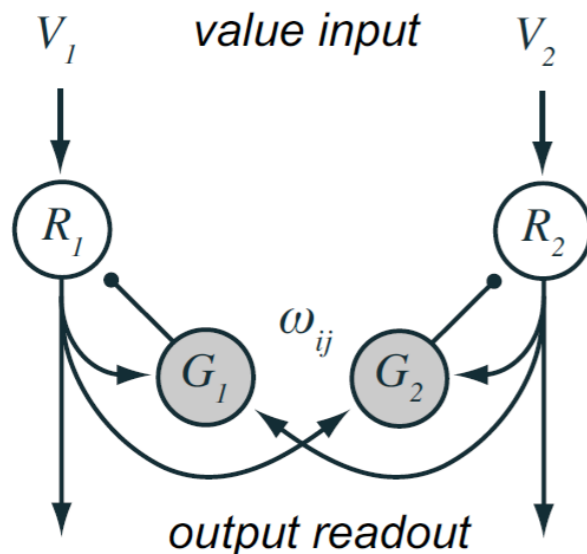
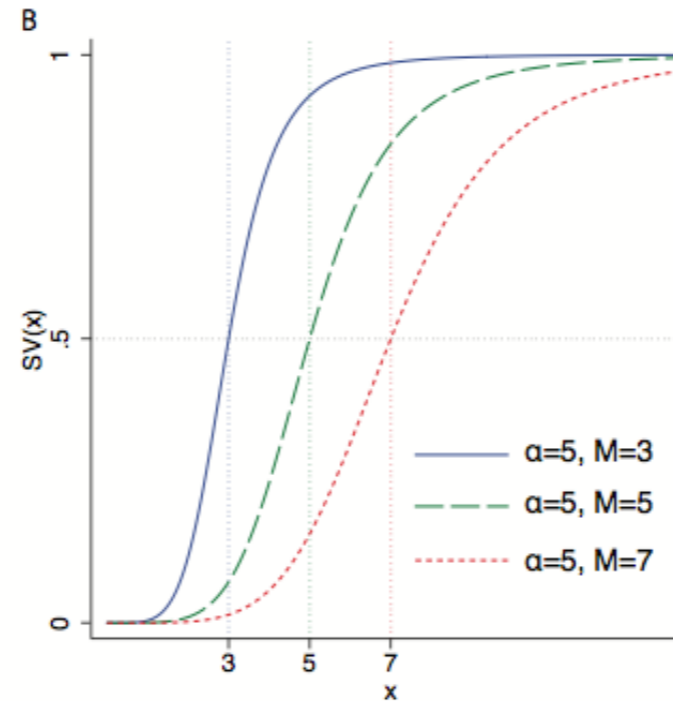
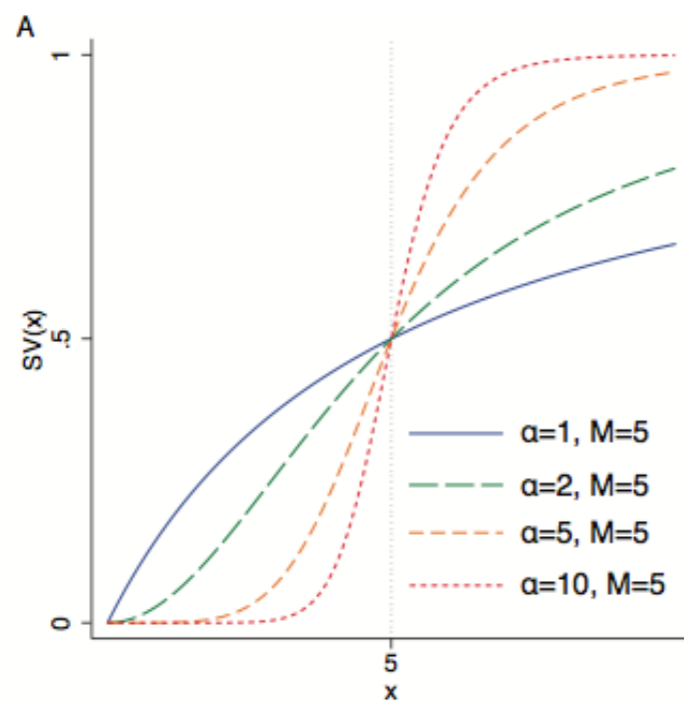
Critical Advantages:

- Precise
- Compact
- **Normative** (people make sense)

Expected Subjective Value Theory

Definition 1. A random choice rule ρ has an **information-processing formulation** if there exist functions $v : X \rightarrow (0, \infty)$ and $F : \mathcal{A} \rightarrow (0, \infty)$ such that for all $A \in \mathcal{A}$:

$$\rho \in \arg \max_{\hat{\rho} \in \mathcal{P}} \left\{ \sum_{x \in A} \hat{\rho}(x, A) v(x) - F(A) (H_{\max}(A) - H(\hat{\rho}, A)) \right\}.$$



Excitatory neurons

$$\tau \frac{dR_i}{dt} = -R_i + \frac{V_i}{1 + G_i}$$

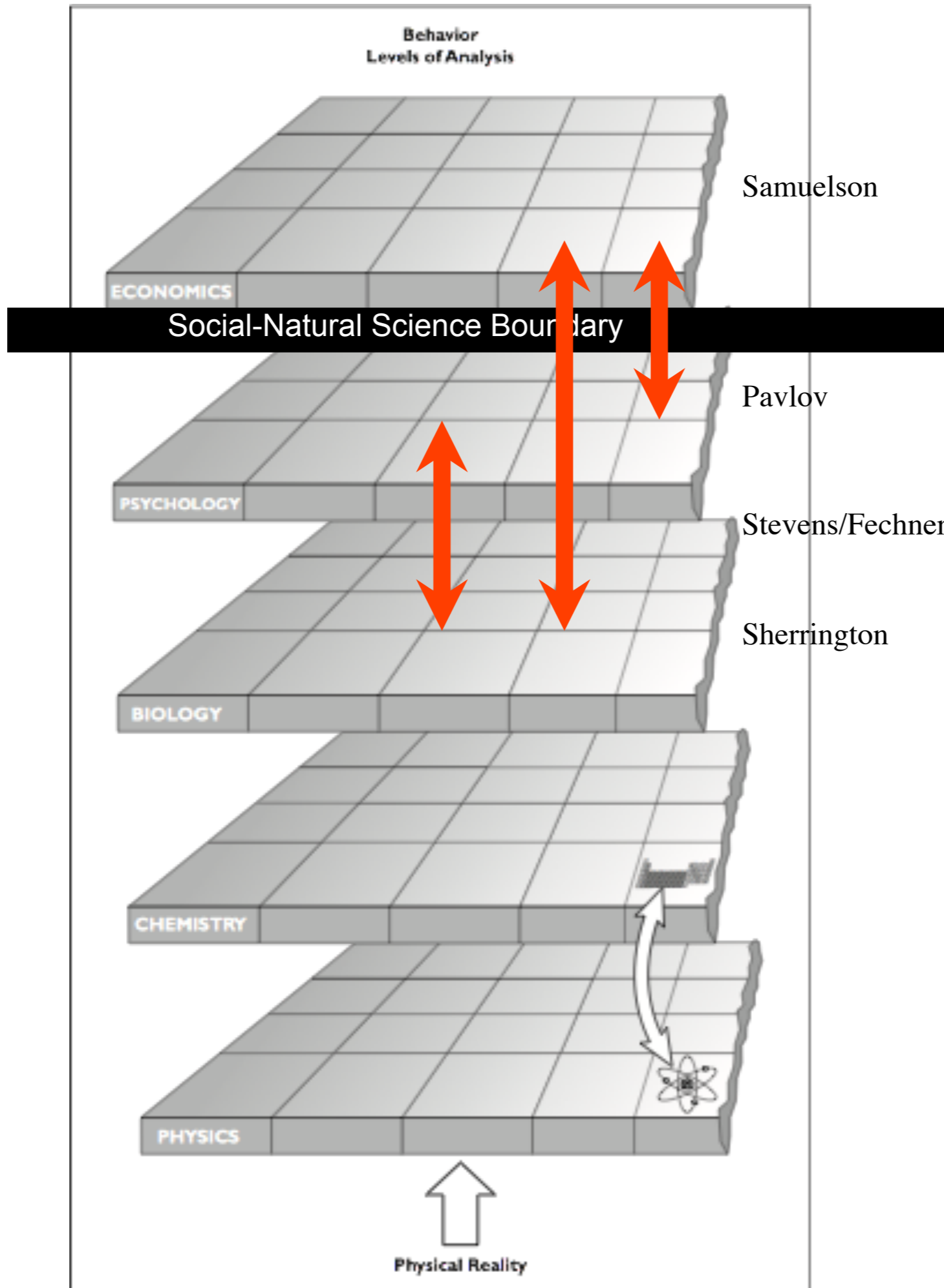
Inhibitory neurons

$$\tau \frac{dG_i}{dt} = -G_i + \sum_{j=1}^N \omega_{ij} R_j$$

Economics

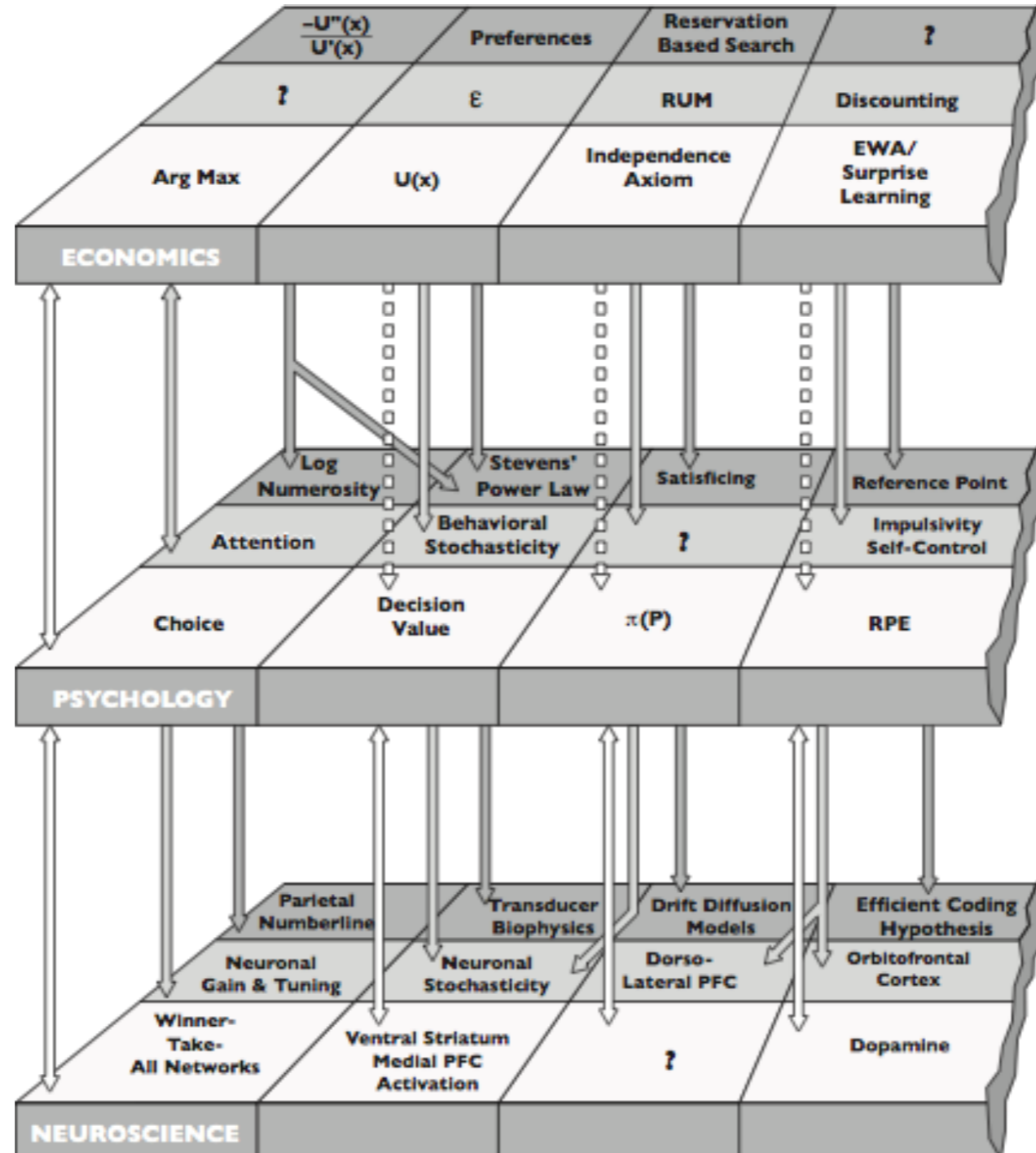
Psychology

Biology



Neuroeconomics

Aligning and Refining Hard Theories





National Institute on Aging ■ ◆ ★ ✨



James S. McDonnell Foundation

