NEUROECONOMICS LABORATORY

Berkeley
UNIVERSITY OF CALIFORNIA

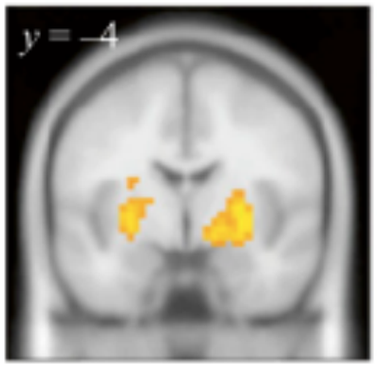
Learning in Games:
Neural and Molecular Mechanisms

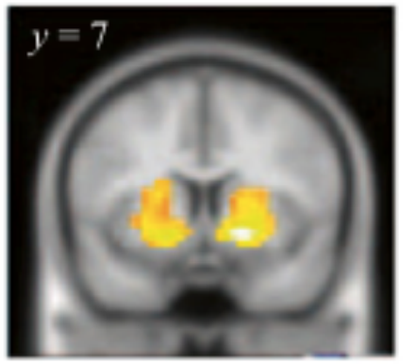
Ming Hsu

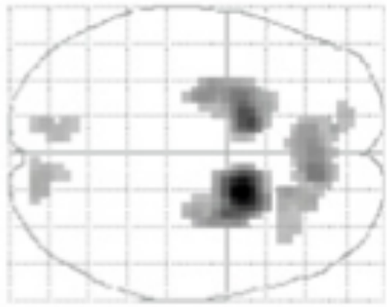
Haas School of Business and
Helen Wills Neuroscience Center
University of California, Berkeley

A Road Map

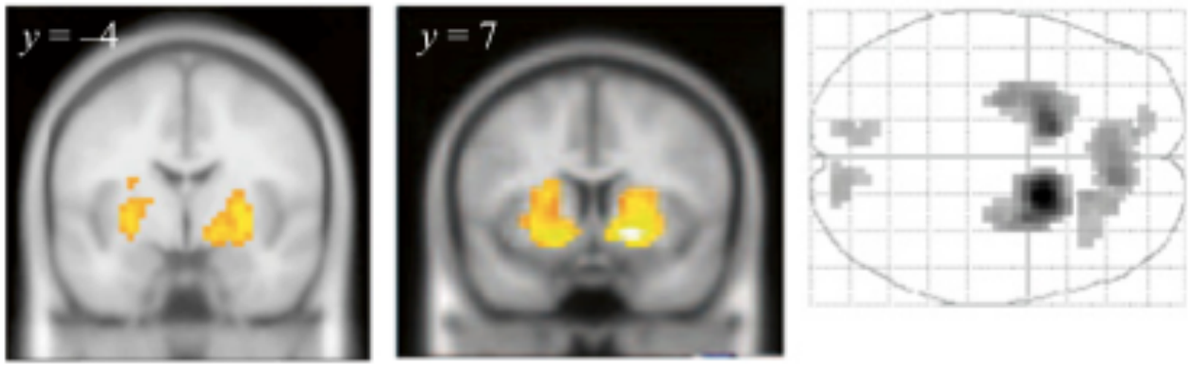
A Road Map

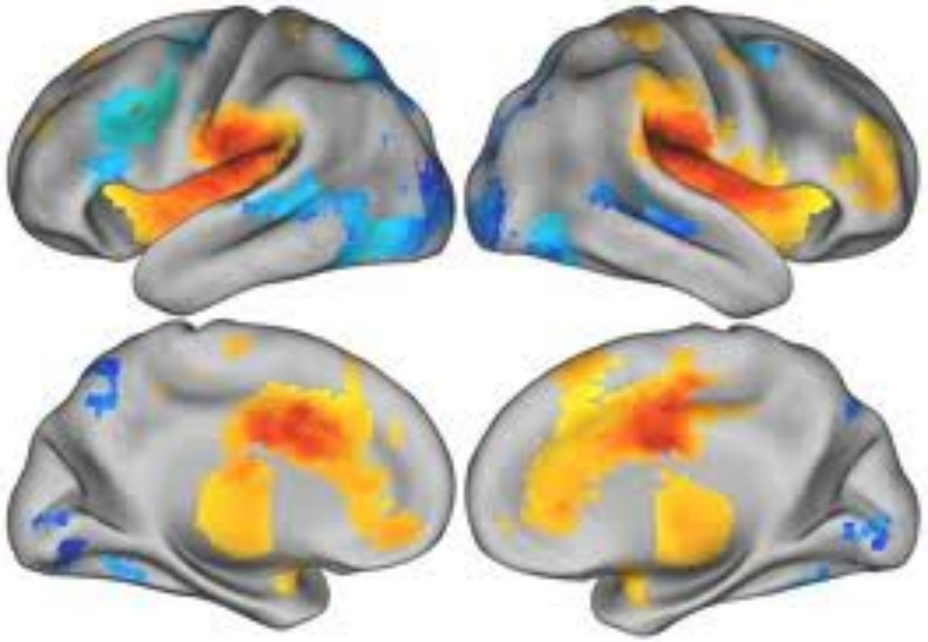




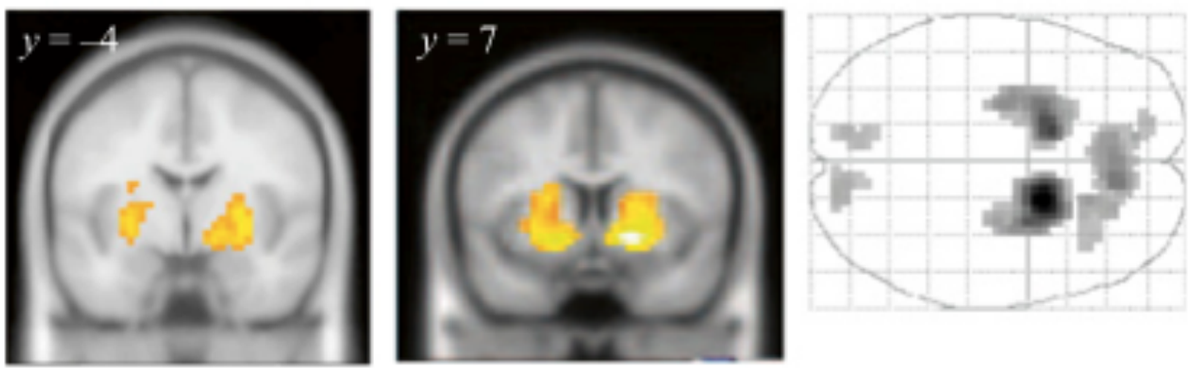


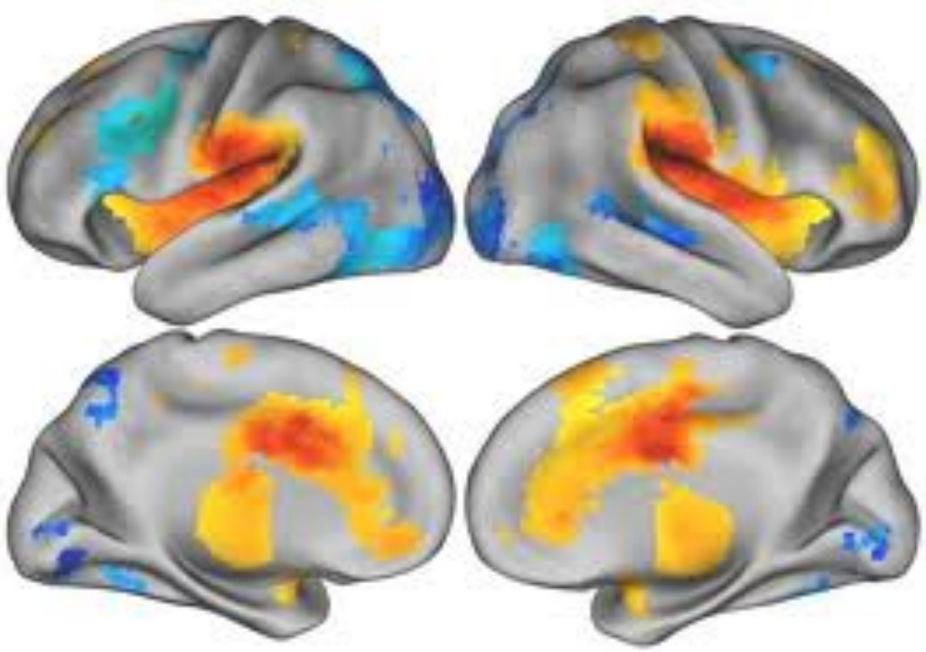
A Road Map





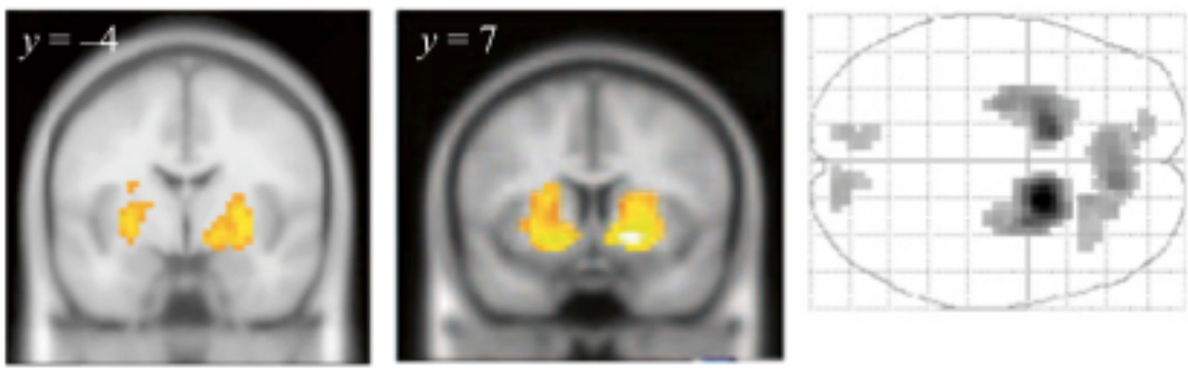
A Road Map

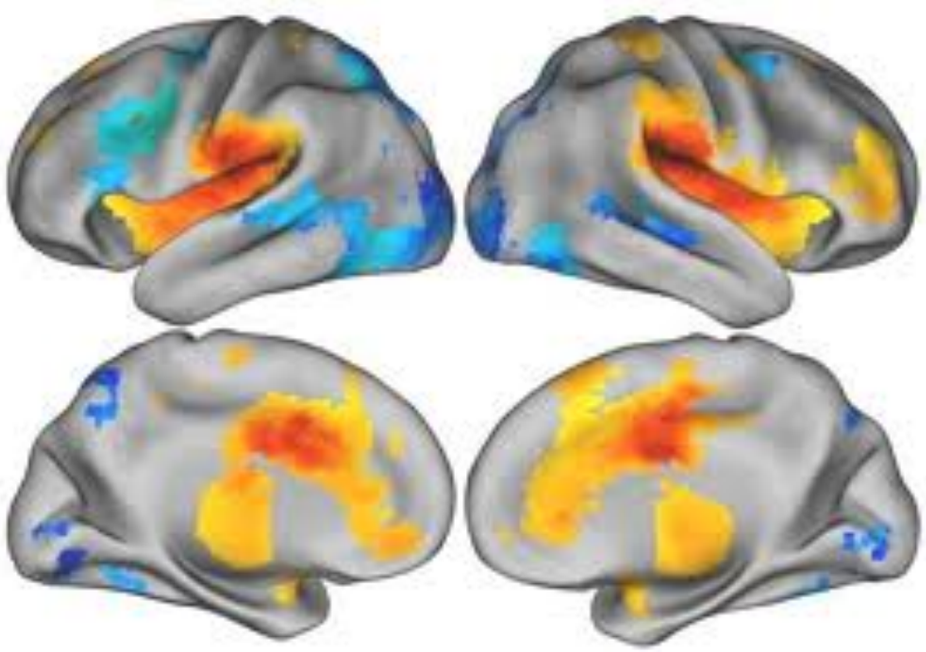




Goals

A Road Map

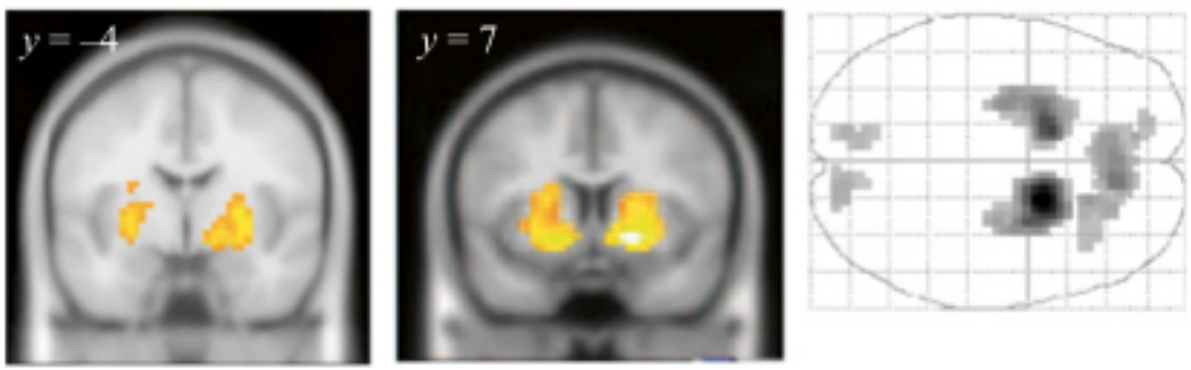


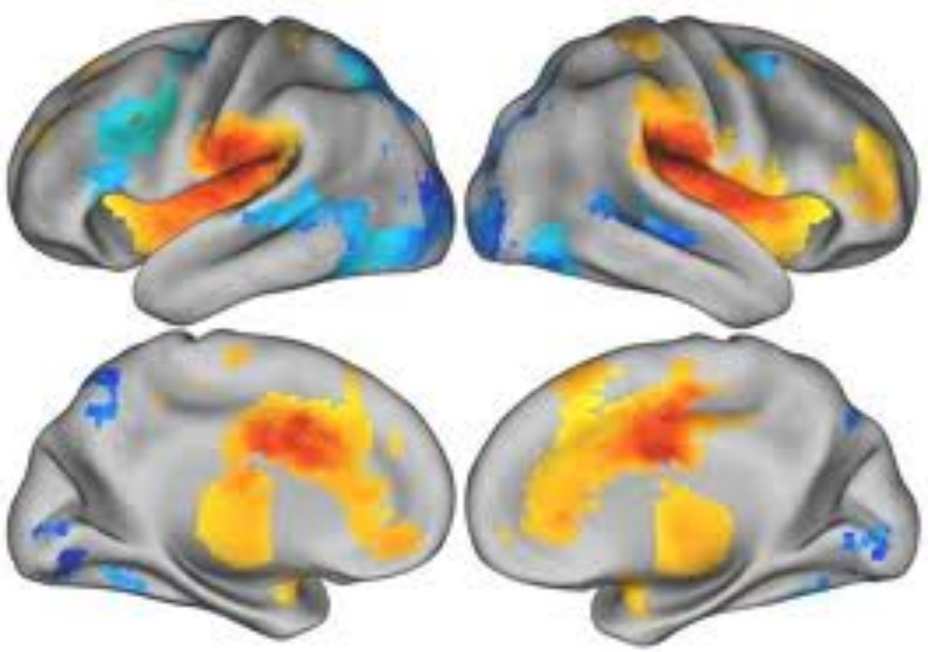


Goals

- Creating maps

A Road Map

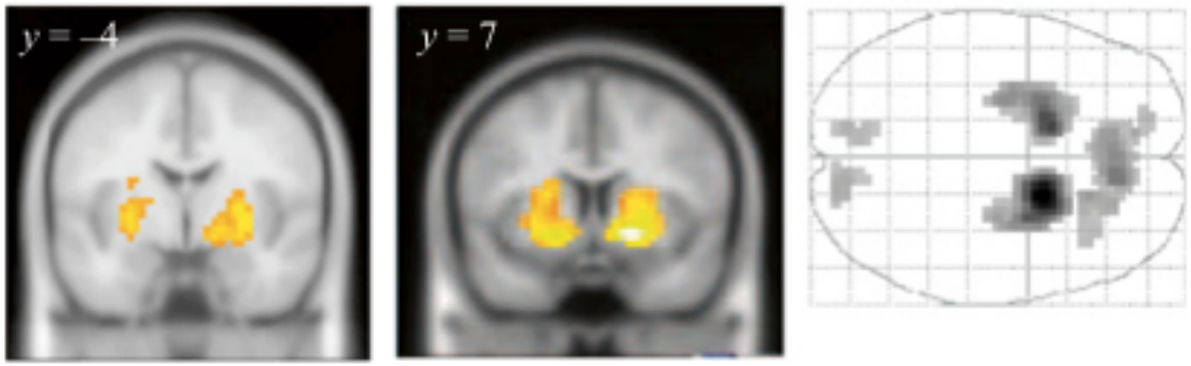


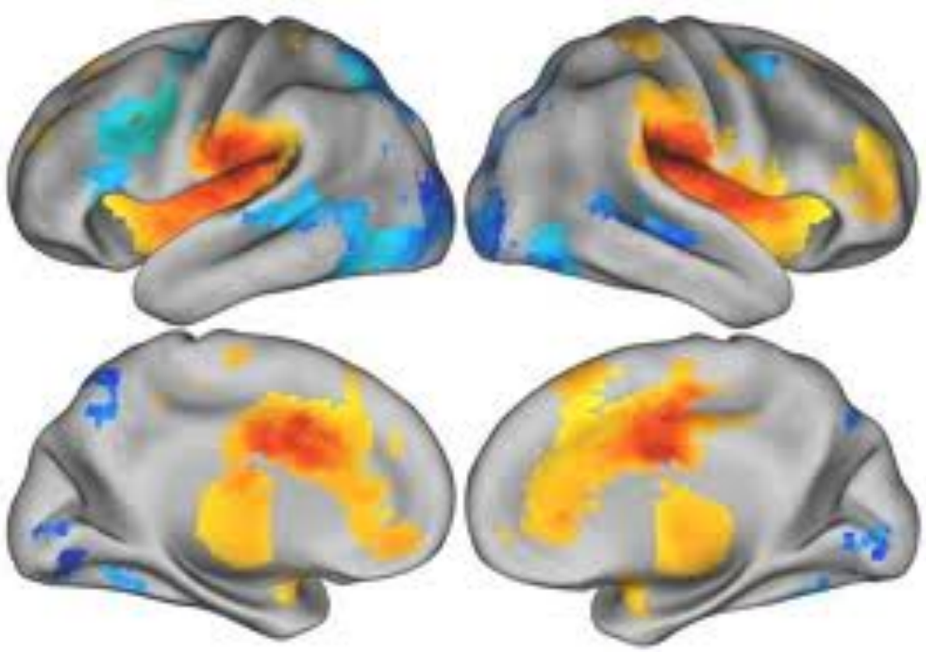


Goals

- Creating maps
- Connect layers

A Road Map

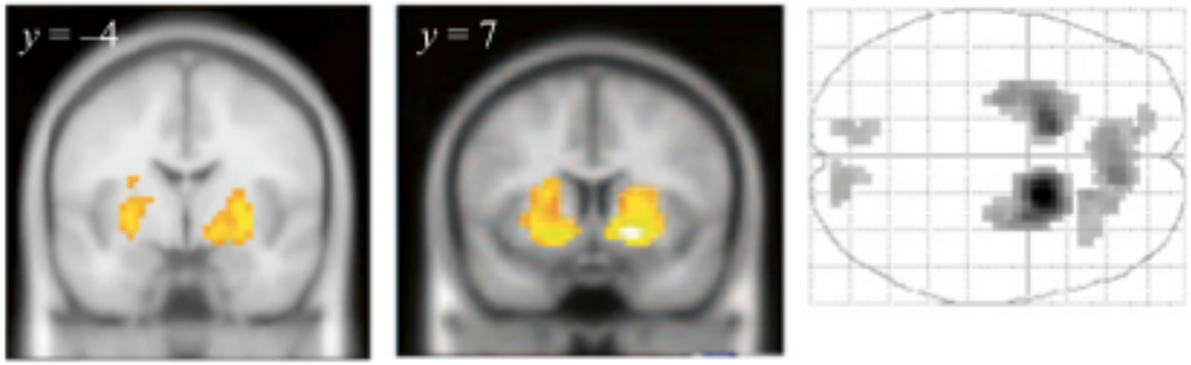


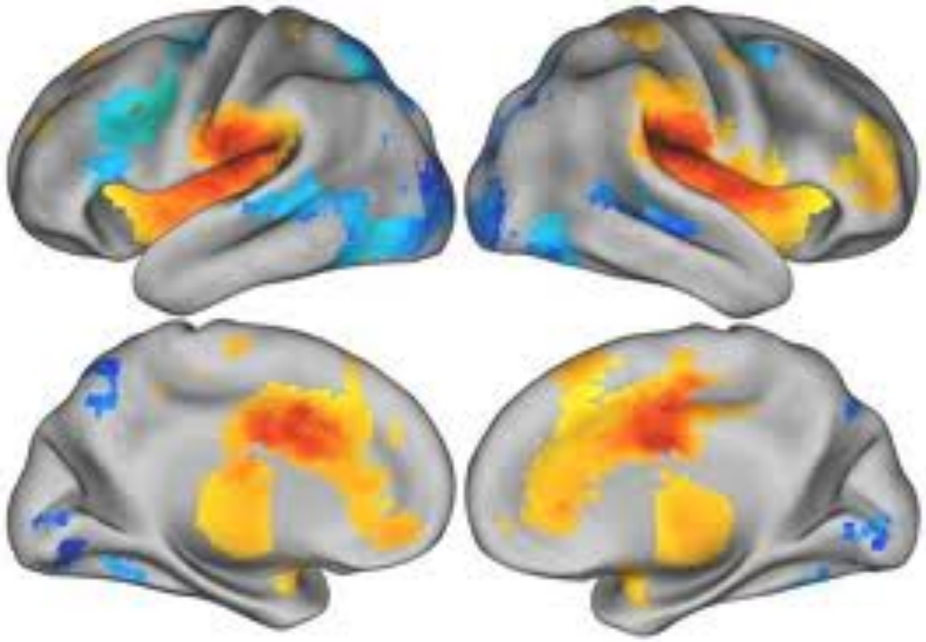


Goals

- Creating maps
- Connect layers
- Inform actions

A Road Map

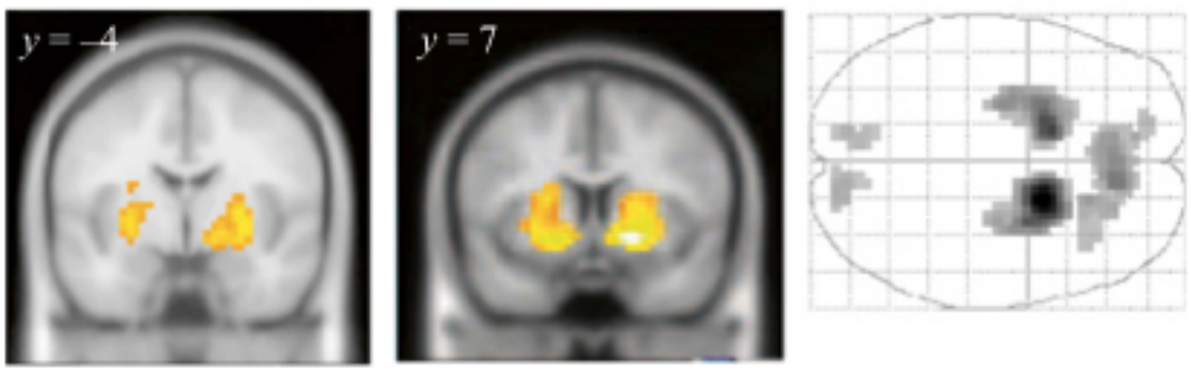


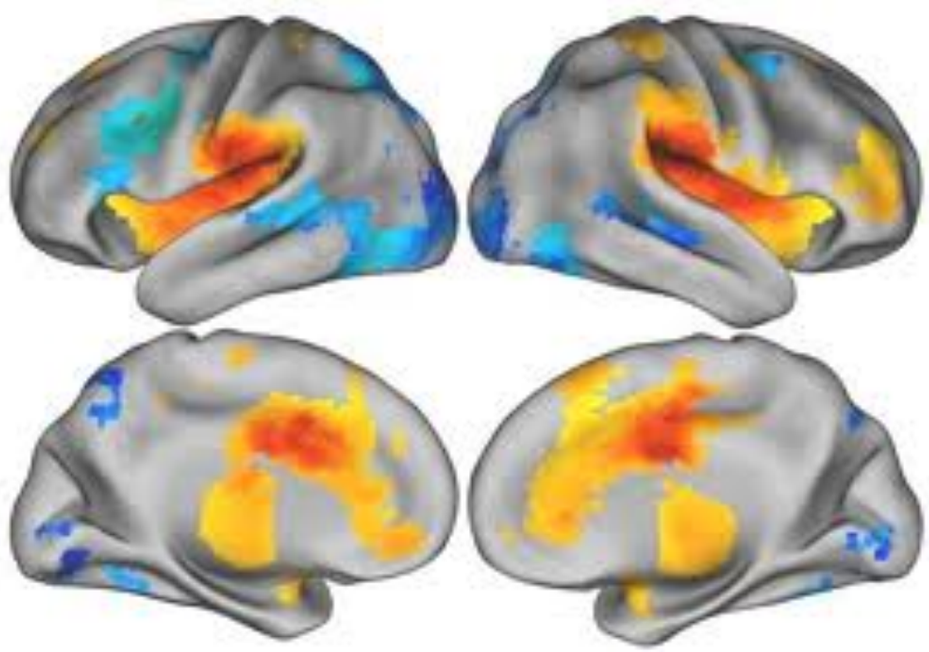


Goals

- Creating maps
- Connect layers
- Inform actions
  - Simplification

A Road Map

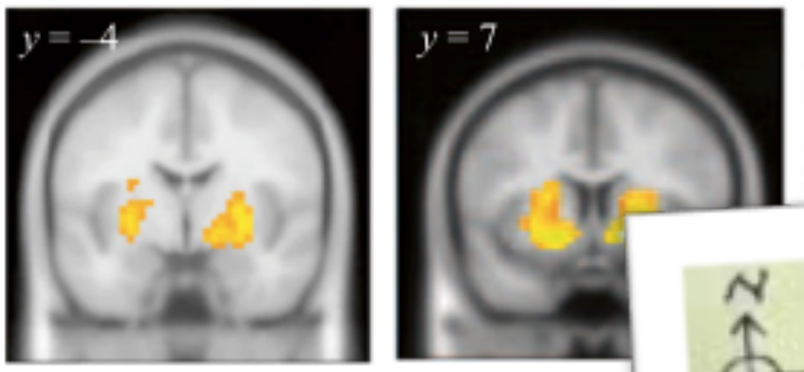


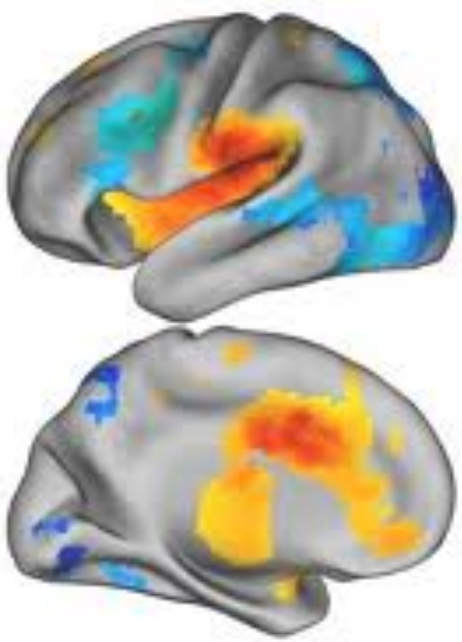


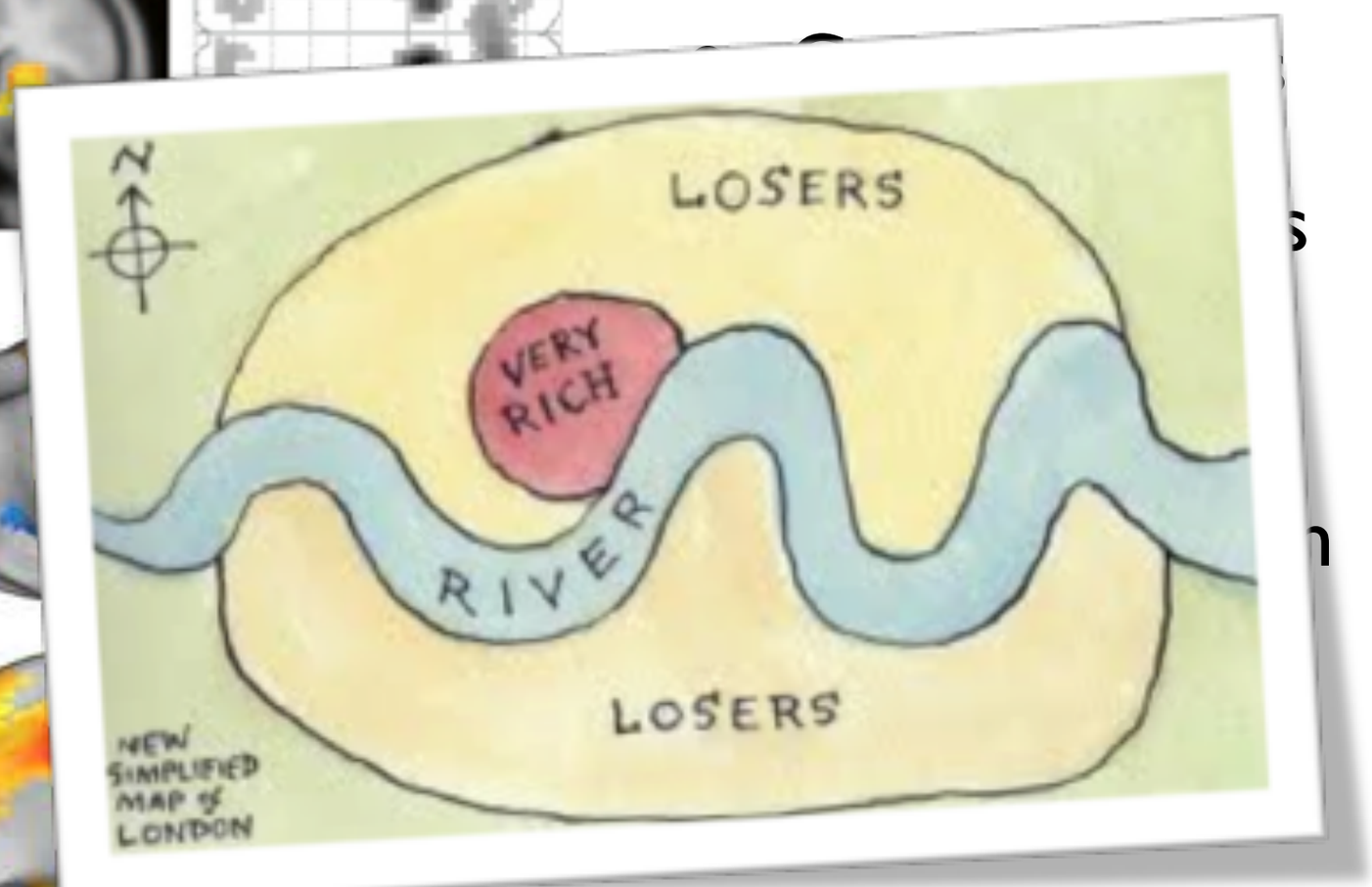
Goals

- Creating maps
- Connect layers
- Inform actions
  - Simplification
  - Actionable

A Road Map







Goals

A Road Map

A Road Map

A Road Map

A Road Map

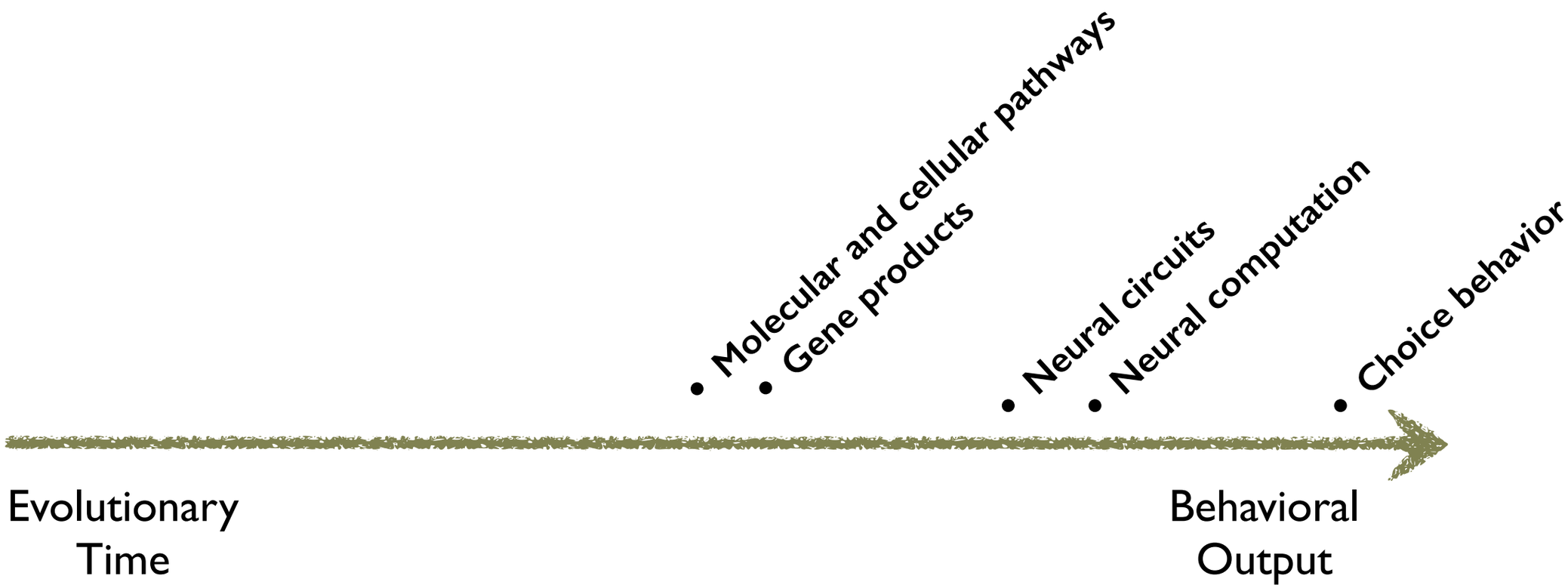
A Road Map


Neural computation
Choice behavior
Evolutionary Time
Behavioral Output

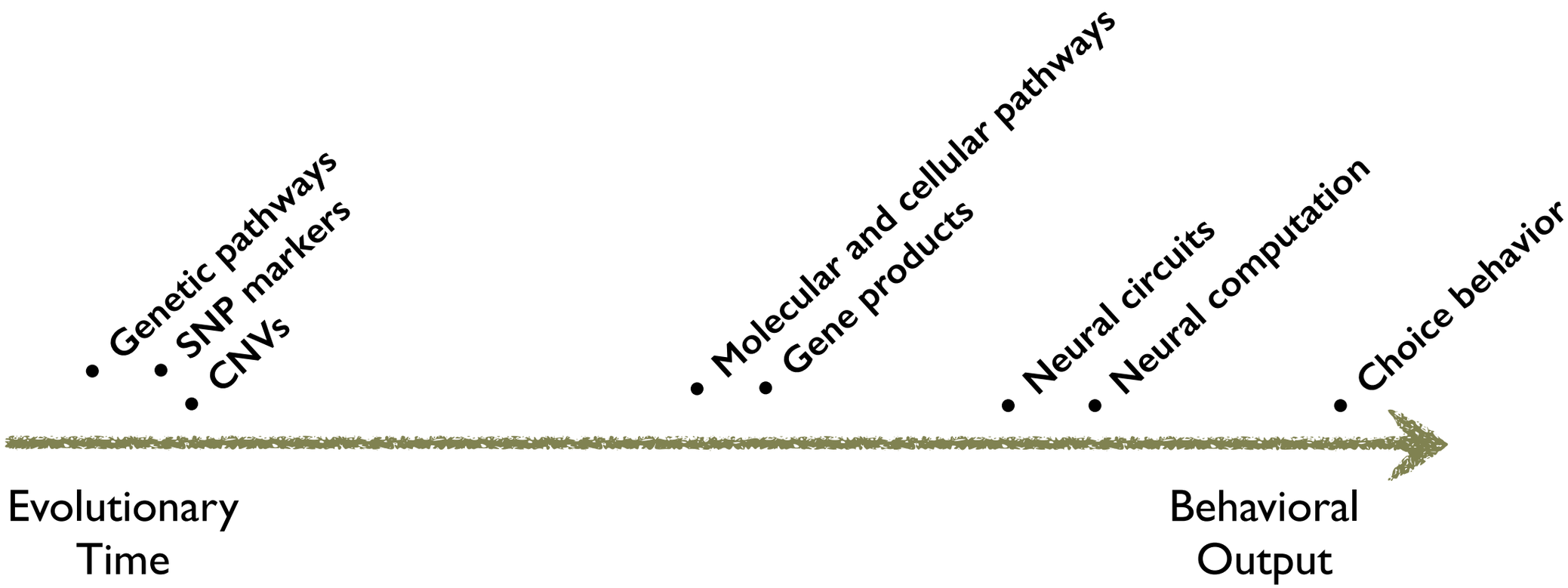
A Road Map


Neural circuits
Neural computation
Choice behavior
Evolutionary Time
Behavioral Output

A Road Map


• Molecular and cellular pathways
• Gene products
• Neural circuits
• Neural computation
• Choice behavior
Evolutionary Time
Behavioral Output

A Road Map


• Genetic pathways
• SNP markers
• CNVs
• Molecular and cellular pathways
• Gene products
• Neural circuits
• Neural computation
• Choice behavior
Evolutionary Time
Behavioral Output

A Road Map

Common Computational Principles

- Genetic pathways
- SNP markers
- CNVs

- Molecular and cellular pathways
- Gene products

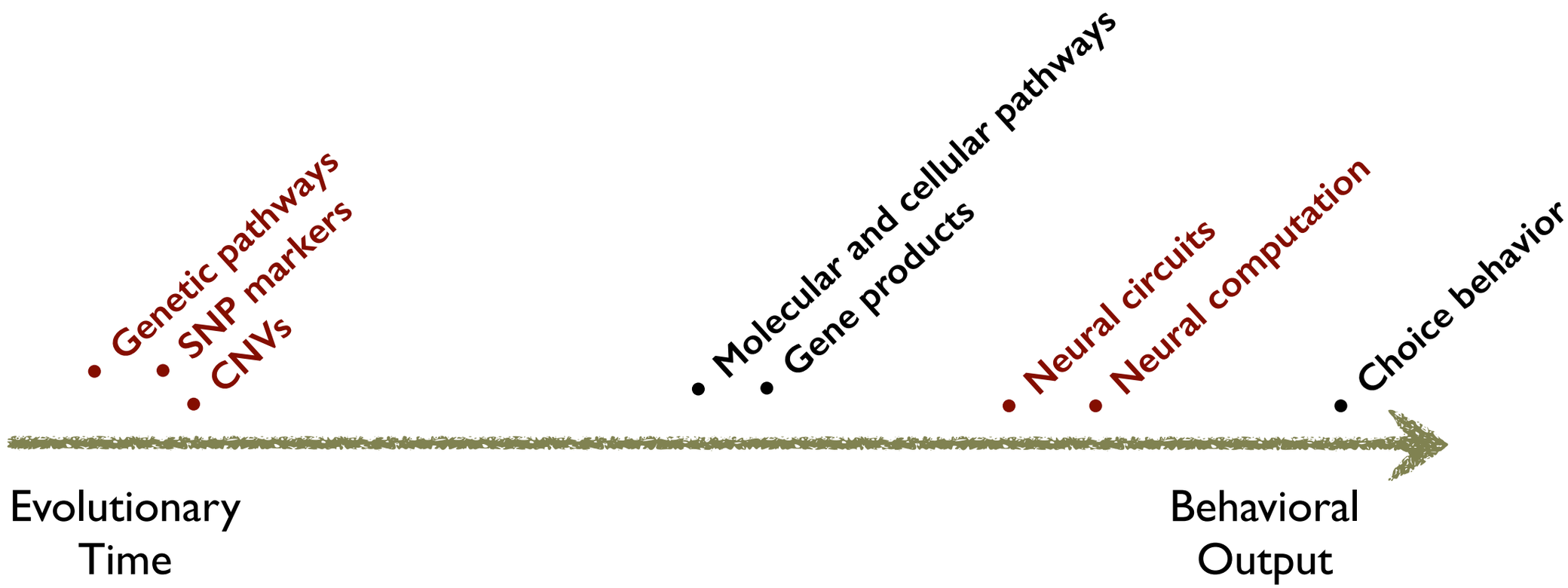
- Neural circuits
- Neural computation

- Choice behavior

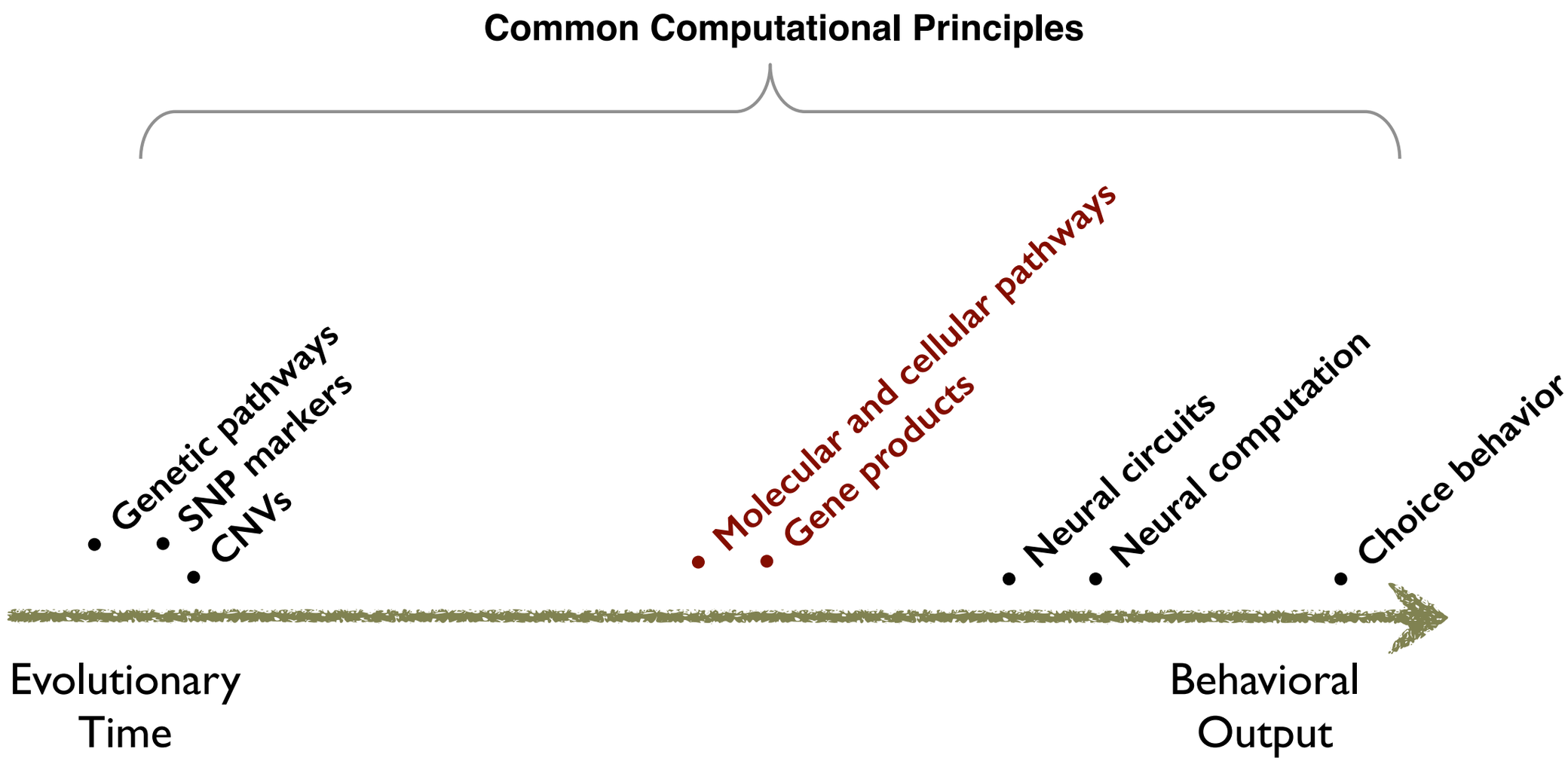
Evolutionary Time → Behavioral Output

A Road Map

Common Computational Principles



A Road Map: Tomorrow

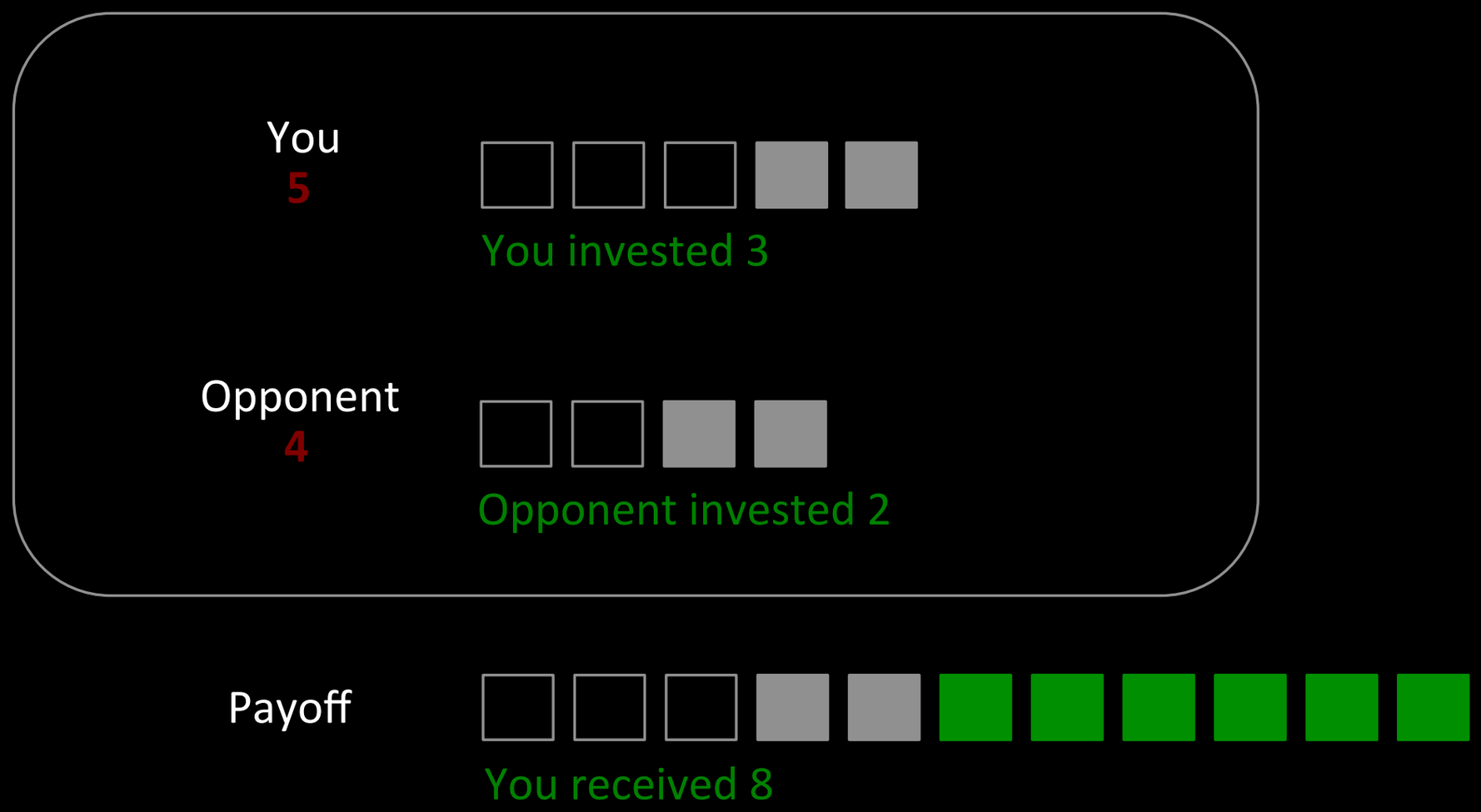


Game Theory

- **Language to <u>quantitatively</u> describe social interactions**


You
5
Opponent
4
Payoff


You
5
You invested 3
Opponent
4
Payoff


You
5
You invested 3
Opponent
4
Opponent invested 2
Payoff
You received 8

Game Theory

- Language to describe social interactions
- **Precise, <u>falsifiable</u> predictions**

Nash Equilibrium

Column Player (Weak)

Row Player (Strong)

	Invest 0		Invest 1		Invest 2		Invest 3		Invest 4		
Invest 0	5	4	5	13	5	12	5	11	5	10	0
Invest 1	14	4	4	3	4	12	4	11	4	10	0.2
Invest 2	13	4	13	3	3	2	3	11	3	10	0
Invest 3	12	4	12	3	12	2	2	1	2	10	0.2
Invest 4	11	4	11	3	11	2	11	1	1	0	0
Invest 5	10	4	10	3	10	2	10	1	10	0	0.6
	0.6		0		0.2		0		0.2		

Strategic Learning

Strategic Learning

- Reinforcement learning through trial and error

- Belief learning through anticipating and responding to action of others

Strategic Learning

$$V_i^k(t) = \begin{cases} \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t), \end{cases}$$

- Reinforcement learning through trial and error

- Belief learning through anticipating and responding to action of others

Strategic Learning

$$V_i^k(t) = \begin{cases} \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t), \end{cases}$$

- Reinforcement learning through trial and error

- Belief learning through anticipating and responding to action of others

Strategic Learning

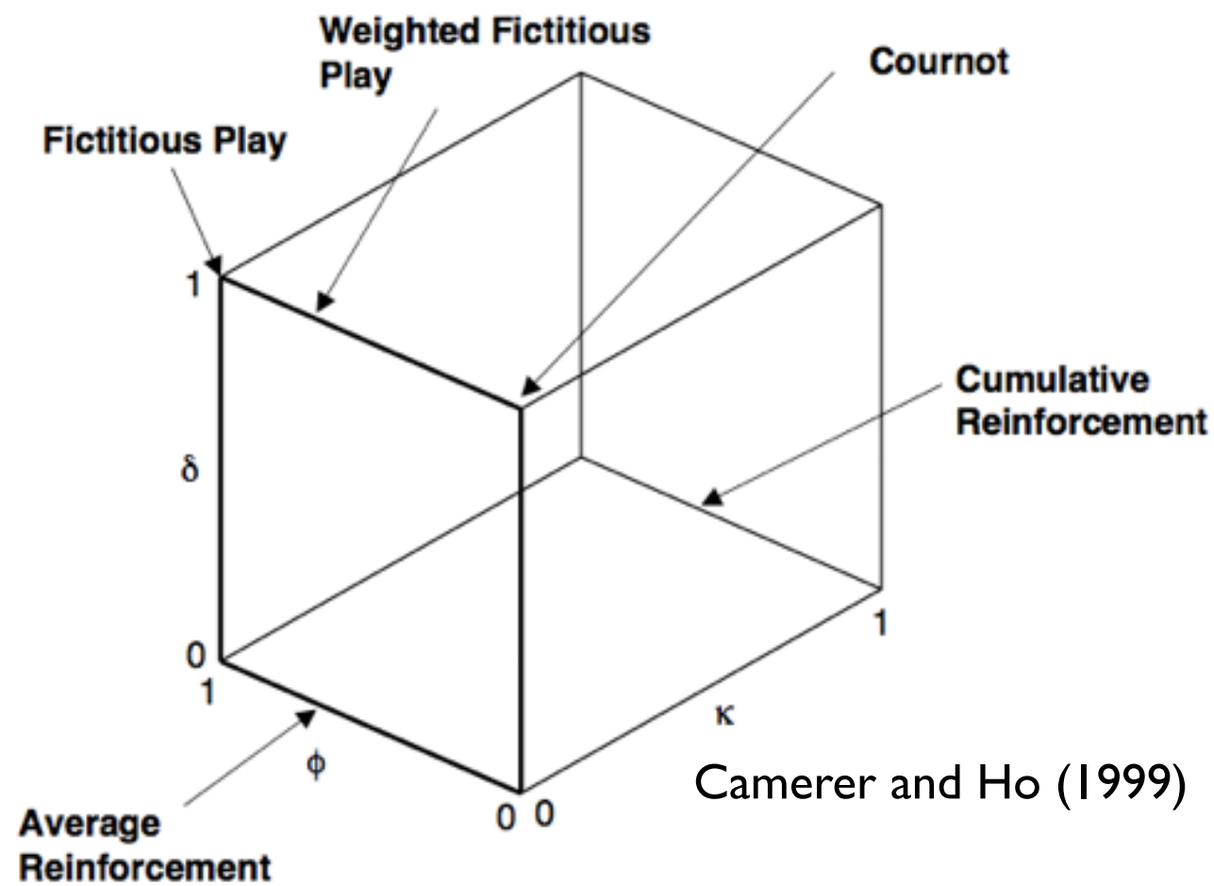
$$V_i^k(t) = \begin{cases} \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t), \end{cases}$$

- Reinforcement learning through trial and error
- Belief learning through anticipating and responding to action of others

Strategic Learning

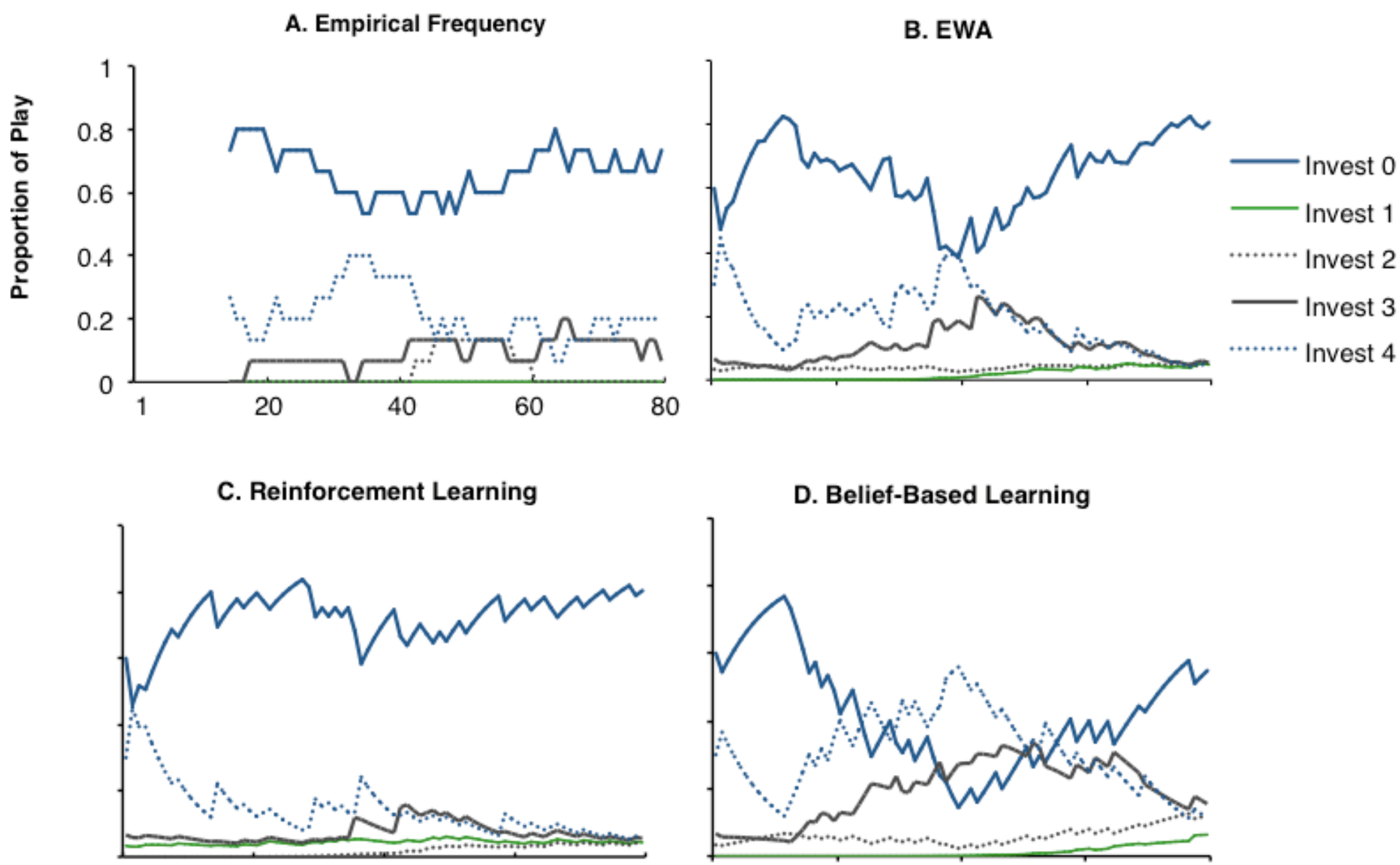
$$V_i^k(t) = \begin{cases} \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t), \end{cases}$$

- Reinforcement learning through trial and error
- Belief learning through anticipating and responding to action of others



Camerer and Ho (1999)

Computational Predictions



With Lusha Zhu

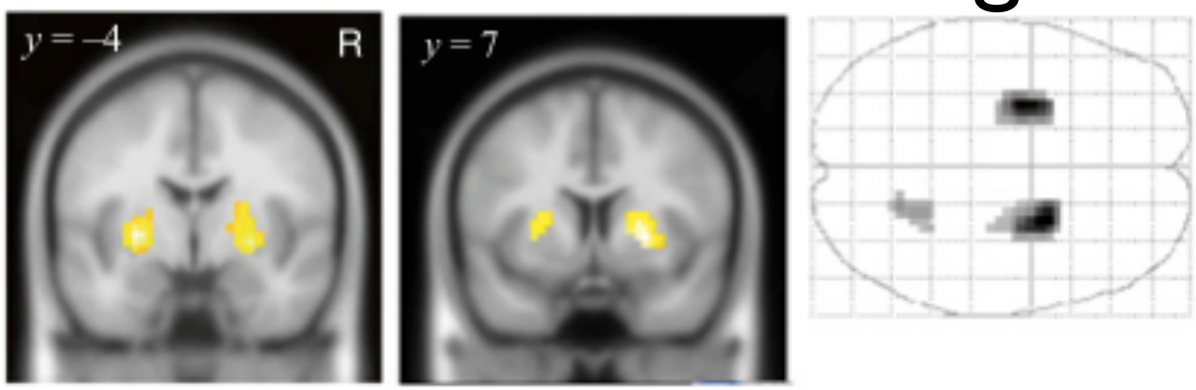
Reward Prediction Errors

Reinforcement Learning

Zhu L, Mathewson KE, Hsu M. *PNAS* (2012)

Reward Prediction Errors

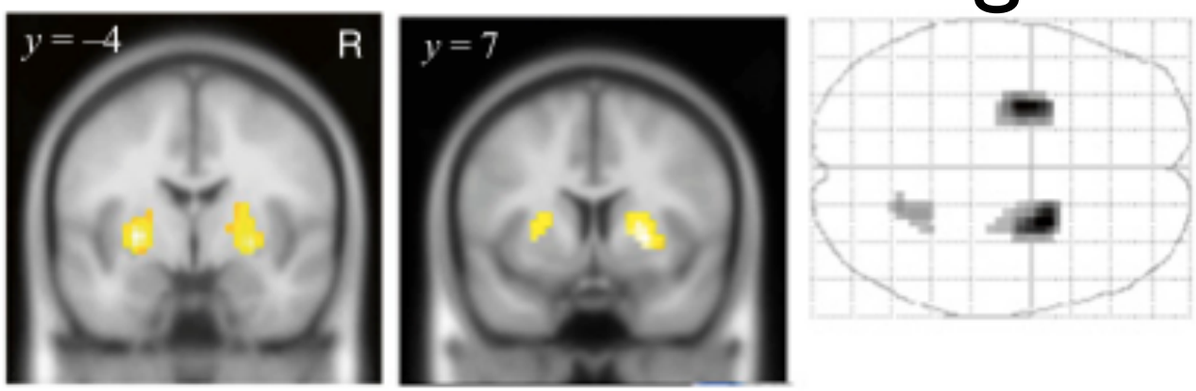
Reinforcement Learning



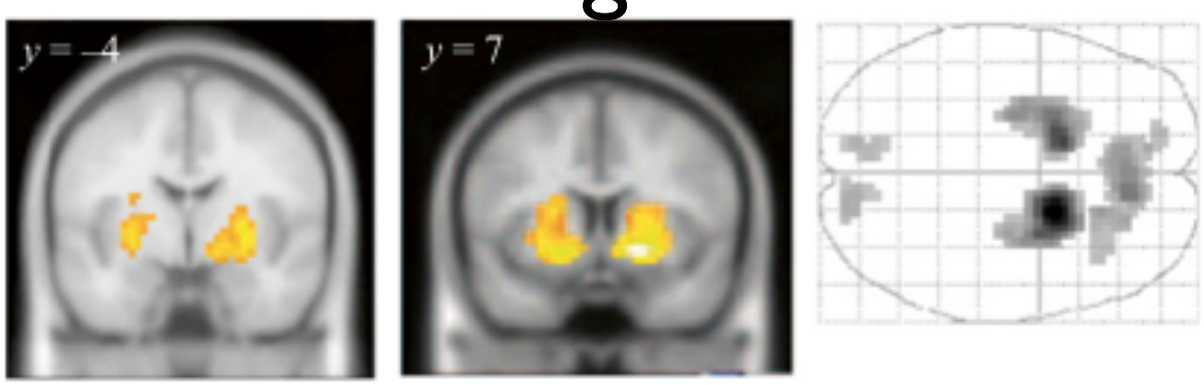
Zhu L, Mathewson KE, Hsu M. *PNAS* (2012)

Reward Prediction Errors

Reinforcement Learning



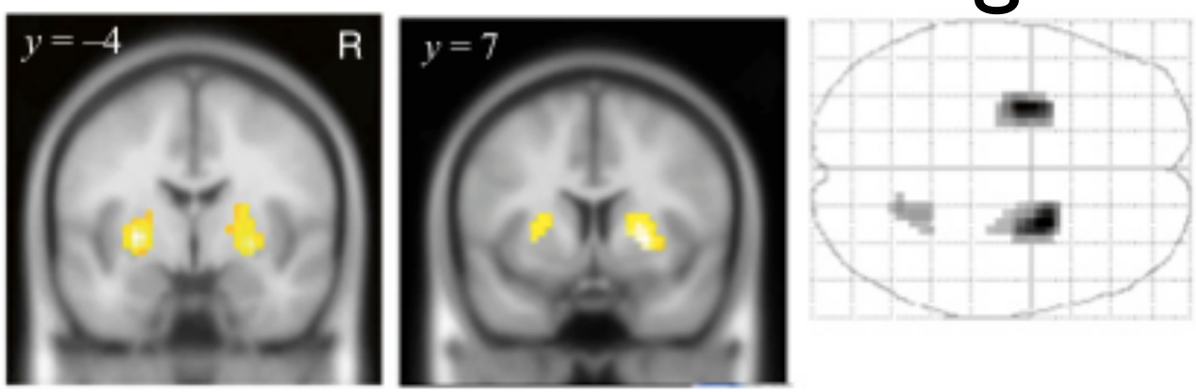
Belief Learning



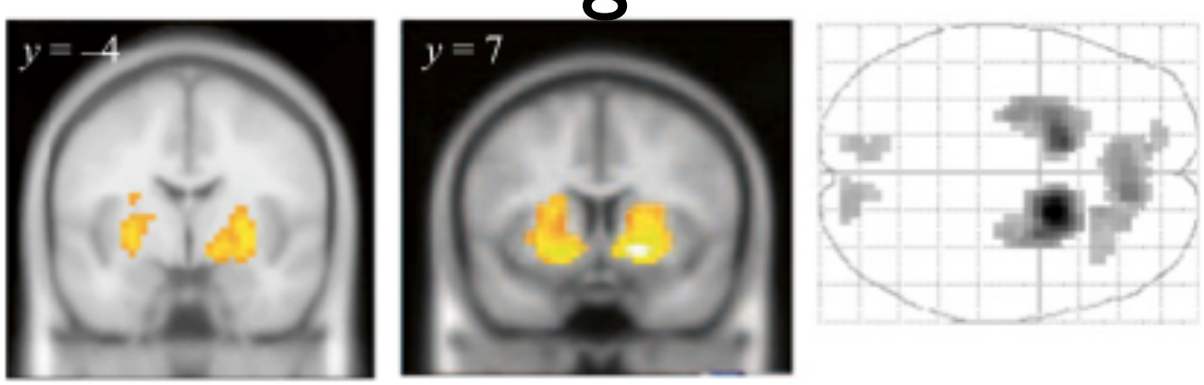
Zhu L, Mathewson KE, Hsu M. *PNAS* (2012)

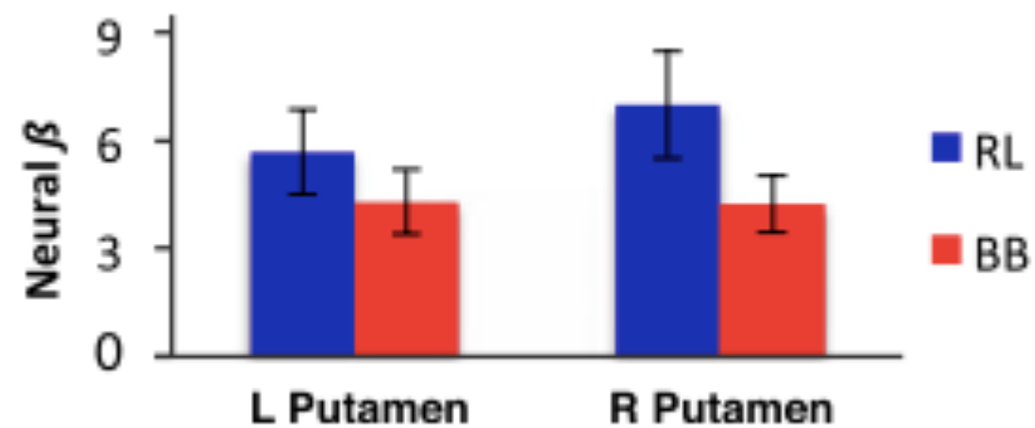
Reward Prediction Errors

Reinforcement Learning



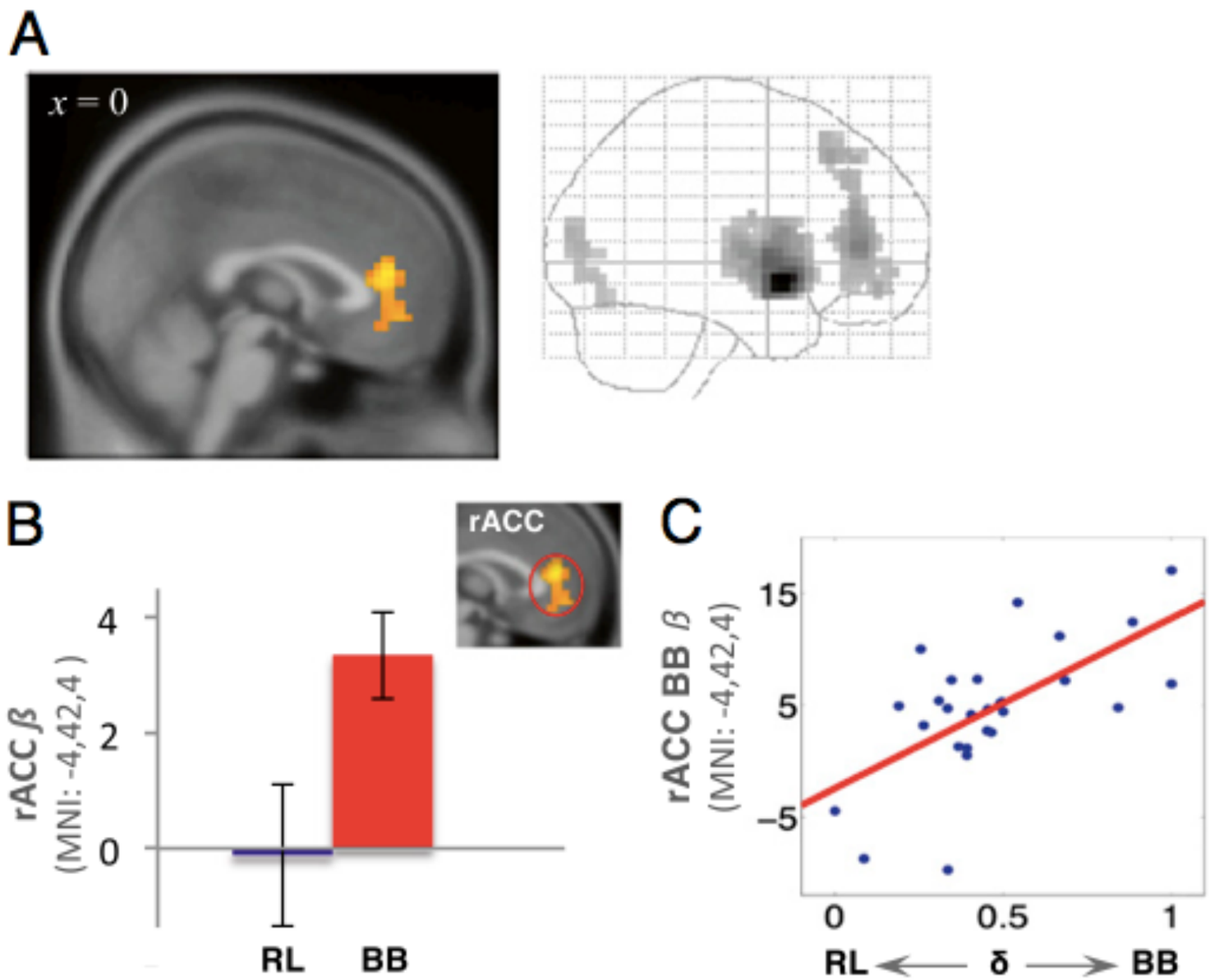
Belief Learning





Zhu L, Mathewson KE, Hsu M. *PNAS* (2012)

Belief Learning



Zhu L, Mathewson KE, Hsu M. *PNAS* (2012)

From Computation to Circuits and Genes

With Lusha Zhu, Bob Knight

With Eric Set, Ignacio Saez, and B2ESS

From Computation to Circuits and Genes

Focal Brain Lesion Patients

With Lusha Zhu, Bob Knight

With Eric Set, Ignacio Saez, and B2ESS

From Computation to Circuits and Genes

Focal Brain Lesion Patients

Dopaminergic Pathway Genes

With Lusha Zhu, Bob Knight

With Eric Set, Ignacio Saez, and B2ESS

From Computation to Circuits and Genes

Focal Brain Lesion Patients

Dopaminergic Pathway Genes

- Causal Mechanisms
- Frontostriatal circuits

With Lusha Zhu, Bob Knight

With Eric Set, Ignacio Saez, and B2ESS

From Computation to Circuits and Genes

Focal Brain Lesion Patients

- Causal Mechanisms
- Frontostriatal circuits

With Lusha Zhu, Bob Knight

Dopaminergic Pathway Genes

- Differential expression
- Functional differences

With Eric Set, Ignacio Saez, and B2ESS

A Puzzle

A Puzzle

- MPFC (OFC) damage:

A Puzzle

- MPFC (OFC) damage:
  - “striking defects in social emotion but generally intact intellect and normal baseline mood”

A Puzzle

- MPFC (OFC) damage:
  - “striking defects in social emotion but generally intact intellect and normal baseline mood”
- BG damage:

A Puzzle

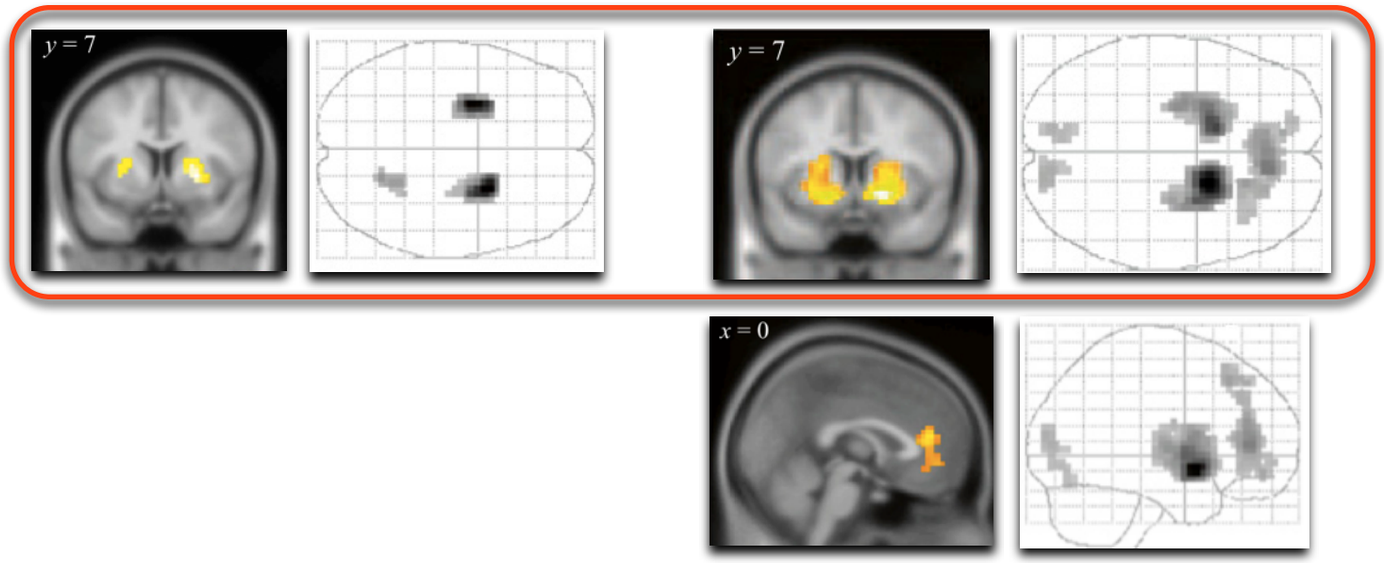
- MPFC (OFC) damage:
  - "striking defects in social emotion but generally intact intellect and normal baseline mood"
- BG damage:
  - " We did not encounter a clinical description of the puerile, profane, slovenly, facetious, irresponsible, grandiose and irascible individual following basal ganglia damage. Nor was there clear description of erosion of foresight, judgement and insight, loss of ability to delay gratification and of the capacity for remorse…" Bhatia et al., Brain, 1994

Hypotheses:

Strategic Treatment

Reinforcement Learning

Belief-Based Learning



Hypothesis 1:

BG is necessary for reinforcement and belief-based learning

BG lesion impaired in both learning

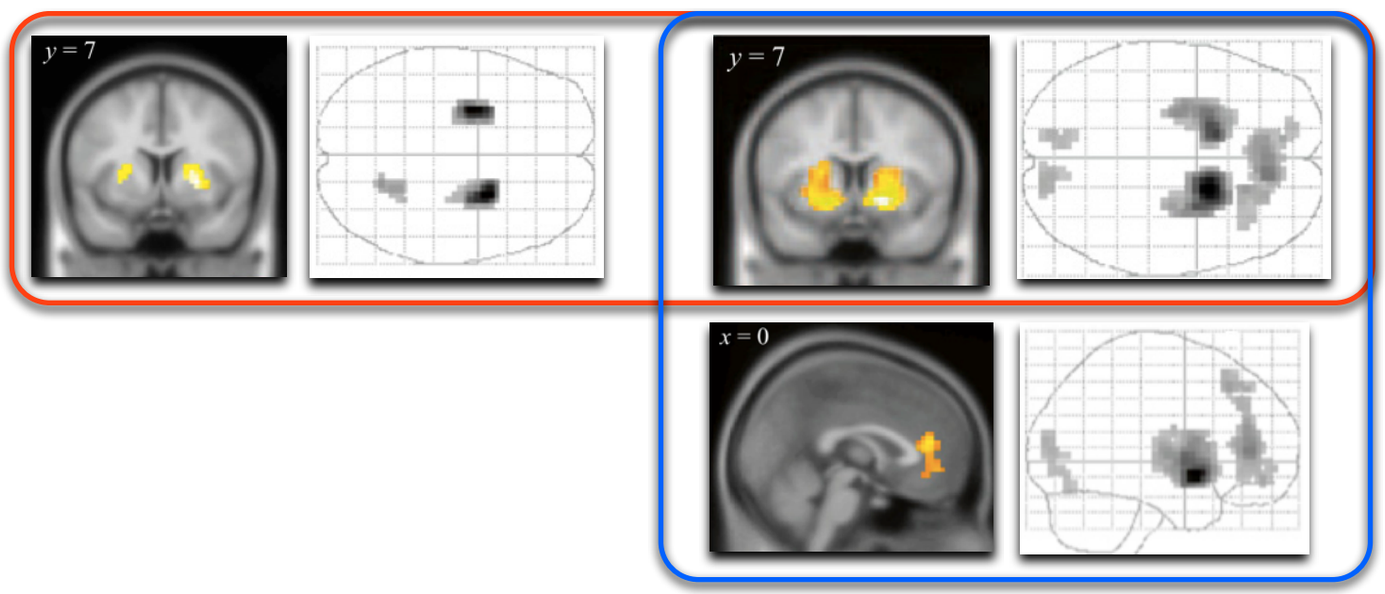
MPFC lesion is only impaired in strategic learning, but intact in non-strategic learning

Previous Behavioral And Neural Findings

Strategic Treatment

Reinforcement Learning

Belief-Based Learning

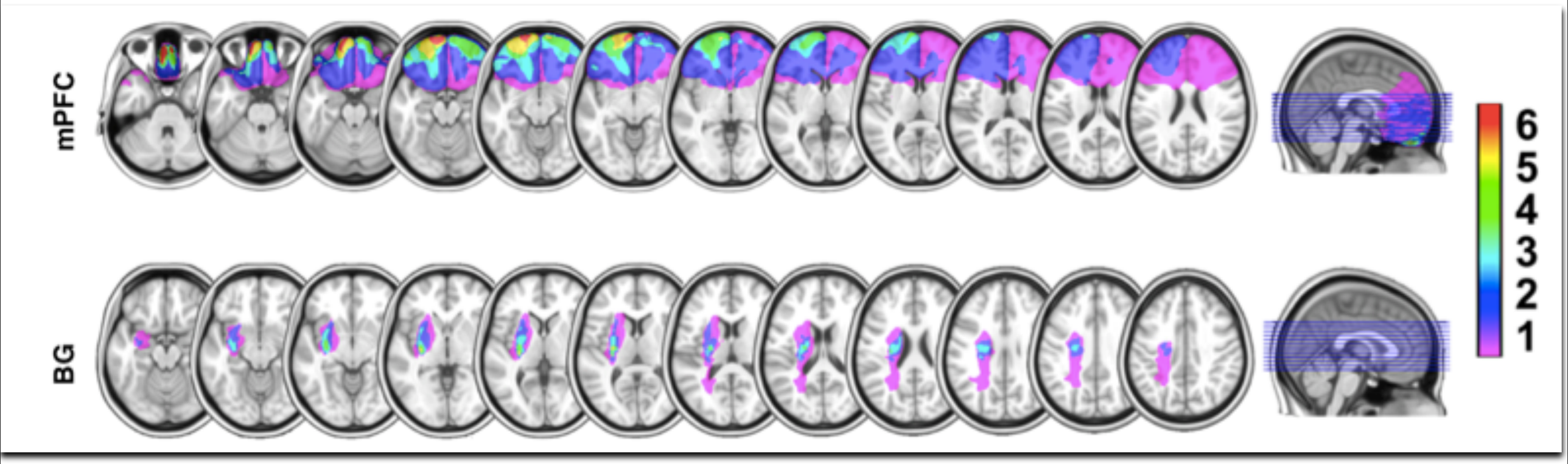


Hypothesis 2:

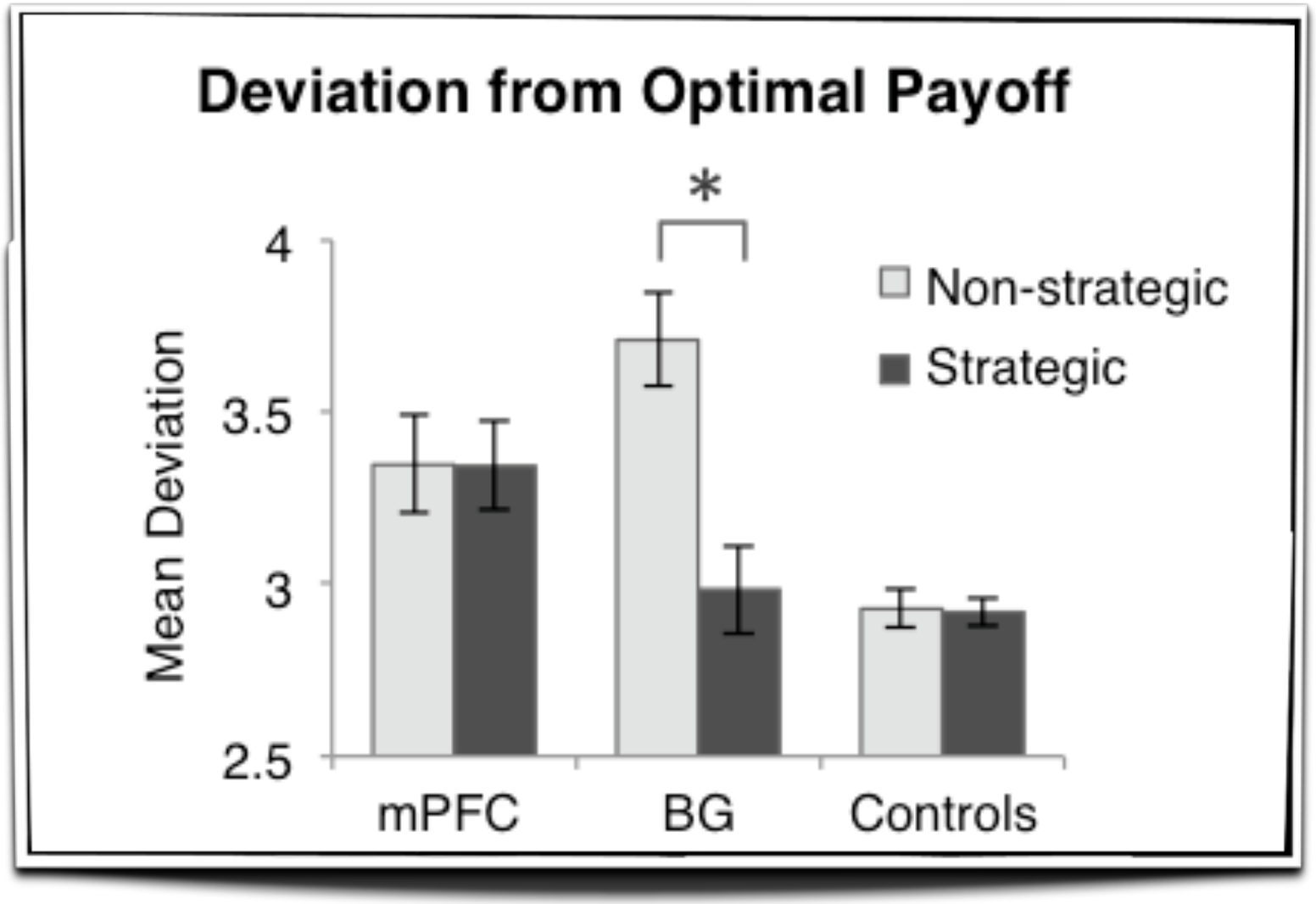
BG is only necessary for reinforcement and MPFC is necessary for belief-based learning.

BG lesion leads to impairment in non-strategic learning but not to strategic learning.

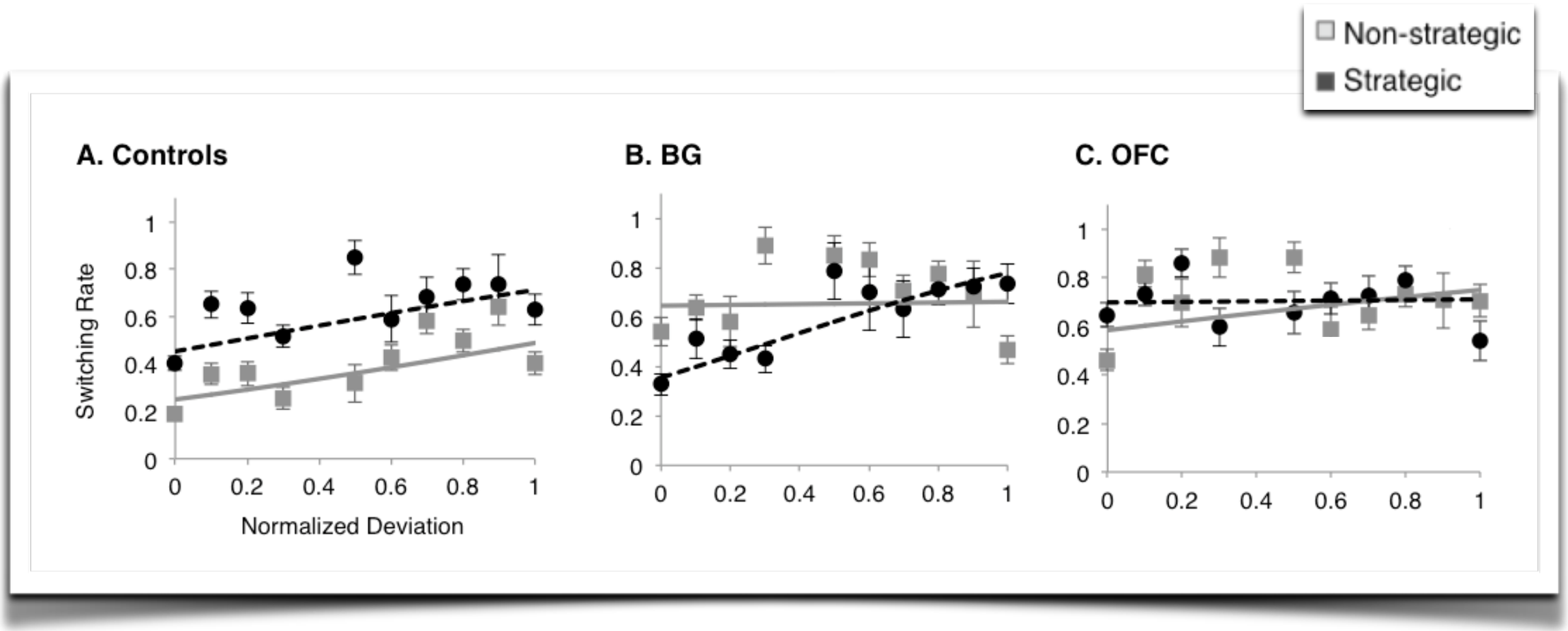
Subjects



Overall performance



Payoff Deviation and Strategy Switching



Hypothesis

$$V_i^k(t) = \begin{cases} \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t), \end{cases}$$

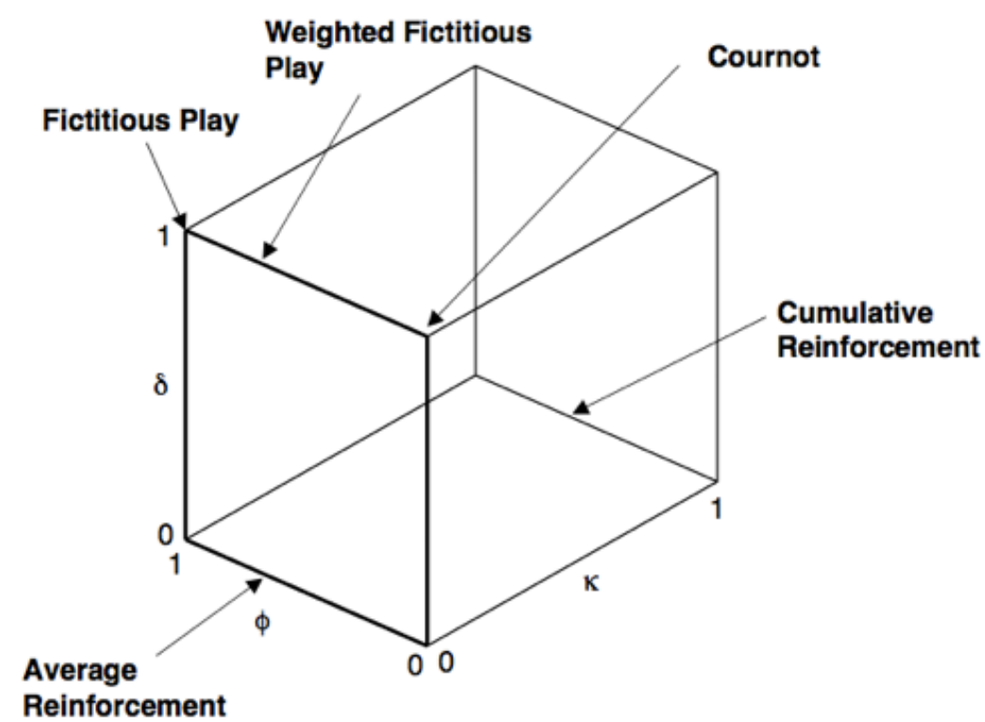
Preservation of strategic learning capacity in BG patients

↑

BG pts invoke compensatory processes that mediate belief-based learning

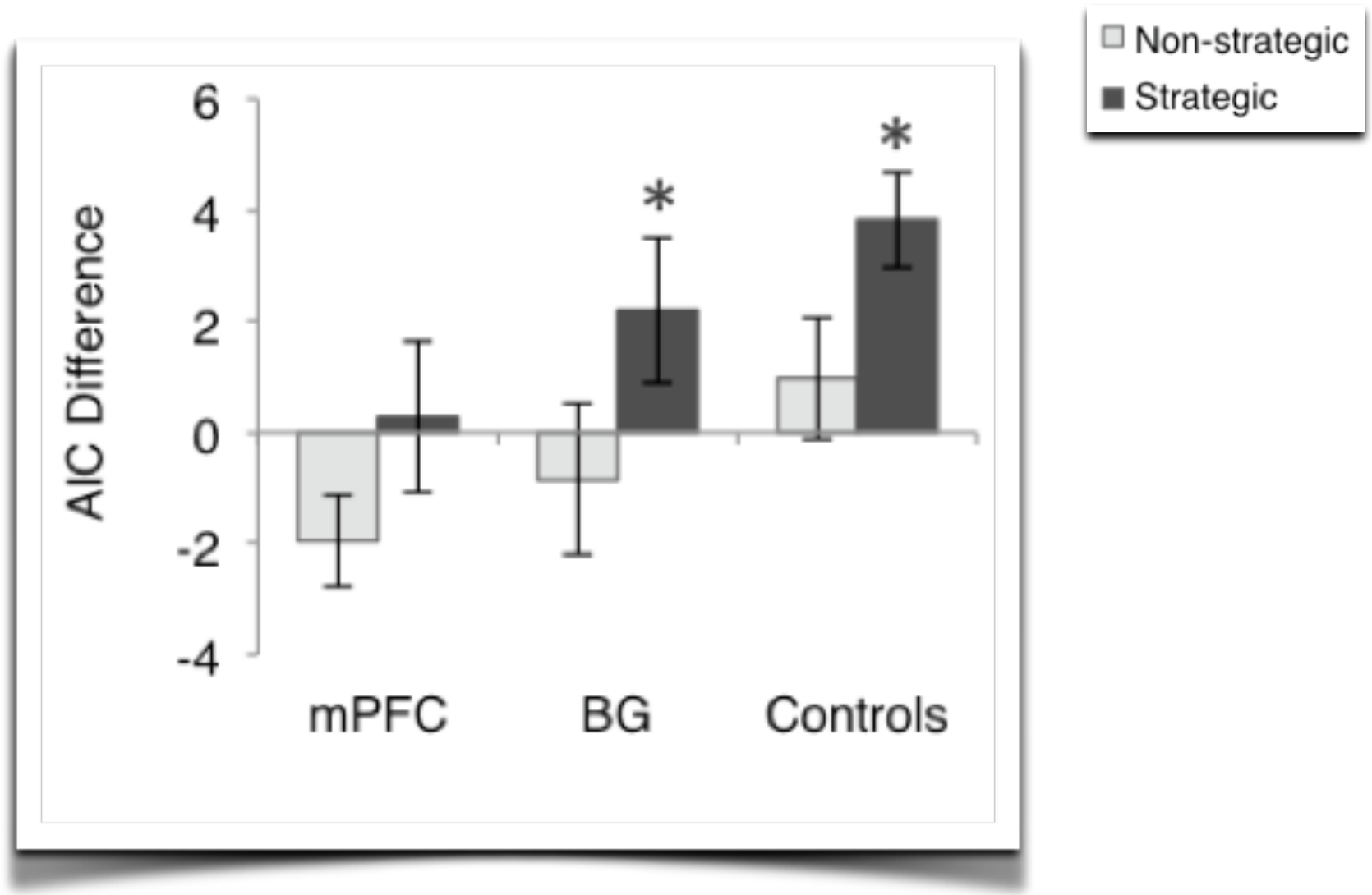
↑

Reinforcement learning vs. Hybrid learning:



Camerer and Ho (1999)

Fit Improvement Over Reinforcement Learning



Summary

- mPFC lesion impaired in both strategic and non-strategic learning.
- BG lesion intact in strategic learning despite deficit in non-strategic learning.
- BG patients may **compensate by invoking higher-order learning in social settings**.

Brain to Genes

- Collaboration with B2ESS Lab in Singapore
- Behavior in Patent Race from 218 subjects with whole genome sequencing data

With Eric Set, Ignacio Saez, and B2ESS

Empirical Approach

- Hypothesis:
  - Variation in DA genes affect key learning parameters

Empirical Approach

- Hypothesis:
  - Variation in DA genes affect key learning parameters

$$V_i^k(t) = \begin{cases} \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t), \end{cases}$$

Empirical Approach

- Hypothesis:
  - Variation in DA genes affect key learning parameters

$$V_i^k(t) = \begin{cases} \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t), \end{cases}$$

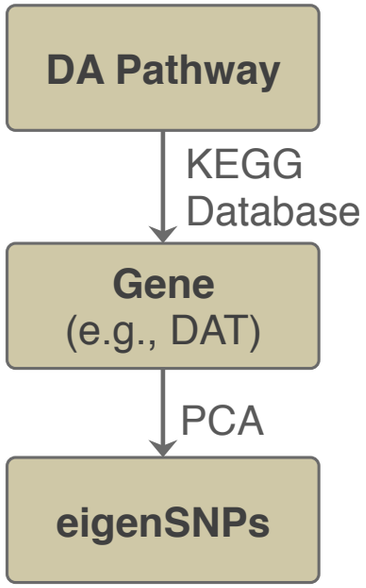
Empirical Approach

- Hypothesis:
  - Variation in DA genes affect key learning parameters

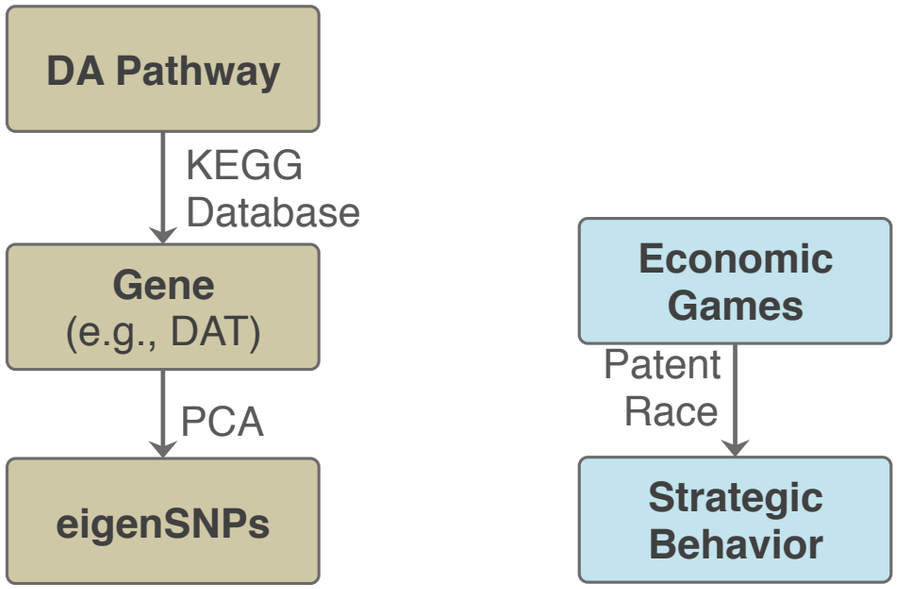
$$V_i^k(t) = \begin{cases} \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k = s_i(t) \\ \dfrac{\phi_i \cdot N(t-1) \cdot V_i^k(t-1) + \delta_i \cdot \pi_i(s_i^k, s_{-i}(t))}{N(t)}, & \text{if } s_i^k \neq s_i(t), \end{cases}$$

Empirical Approach

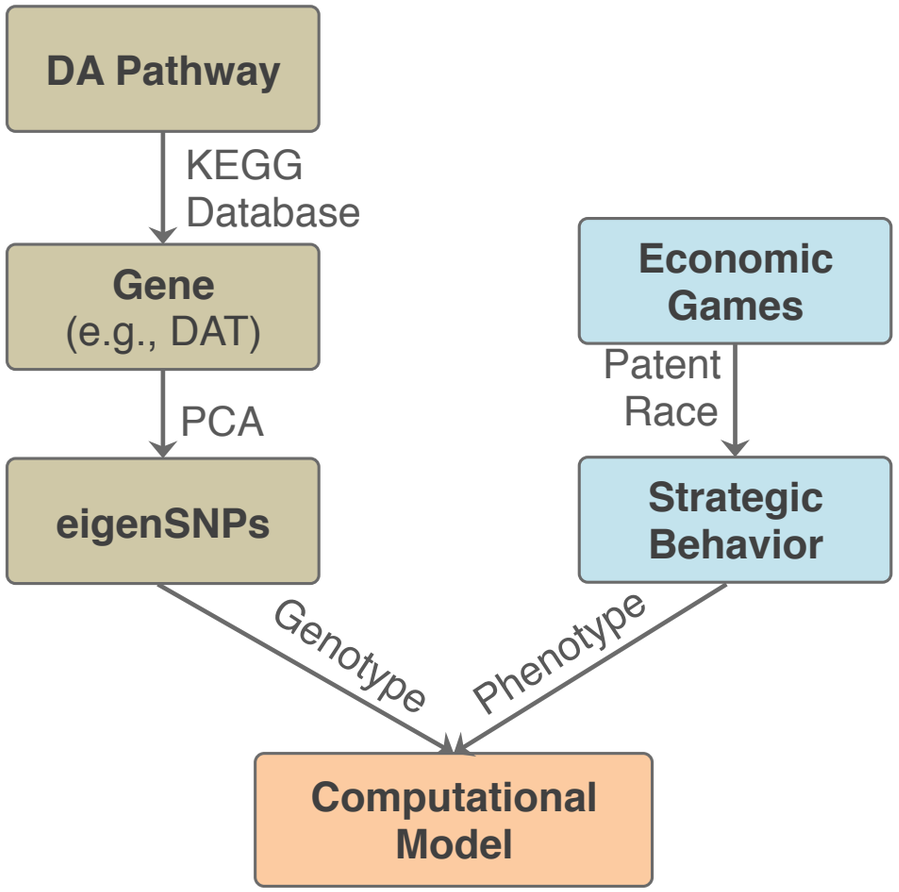
Empirical Approach



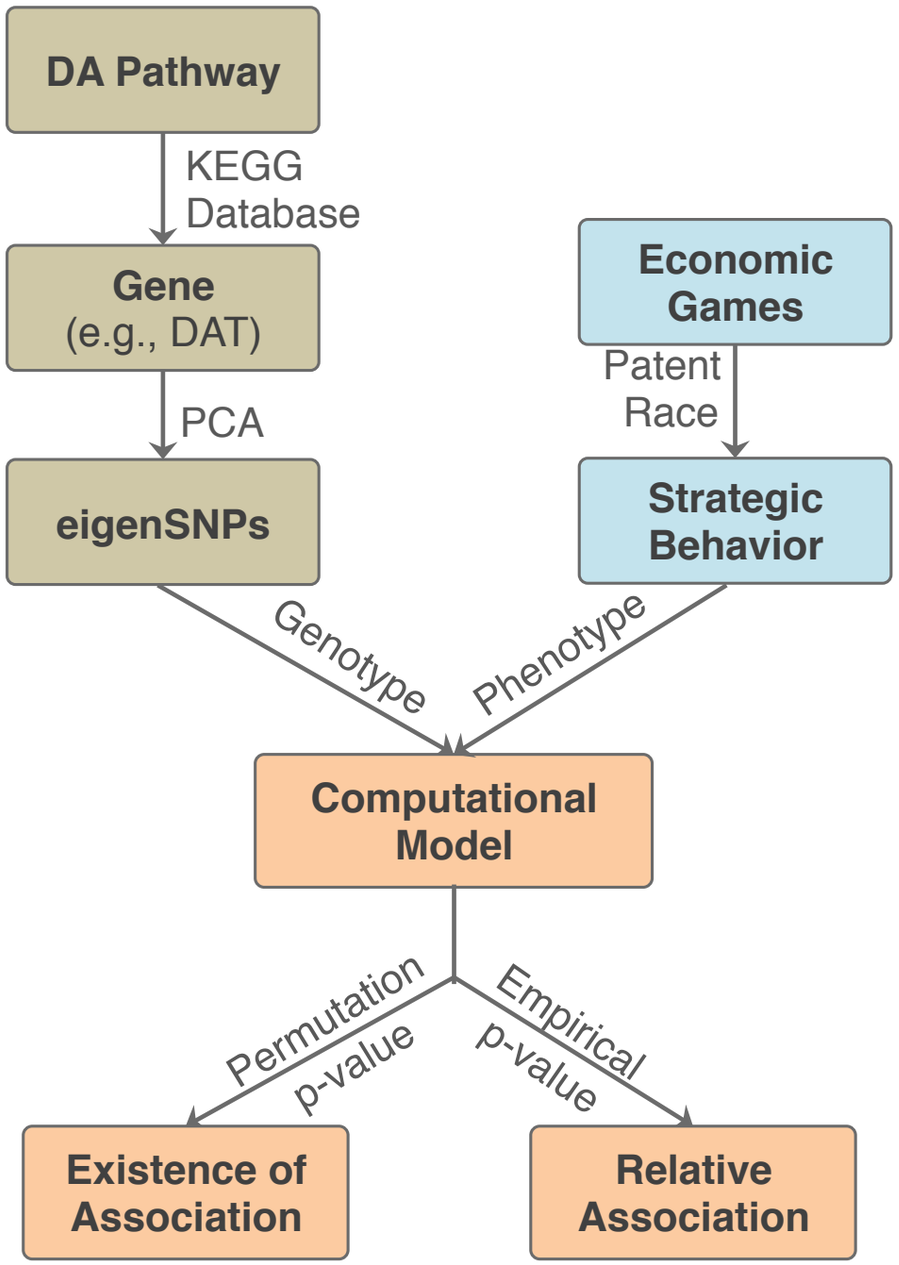
Empirical Approach



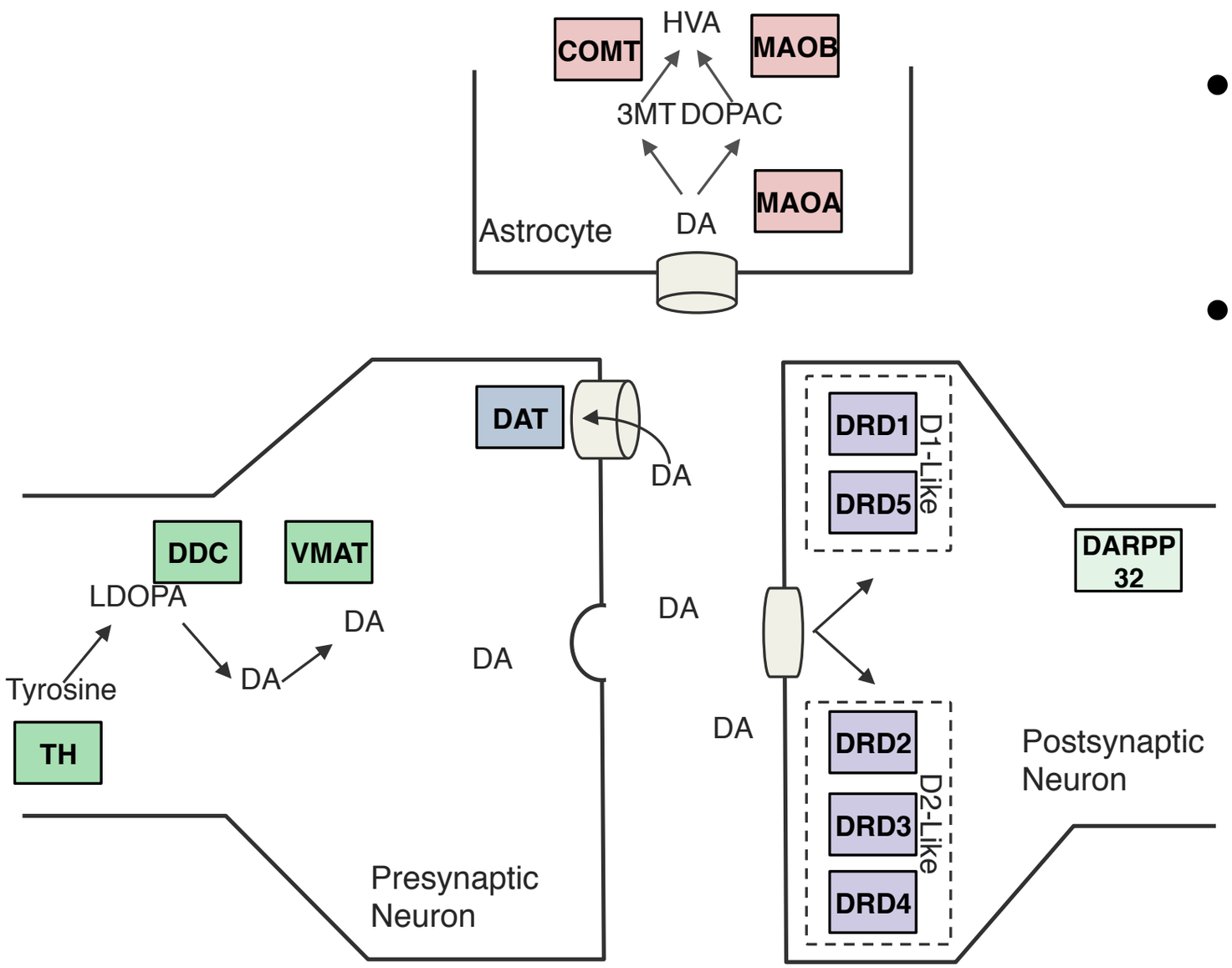
Empirical Approach



Empirical Approach

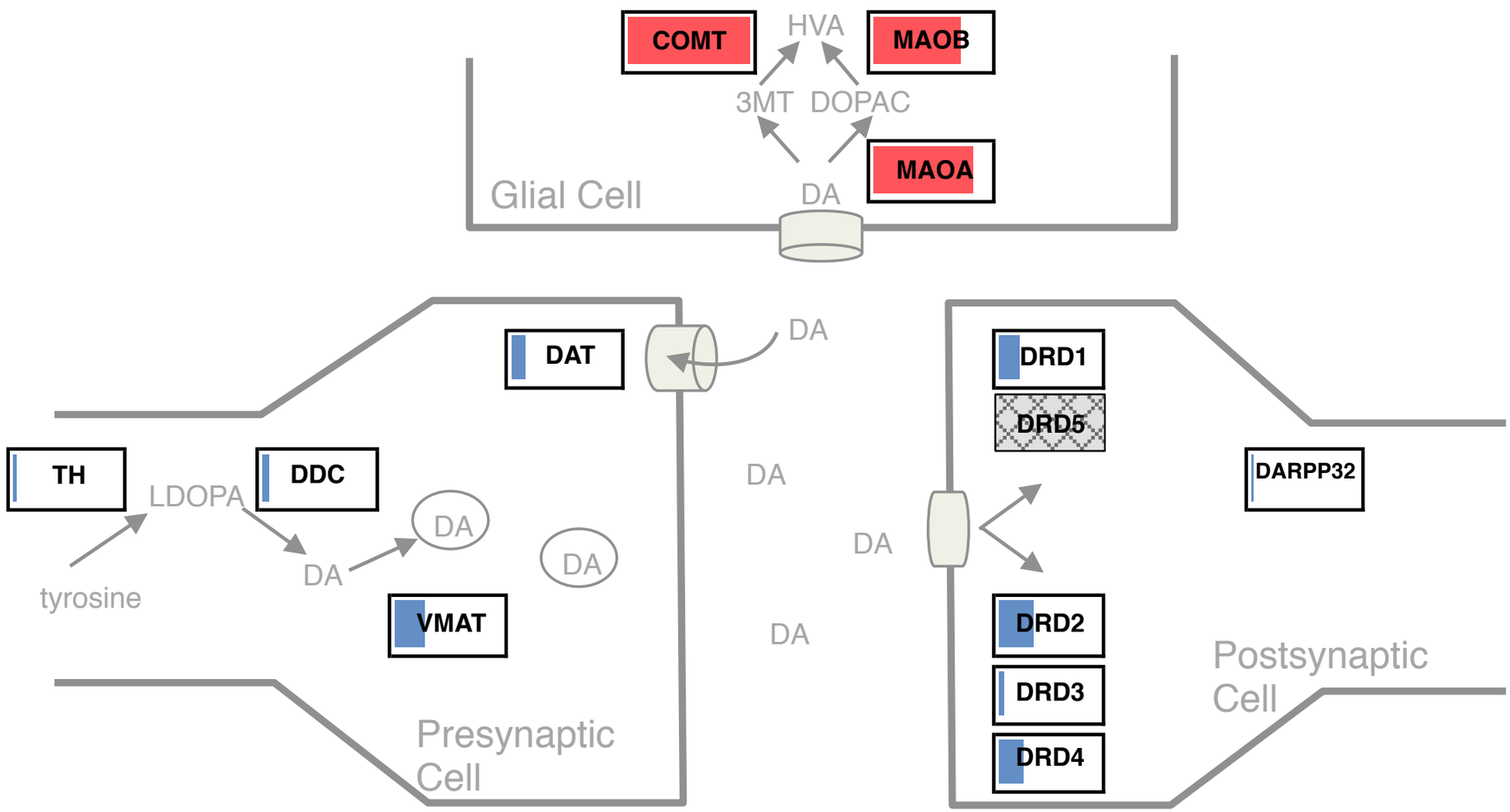


DA Pathway

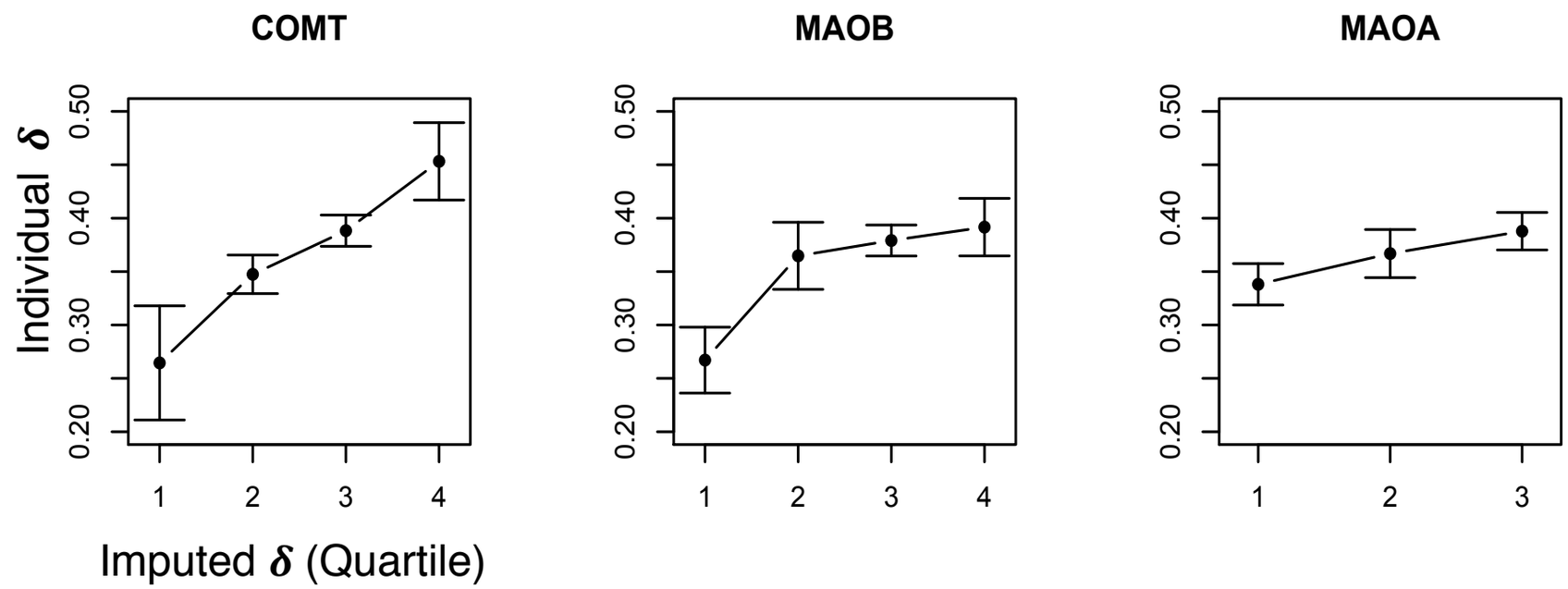


- **Striatal**:
  - DAT1, DRD2, DRD3
- **Prefrontal**:
  - COMT, MAOA, MAOB

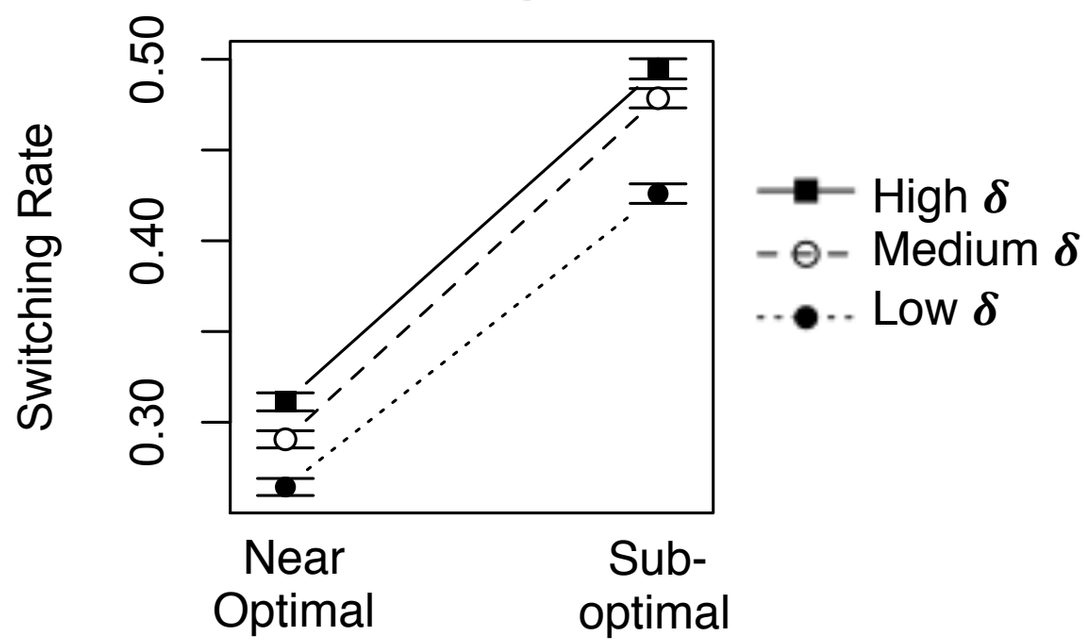
Belief Learning



Imputed Belief Learning Parameter

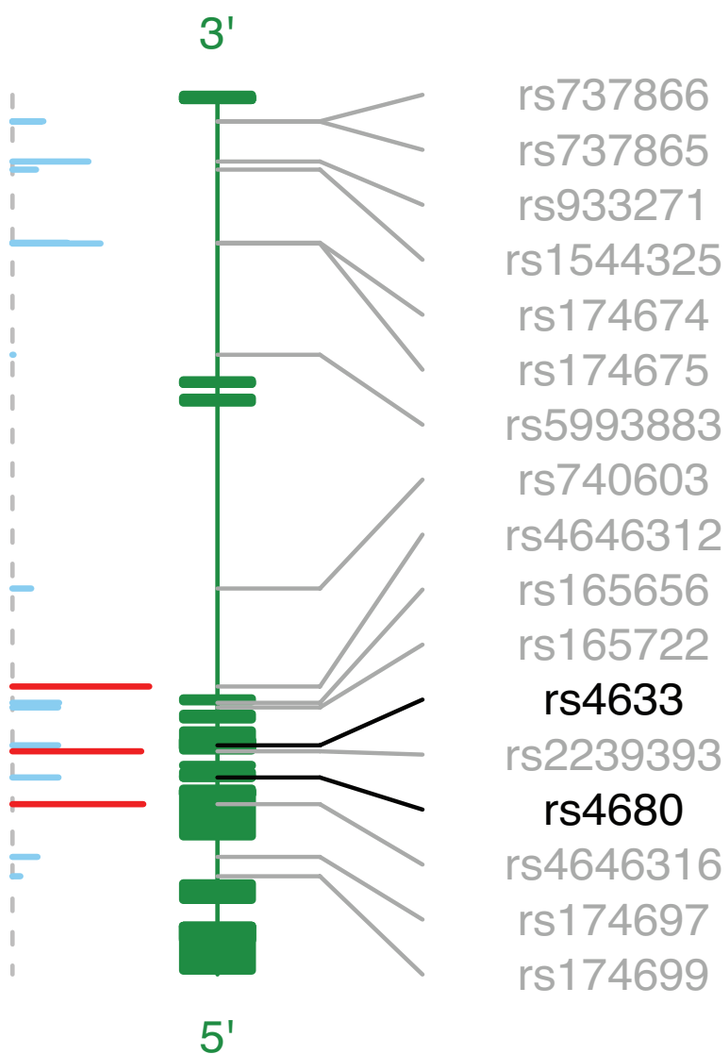


Effect on Switching Rate



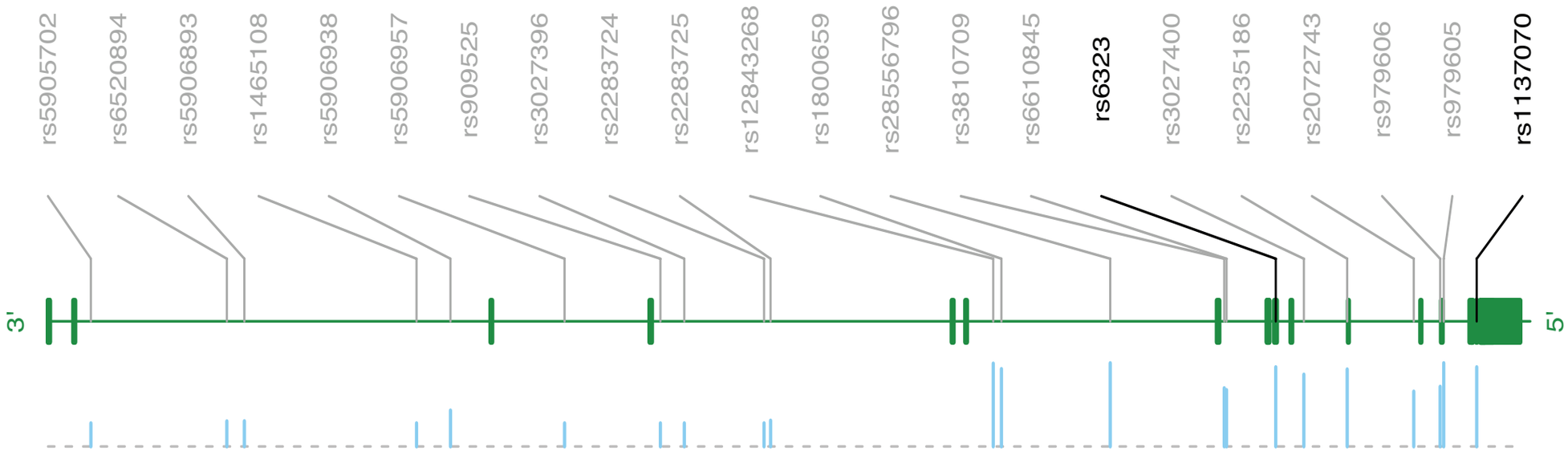
SNP-level Effects

COMT (22q11.21–28.4Kbp)

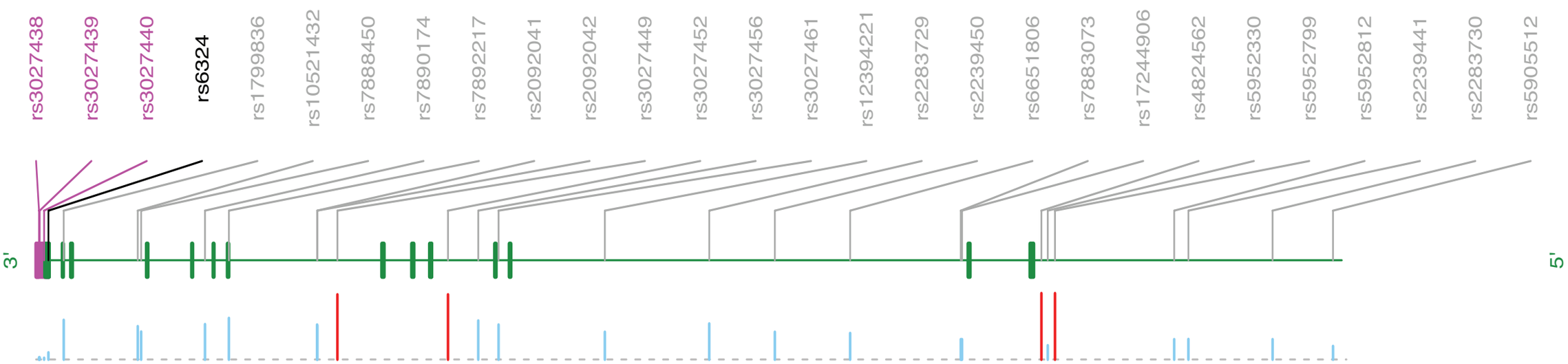


SNP-level Effects

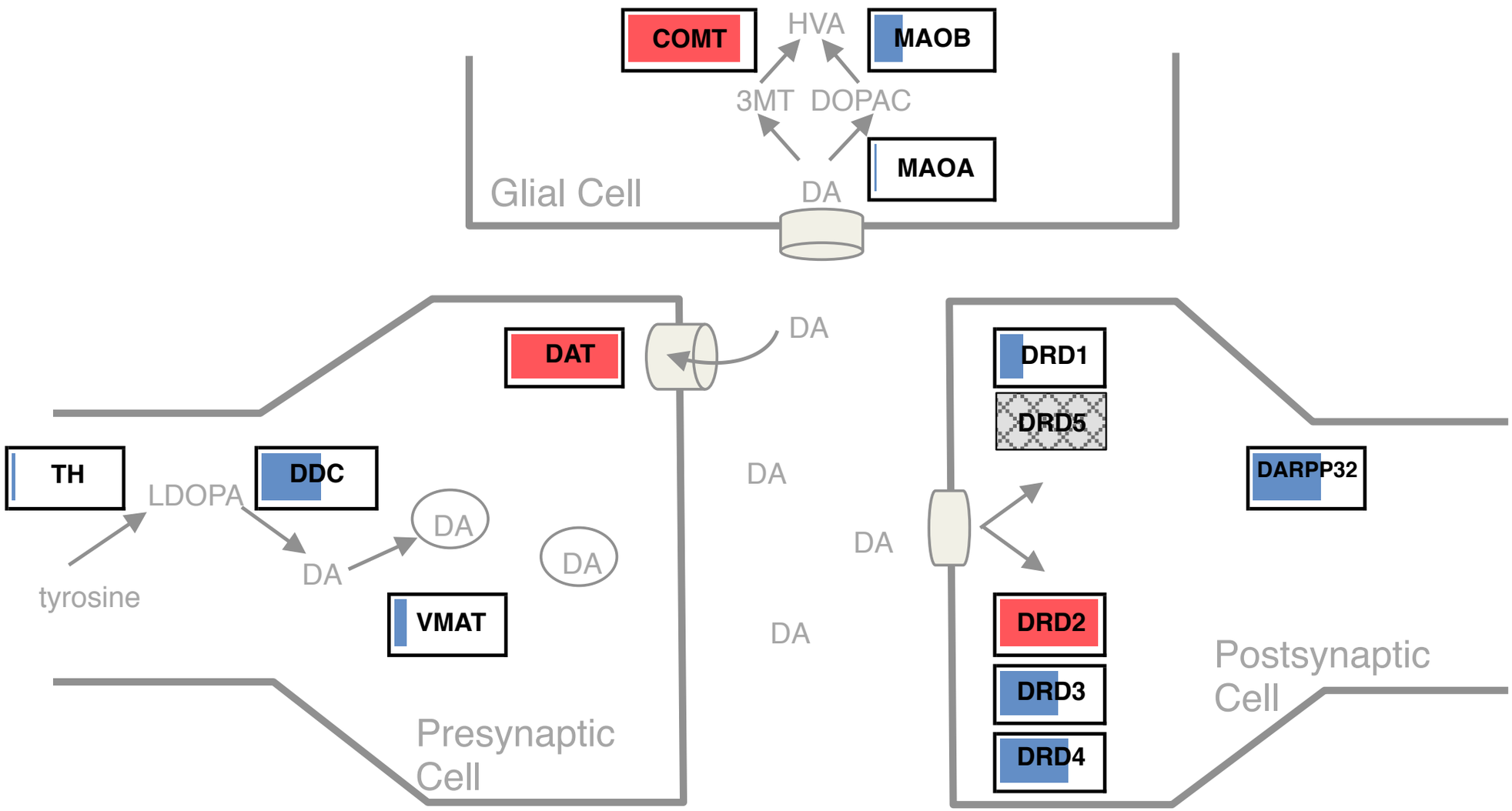
MAOA (Xp11.3–90.6Kbp)



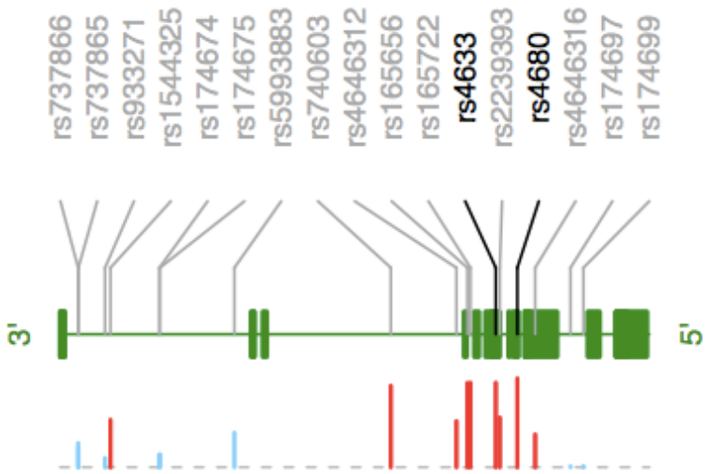
MAOB (Xp11.23–115.9Kbp)



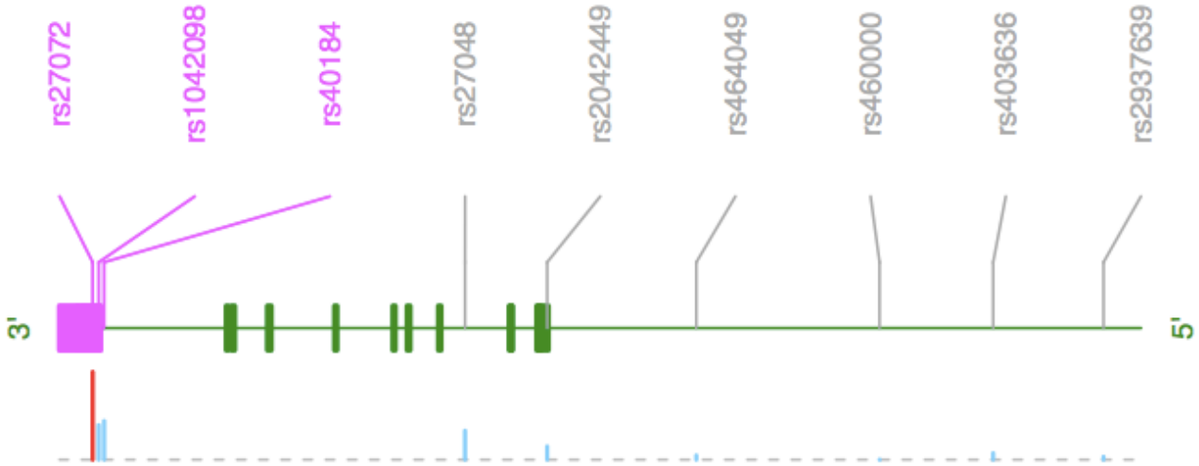
Learning Rate



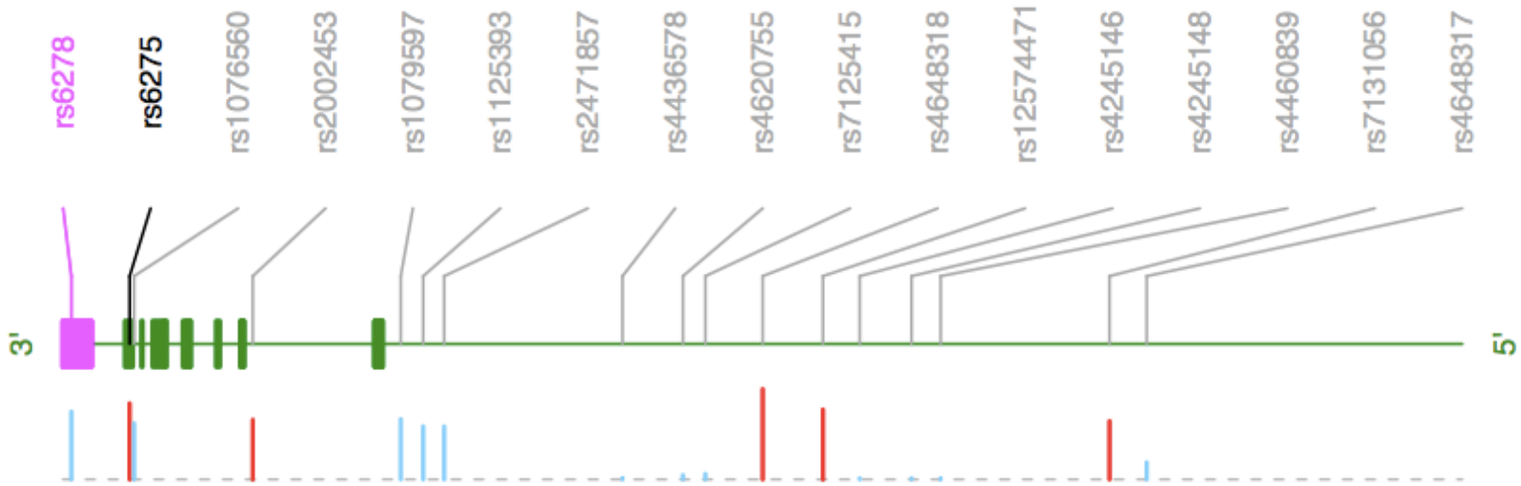
SNP-level Effect



COMT – 22q11.21 – 28.4Kbp

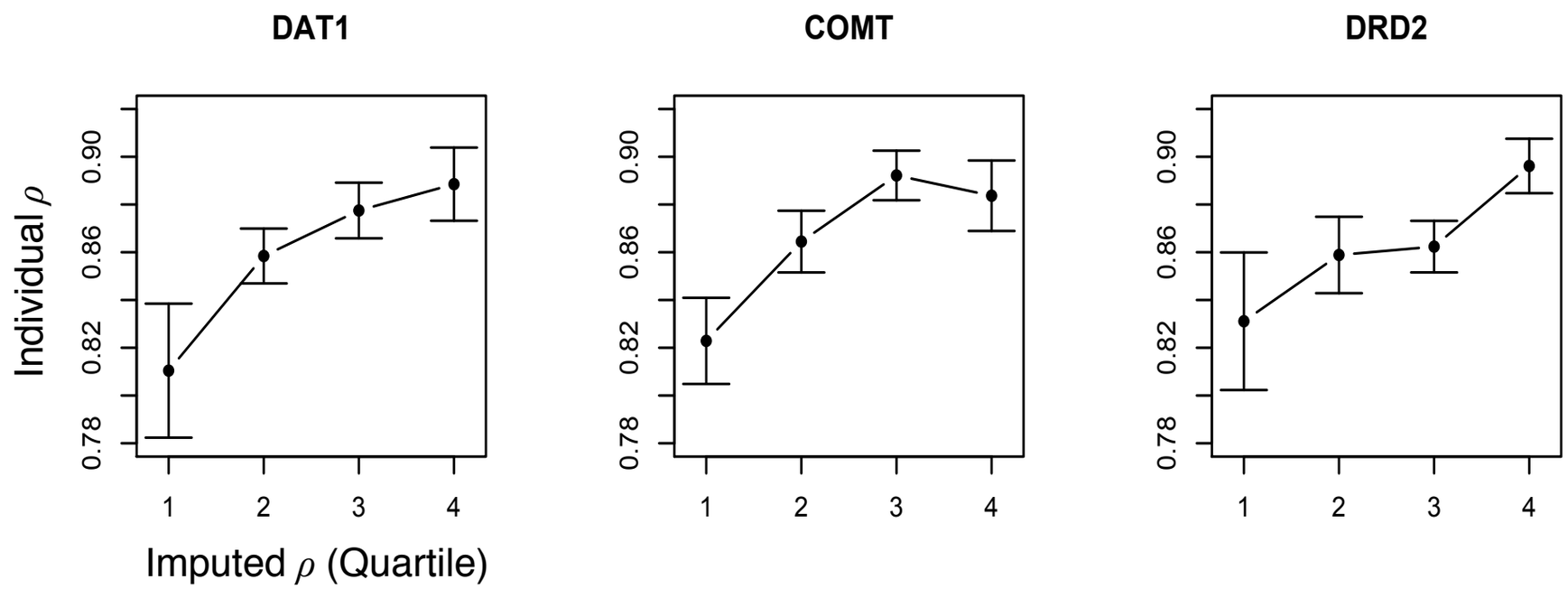


DAT1 – 5p15.33 – 52.6Kbp

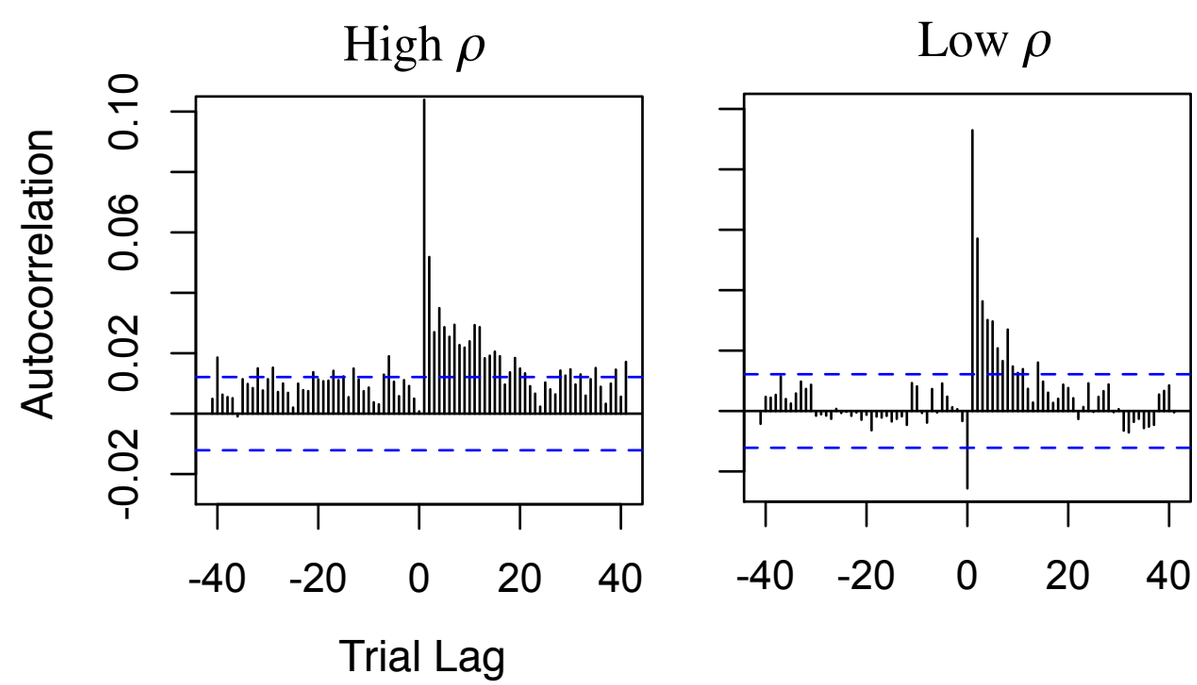


DRD2 – 11q23 – 66.1Kbp

Imputed Discounting Parameter



Effect of Learning Rate on Choice

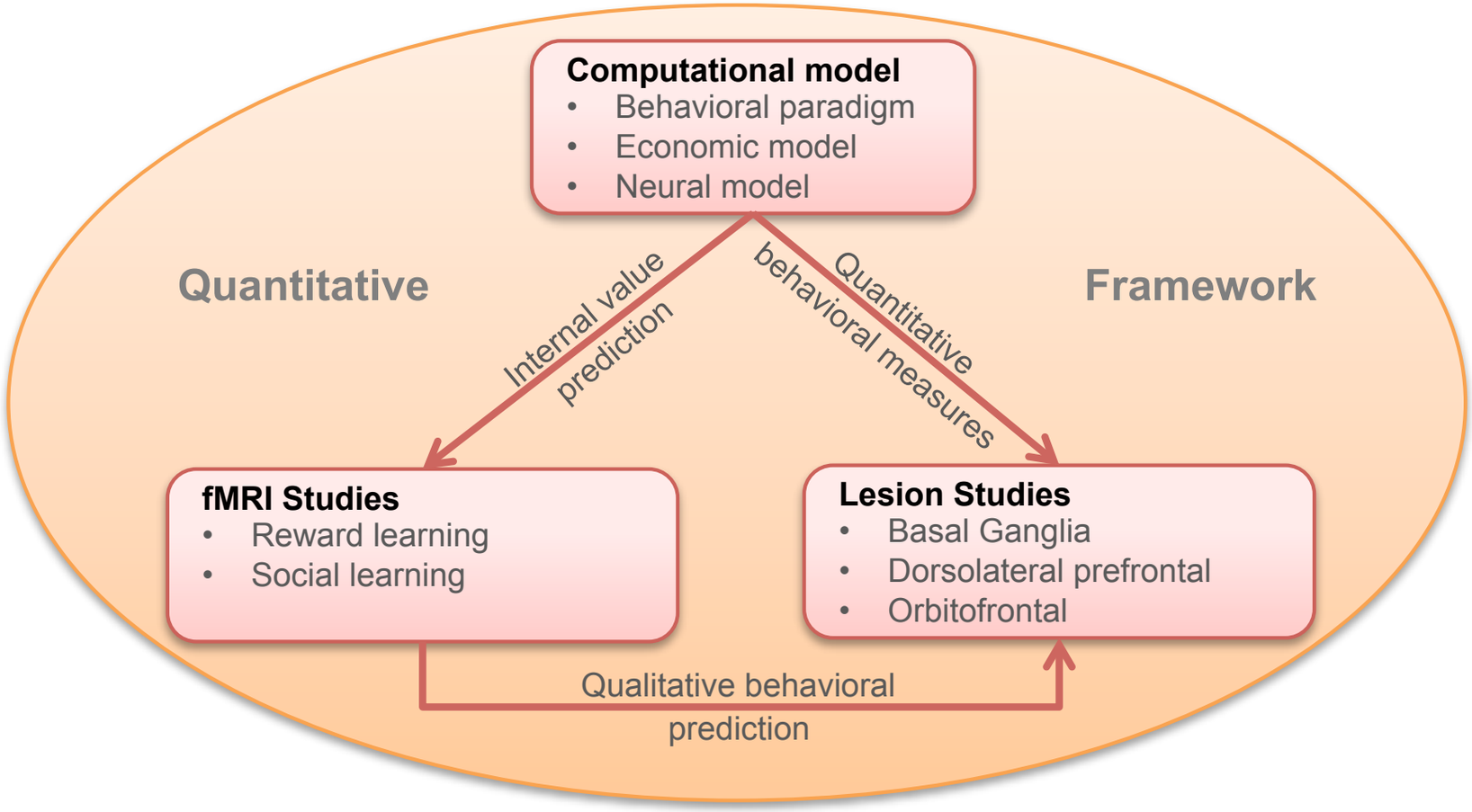


						Belief Learning (δ)				Discounting (ρ)			
Function	**Gene**	**SNPs**	**PCs**	**% Var**	k[1]	**LLR**	p_{unc}	p_{perm}	p_{emp}	**LLR**	p_{unc}	p_{perm}	p_{emp}
Synthesis	TH	2	2	100%	1089	0.98	0.374	0.898	0.888	0.9	0.400	0.913	0.932
	DDC	20	4	90%	162	3.45	0.141	0.943	0.963	29.3	0.000	0.257	0.278
	VMAT	16	8	92%	22	31.1	0.000	0.420	0.410	12.3	0.002	0.969	1.000
Transport/ Clearance	DAT1	9	5	93%	73	9.35	0.002	0.821	0.808	68.6	0.000	0.024	0.027
	VNTR[3]					0.22	0.510	0.796		34.5	0.000	0.008	
	Joint					9.7	0.007	0.877		86.9	0.000	0.014	
	COMT	17	4	91%	191	57.3	0.000	0.005	0.005	49.8	0.000	0.038	0.031
	MAOA	22	1	94%	4	12.3	0.000	0.082	0.25	0.2	0.495	0.834	1.000
	VNTR[4]					0.7	0.247	0.687		2.2	0.034	0.498	
	Joint					27.7	0.000	0.029		4.1	0.017	0.691	
	MAOB	28	3	95%	70	32.7	0.000	0.035	0.029	10.2	0.000	0.585	0.586
Receptor	DRD1	5	3	99%	275	9.22	0.000	0.522	0.510	9.76	0.000	0.639	0.647
	DRD2	17	5	94%	159	24.8	0.000	0.295	0.296	67.5	0.000	0.036	0.025
	DRD3	6	3	97%	289	2.49	0.174	0.881	0.917	23.9	0.000	0.219	0.201
	DRD4	1	1	100%	975	3.40	0.009	0.335	0.396	9.46	0.000	0.193	0.183
	VNTR[2]					11.9	0.000	0.247		12.5	0.000	0.314	
	Joint					12.0	0.000	0.398		25.8	0.000	0.207	

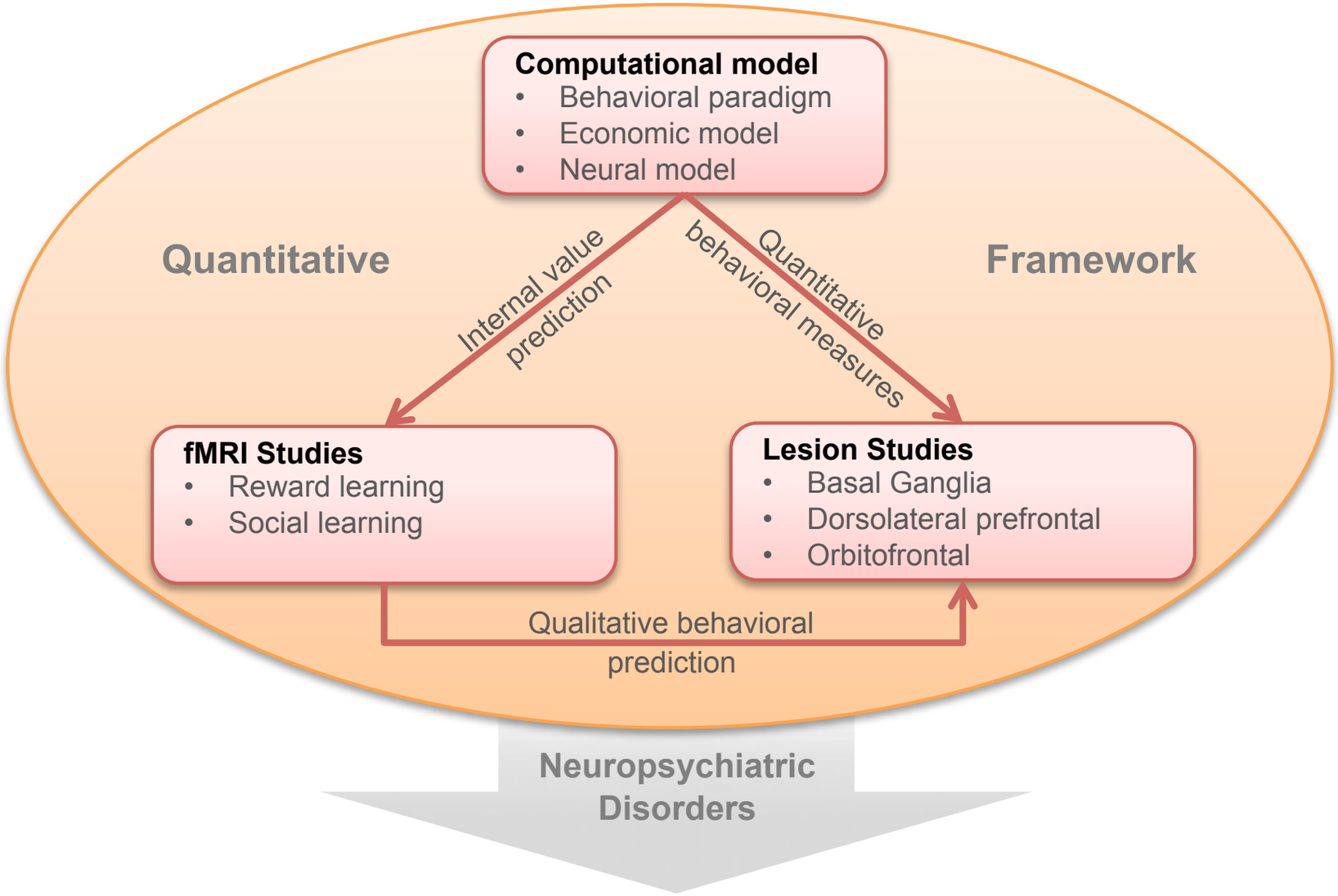
[1] Number of matched comparison genes chosen at random from the GWAS dataset, subject to availability.
[2] DRD4 classification: 2R carrier, 4R4R, all others.
[3] DAT classification: 9R carrier, all others.
[4] MAOA classification: 3R3R, all others.

Bigger Picture

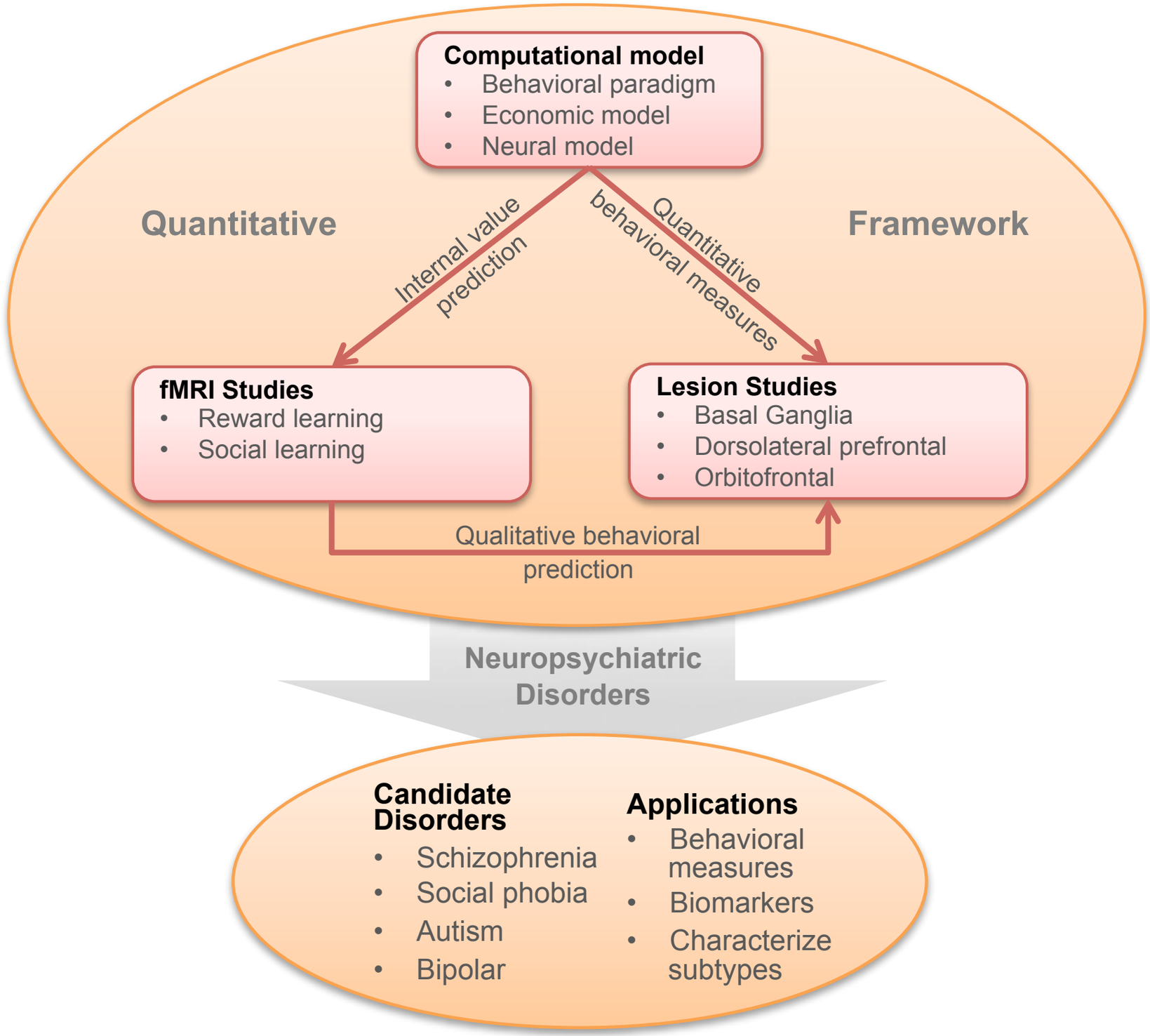
Bigger Picture



Bigger Picture



Bigger Picture

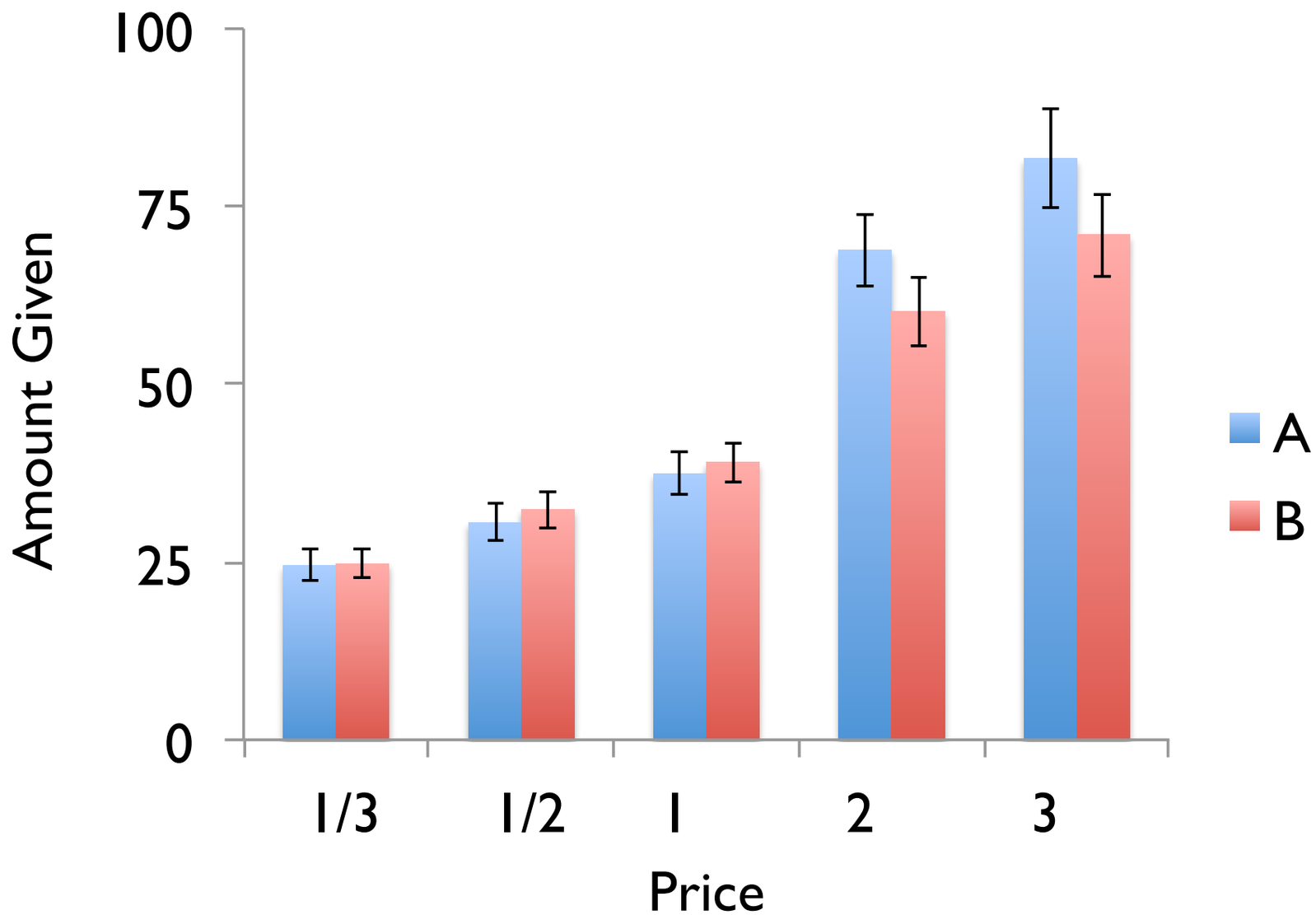


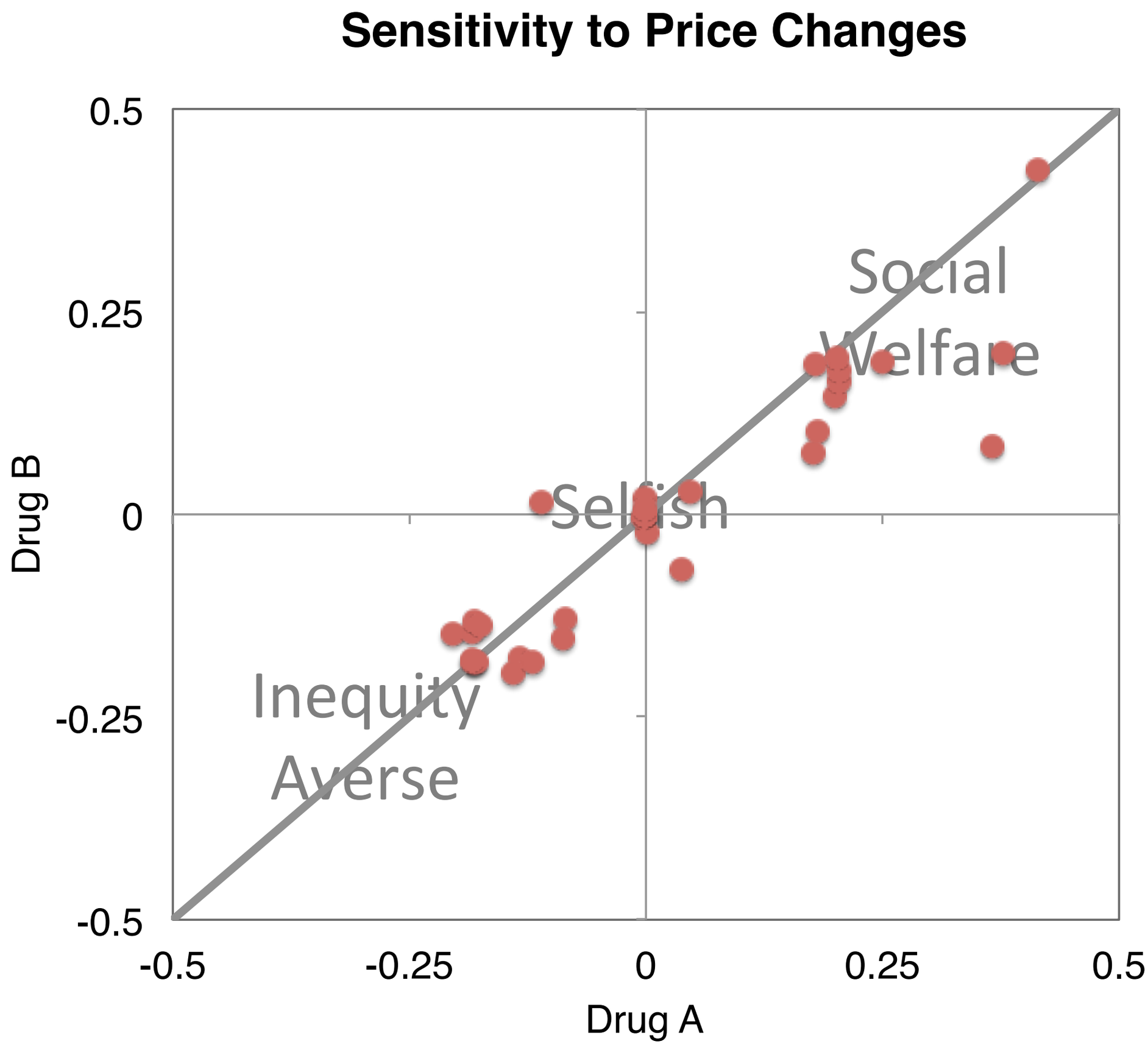
Acknowledgements

- Neuroecon Lab
  - Anna Jenkins
  - Ignacio Saez
  - Eric Set
  - Kenji Kobayashi
  - Yuping Chen
  - Dan Walsh
- Knight Lab
  - Robert T. Knight
  - Donatella Scabini
- B2SEE Lab
  - Soohong Chew
  - Richard Ebstein
  - Songfa Zhong
- NIMH
- Hellman Foundation
- RWJ Foundation
- Risk Management Institute

END

Tolcapone Changes Fairness Concerns in Dictator Game



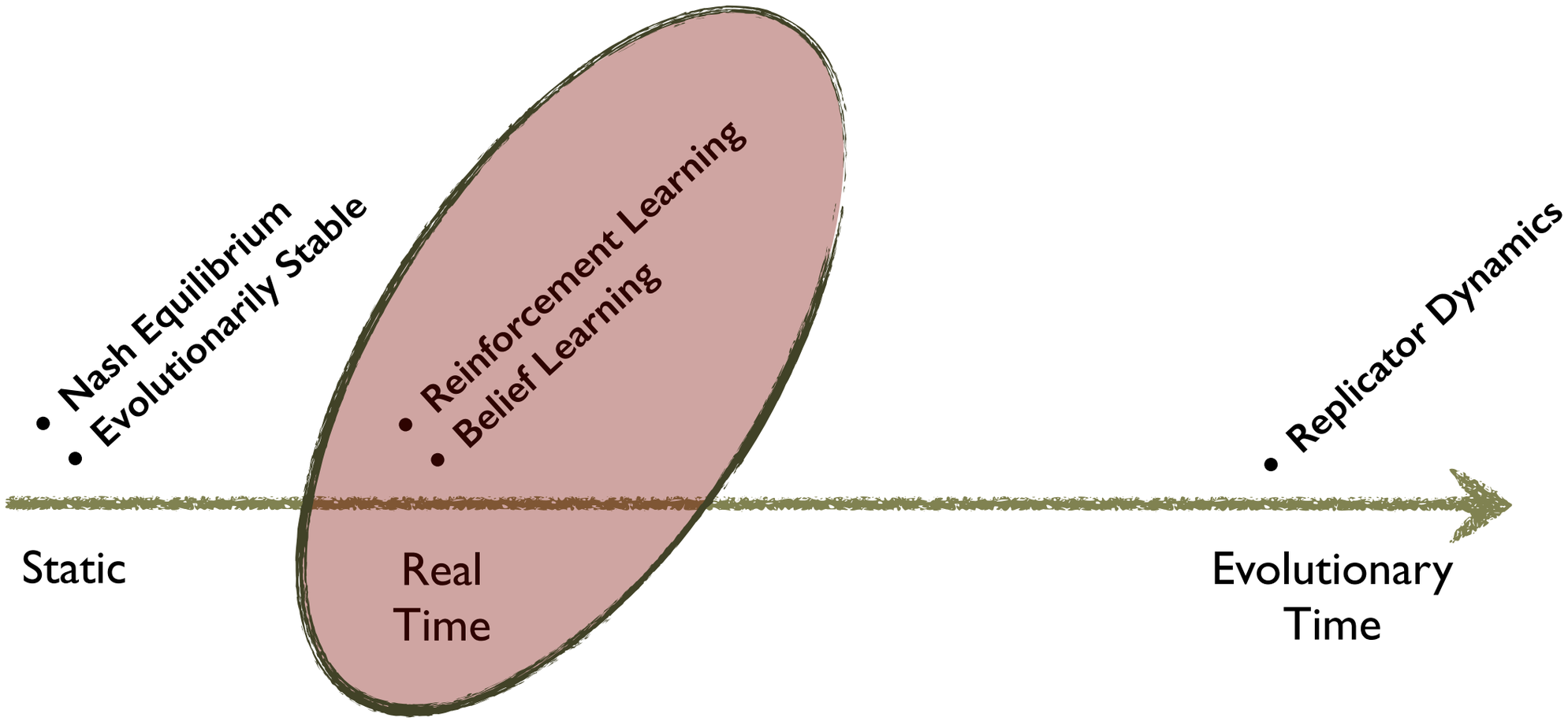

Sensitivity to Price Changes
0.5
0.25
0
-0.25
-0.5
Drug B
-0.5
-0.25
0
0.25
0.5
Drug A
Social Welfare
Selfish
Inequity Averse

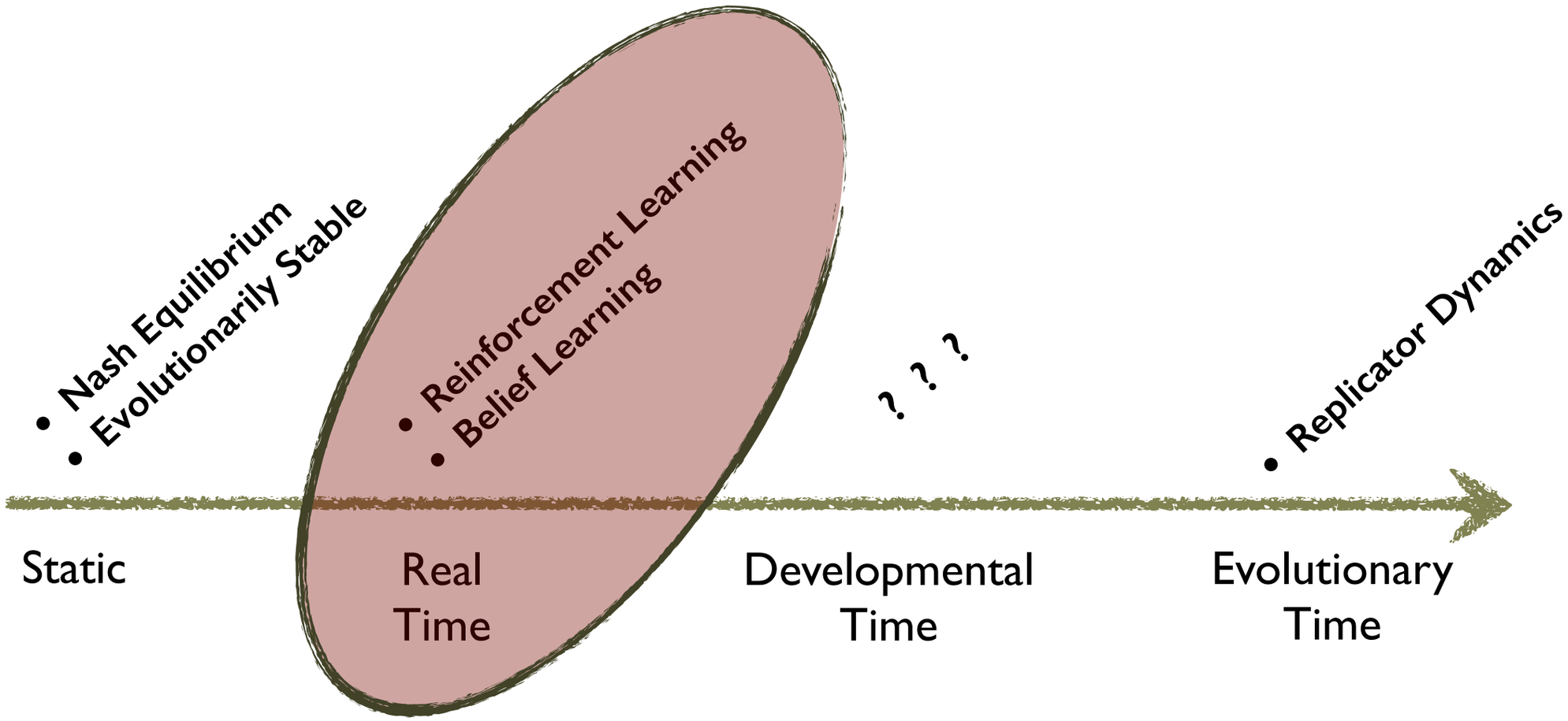
Static

- Nash Equilibrium
- Evolutionarily Stable

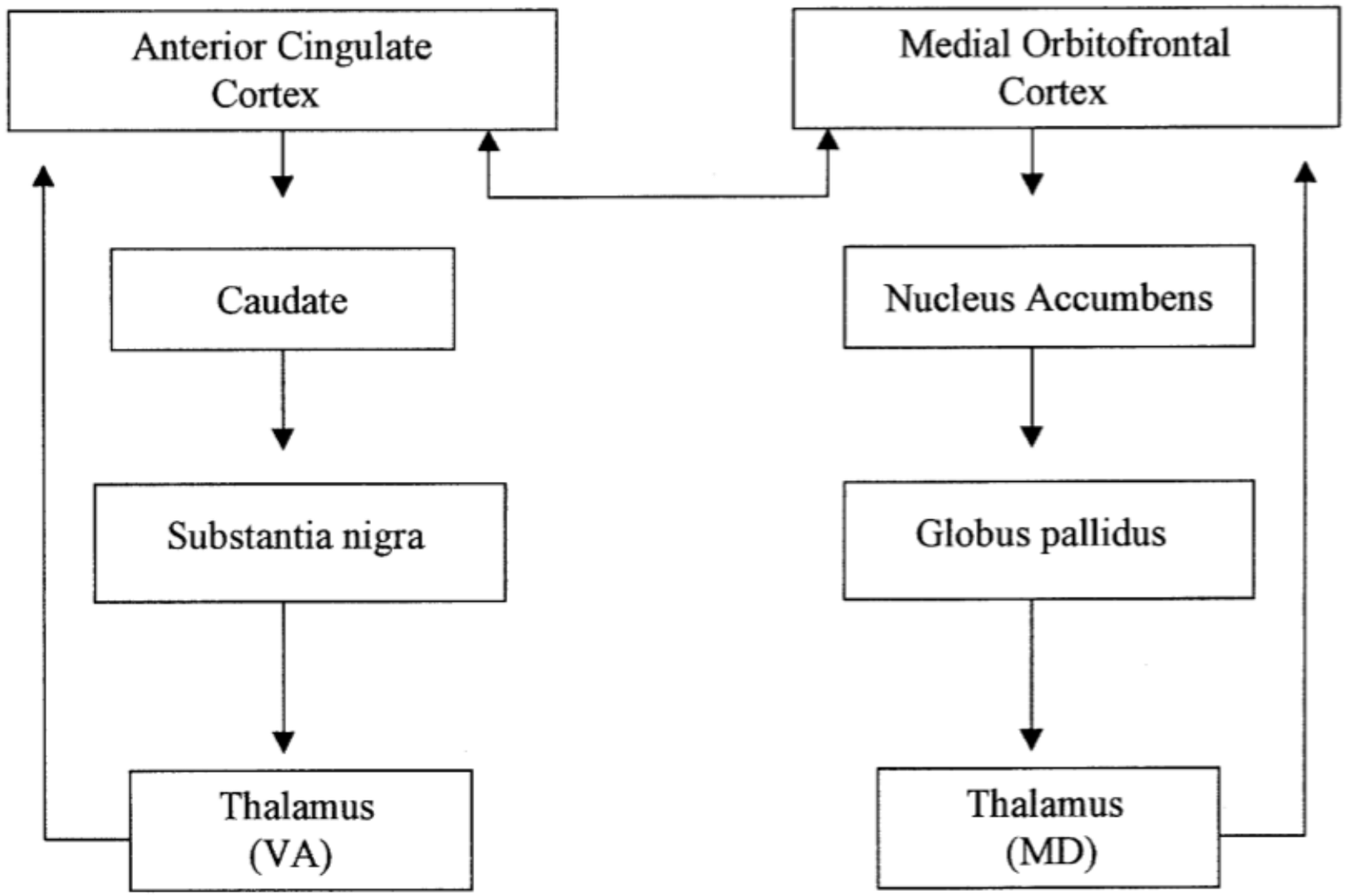
Static


• Nash Equilibrium
• Evolutionarily Stable
• Replicator Dynamics
Static
Evolutionary Time


• Nash Equilibrium
• Evolutionarily Stable
• Reinforcement Learning
• Belief Learning
• Replicator Dynamics
Static
Real Time
Evolutionary Time


• Nash Equilibrium
• Evolutionarily Stable
• Reinforcement Learning
• Belief Learning
? ? ?
• Replicator Dynamics
Static
Real Time
Developmental Time
Evolutionary Time

Neural Circuits



Mega MS, Cummings JL (1994) *Journal of Neuropsychiatry and Clinical Neurosciences* 6:358–370.

Neural Circuits

Diseases	Cortical level	Subcortical level
Neurodegenerative		
Alzheimer's disease	+	–
Corticobasal degeneration	+	+
Frontotemporal dementia	+	–
Huntington's disease	–	+
Parkinson's disease	–	+
Progressive supranuclear palsy	–	+
Multiple system atrophy	+	+
Psychiatric		
Schizophrenia	+	+
Obsessive compulsive disorder	–	+
Depression	+	+
Gilles de la Tourette's syndrome	–	+

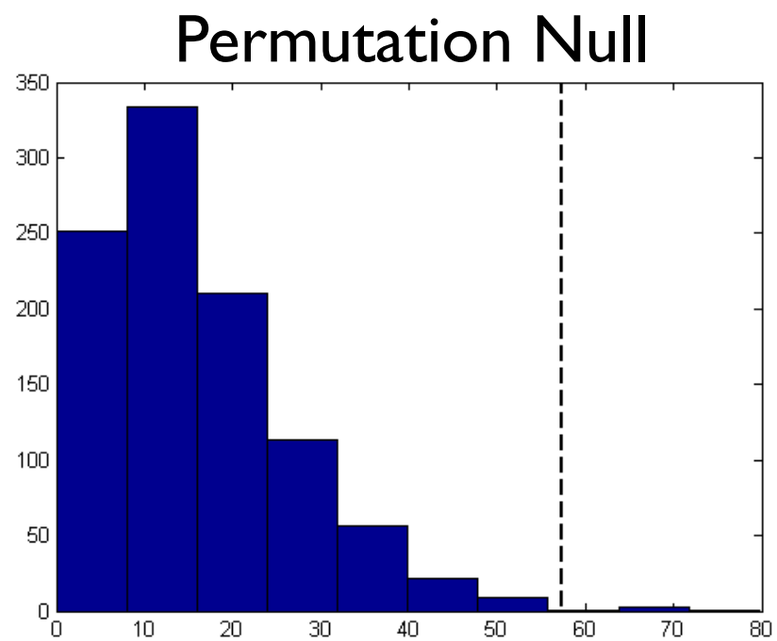
Mega MS, Cummings JL (1994) *Journal of Neuropsychiatry and Clinical Neurosciences* 6:358–370.

Frontostriatal circuits

- In non-social domain: (imaging + lesion)
  - mPFC: Stimulus-reward learning, flexible behavior, etc
  - BG: reward prediction error
- In social domain: (imaging + ??)
  - mPFC: other's reward, social preferences.
  - BG: social reward reinforcement, trust

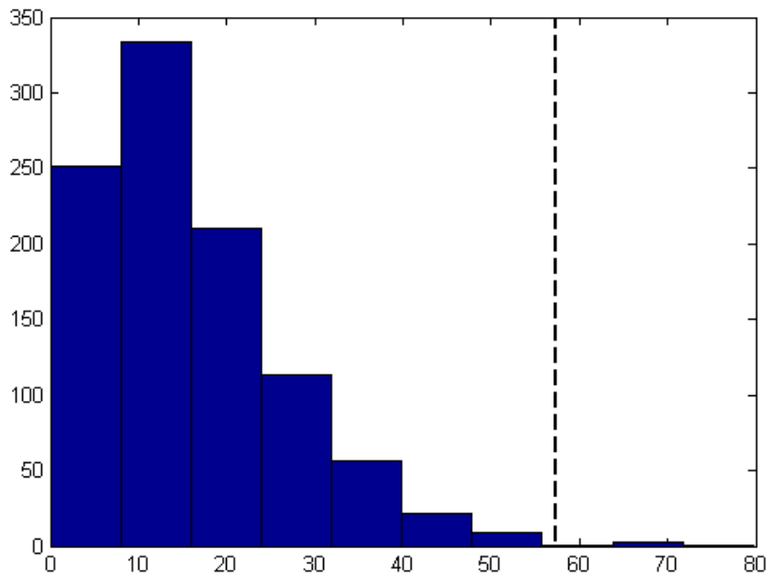
Catechol-O-methyl transferase
(COMT)

Catechol-O-methyl transferase
(COMT)

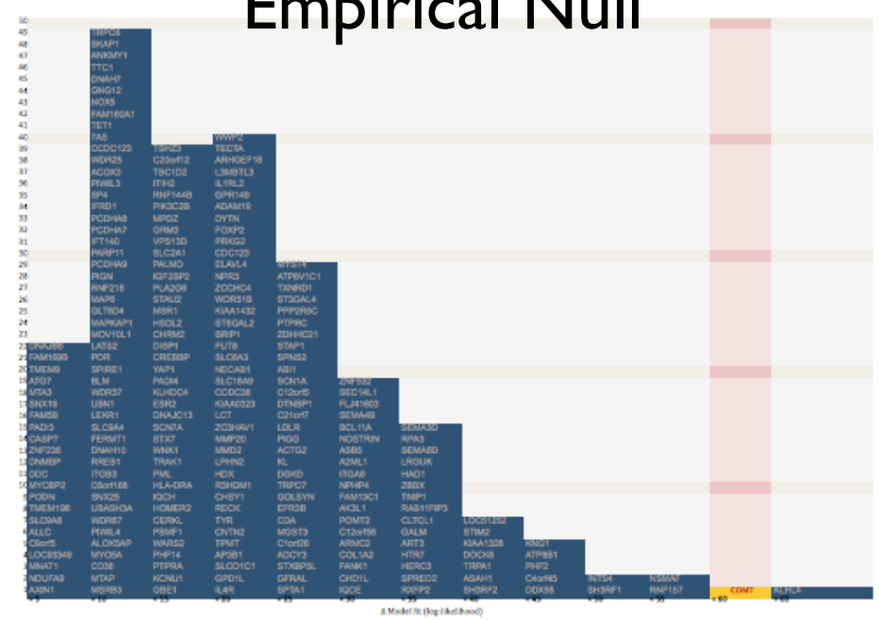


Catechol-O-methyl transferase
(COMT)

Permutation Null

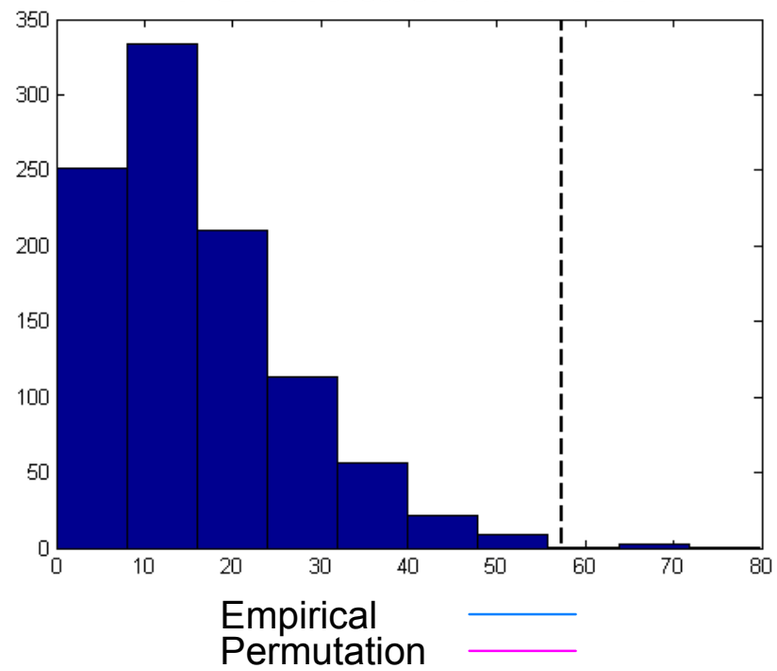


Empirical Null

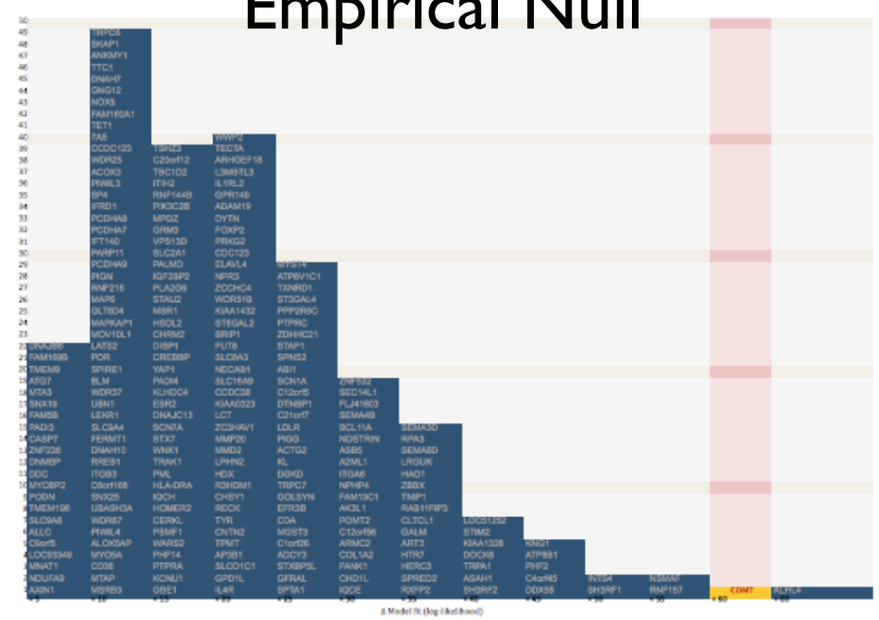


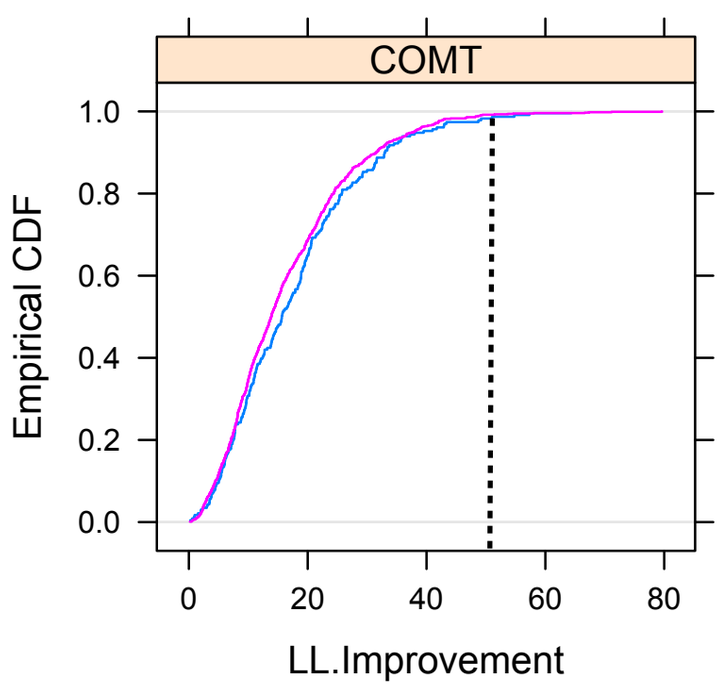
Catechol-O-methyl transferase
(COMT)

Permutation Null



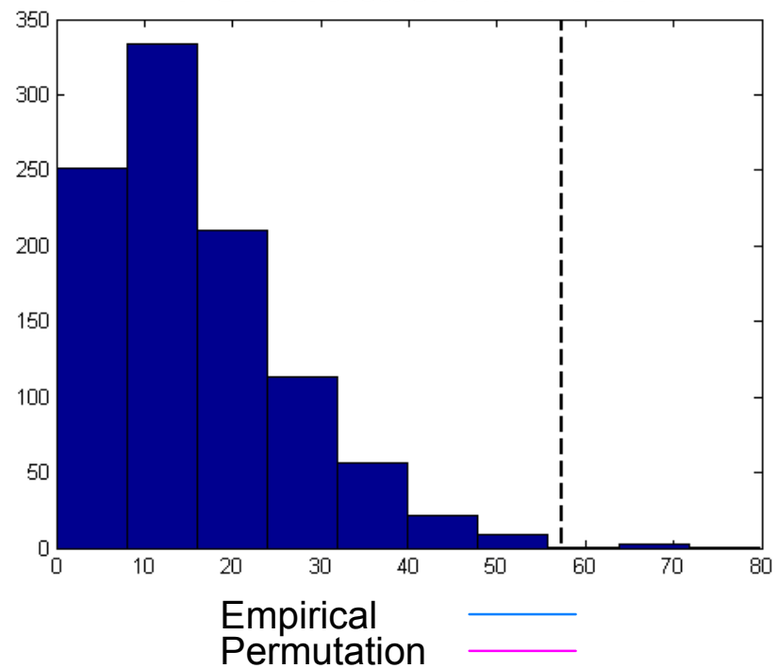
Empirical Null



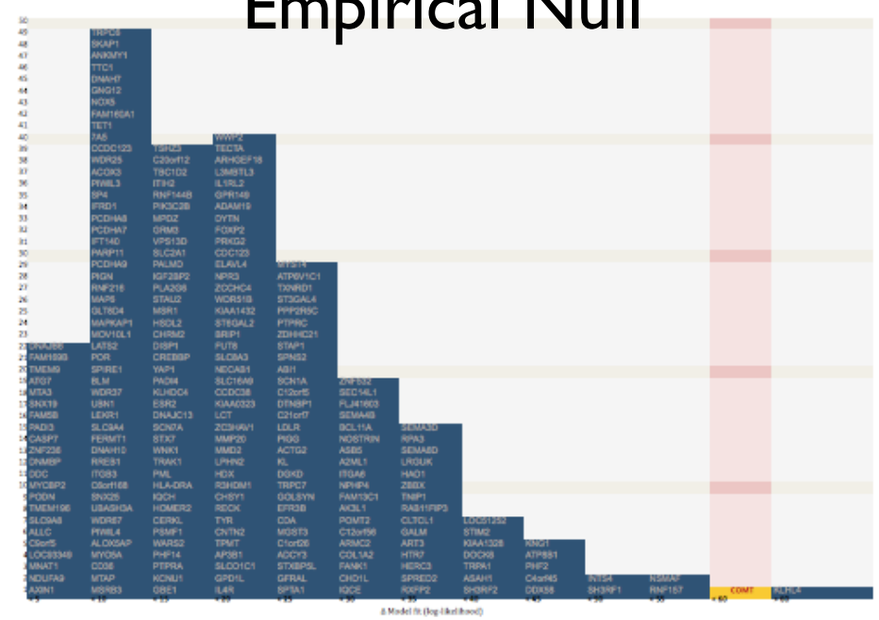


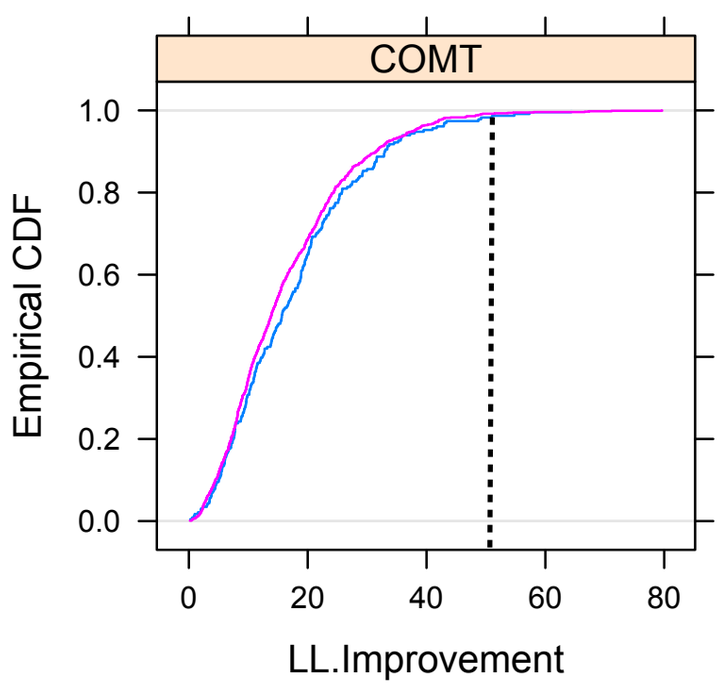
Catechol-O-methyl transferase
(COMT)

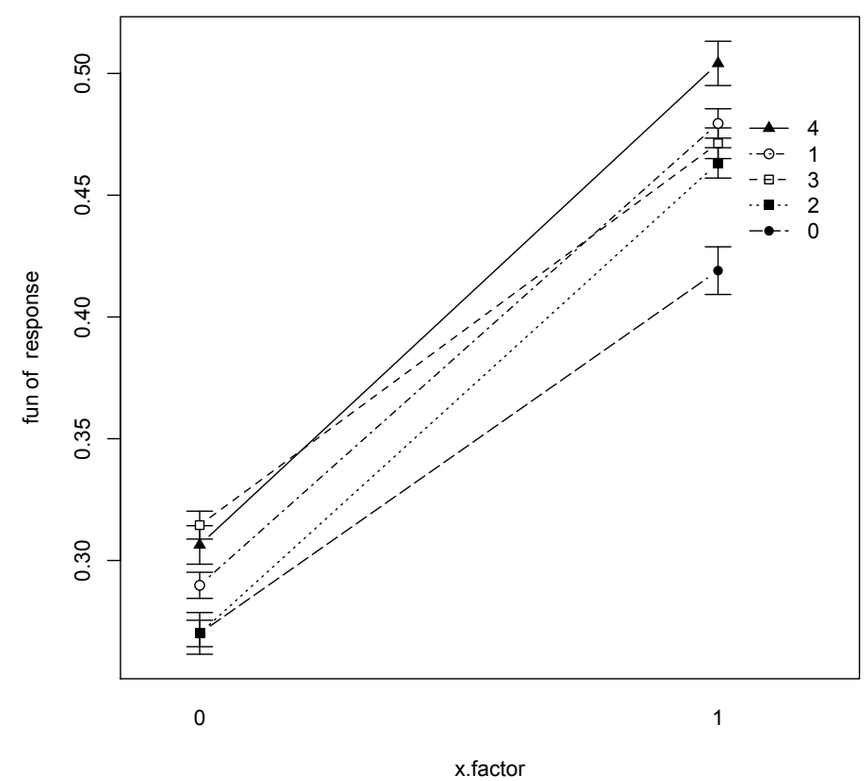
Permutation Null



Empirical Null







Current Direction

Current Direction

- Expand range of <u>behavior</u>

Current Direction

- Expand range of <u>behavior</u>
- Expand set of <u>measures</u> and <u>manipulations</u>