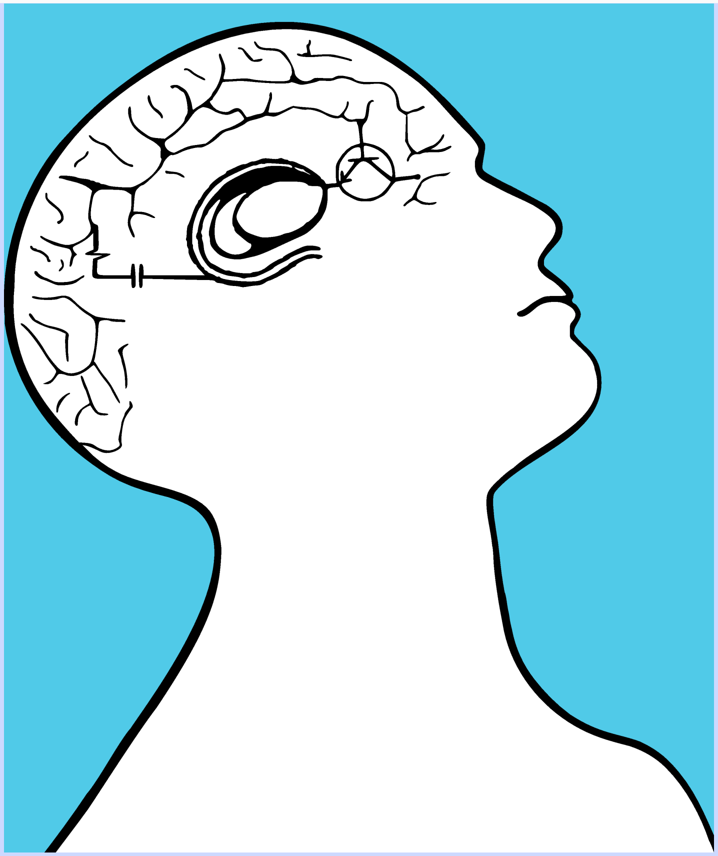
Interactive dynamics of corticostriatal function: Linking levels of computation



Michael Frank
Laboratory for Neural Computation and Cognition
Brown University

Multiple approaches to Computational Psychiatry

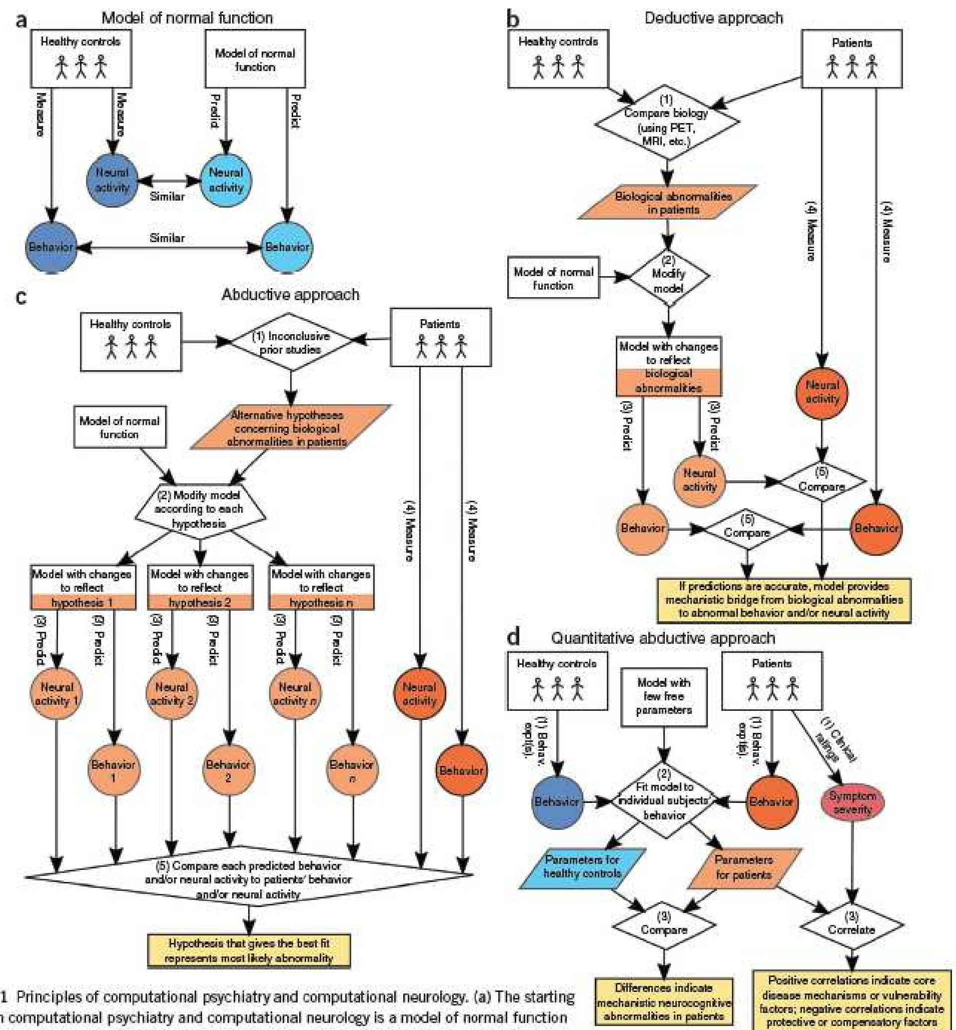


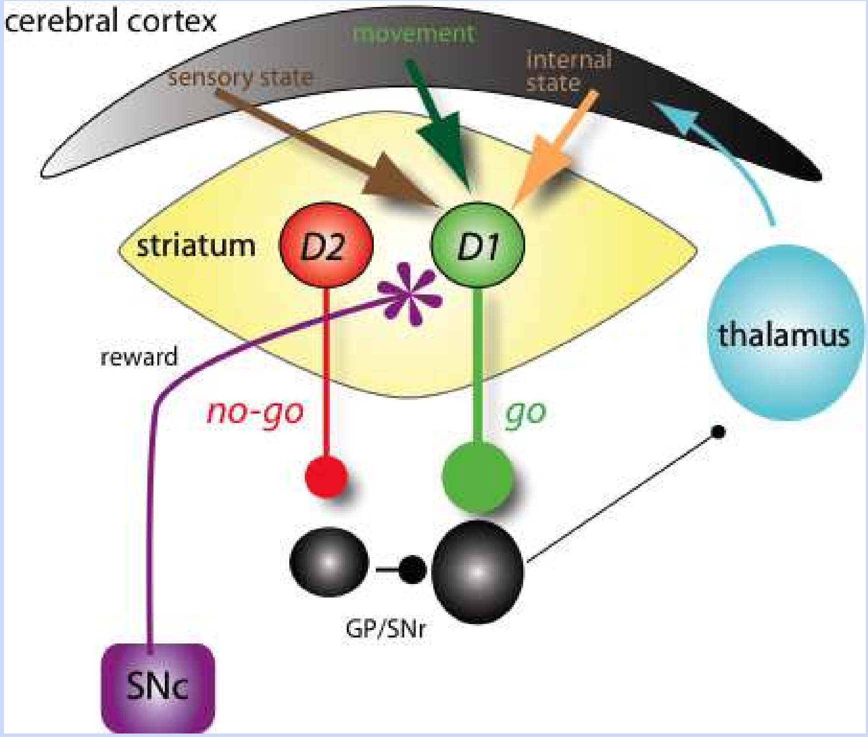
Figure 1 Principles of computational psychiatry and computational neurology. (a) The starting point in computational psychiatry and computational neurology is a model of normal function that captures key aspects of behavior, neural activity or both. Models at various levels of

Maia & Frank 2011, Nat Neurosci

Theories of striatal dopamine

- Performance (motor function, motivation: incentive salience)
- Reinforcement learning (prediction errors)
- *Dynamic interaction between the two*.
  Dopamine regulates striatal activity $\leftrightarrow$ learning
- Converging evidence in:
  - rodents
  - humans
  - models

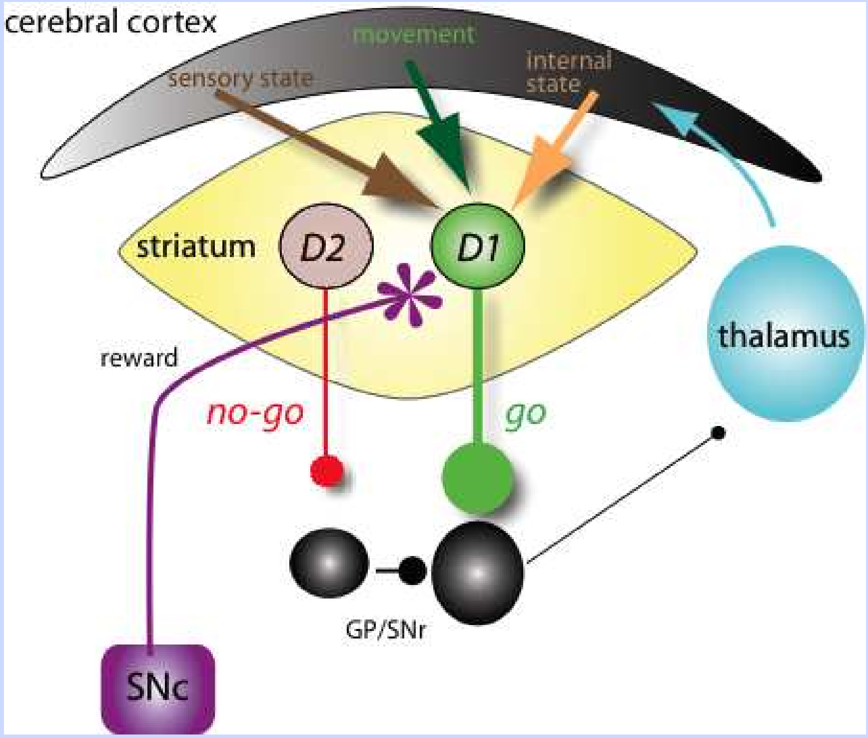
“Go” and “NoGo” pathways in the BG

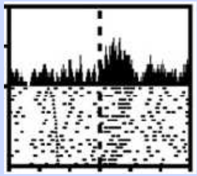


- “Direct pathway”: Go neurons gate actions
- “Indirect pathway”: NoGo neurons prevent gating
- Dopamine activates Go (D1), inhibits NoGo (D2)
- Phasic DA signals drive learning via dynamic activation effects

D1 effects on BG learning: Positive PE

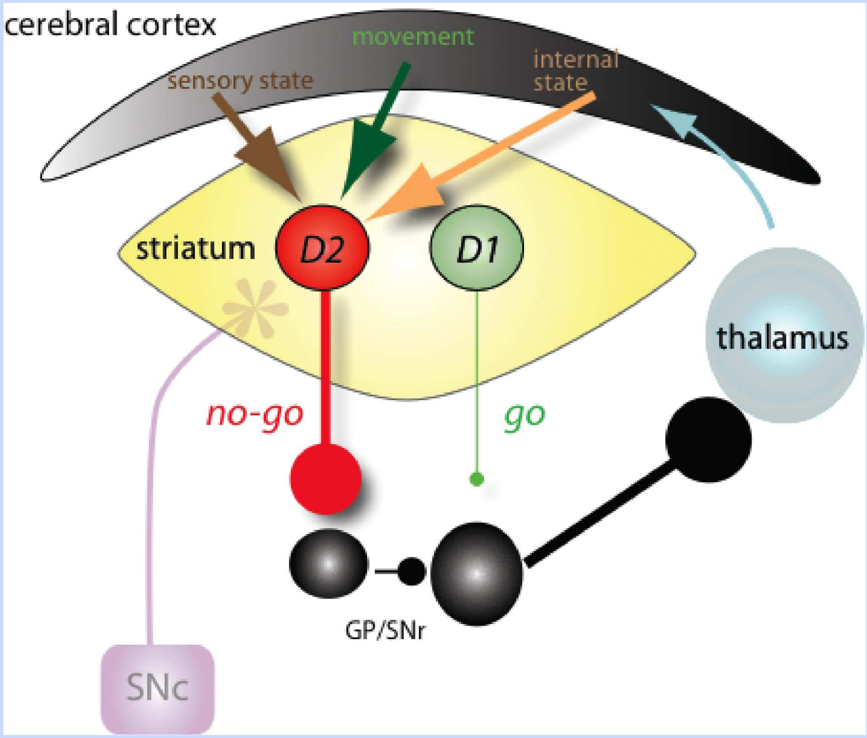
D1 effects on BG learning: Positive PE

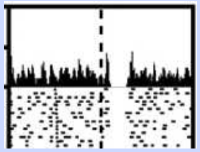




Three factor learning: presynaptic, postsynaptic and DA

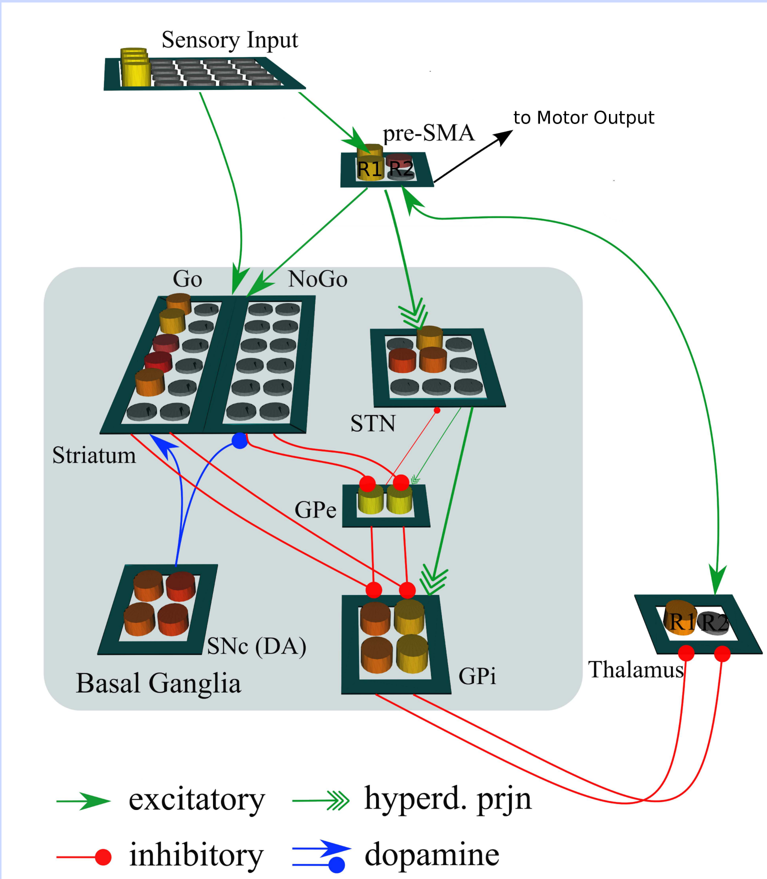
D2 effects on BG learning: Negative PE





Frank 2005

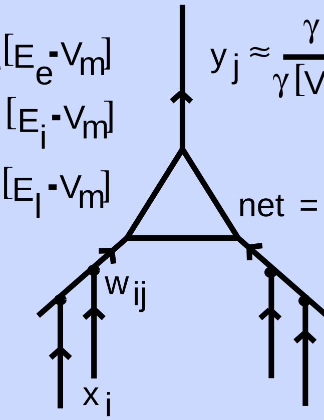
Neural circuit model of BG in learning / decision making



$$c\dot{V}_m = g_e \bar{g}_e [E_e - V_m] + g_i \bar{g}_i [E_i - V_m] + g_l \bar{g}_l [E_l - V_m] + \ldots$$

$$y_j \approx \frac{\gamma[V_m-\Theta]_+}{\gamma[V_m-\Theta]_+ + 1}$$

$$net = g_e \approx < x_i w_{ij} > + \frac{\beta}{N}$$



Frank, 2005, 2006 *J Cog Neurosci, Neural Networks*

Support for go/nogo learning mechanisms in rodents

Distinct Roles of Synaptic Transmission in Direct and Indirect Striatal Pathways to Reward and Aversive Behavior

Takatoshi Hikida,[1,2] Kensuke Kimura,[1,3] Norio Wada,[1] Kazuo Funabiki,[1] and Shigetada Nakanishi[1,*]

Distinct roles for direct and indirect pathway striatal neurons in reinforcement

Alexxai V Kravitz[1,4], Lynne D Tye[1,2,4] & Anatol C Kreitzer[1–3]

Phasic Firing in Dopaminergic Neurons Is Sufficient for Behavioral Conditioning

Hsing-Chen Tsai,[1,2]* Feng Zhang,[2,]* Antoine Adamantidis,[3] Garret D. Stuber,[4] Antonello Bonci,[4] Luis de Lecea,[3] Karl Deisseroth[2,3]†

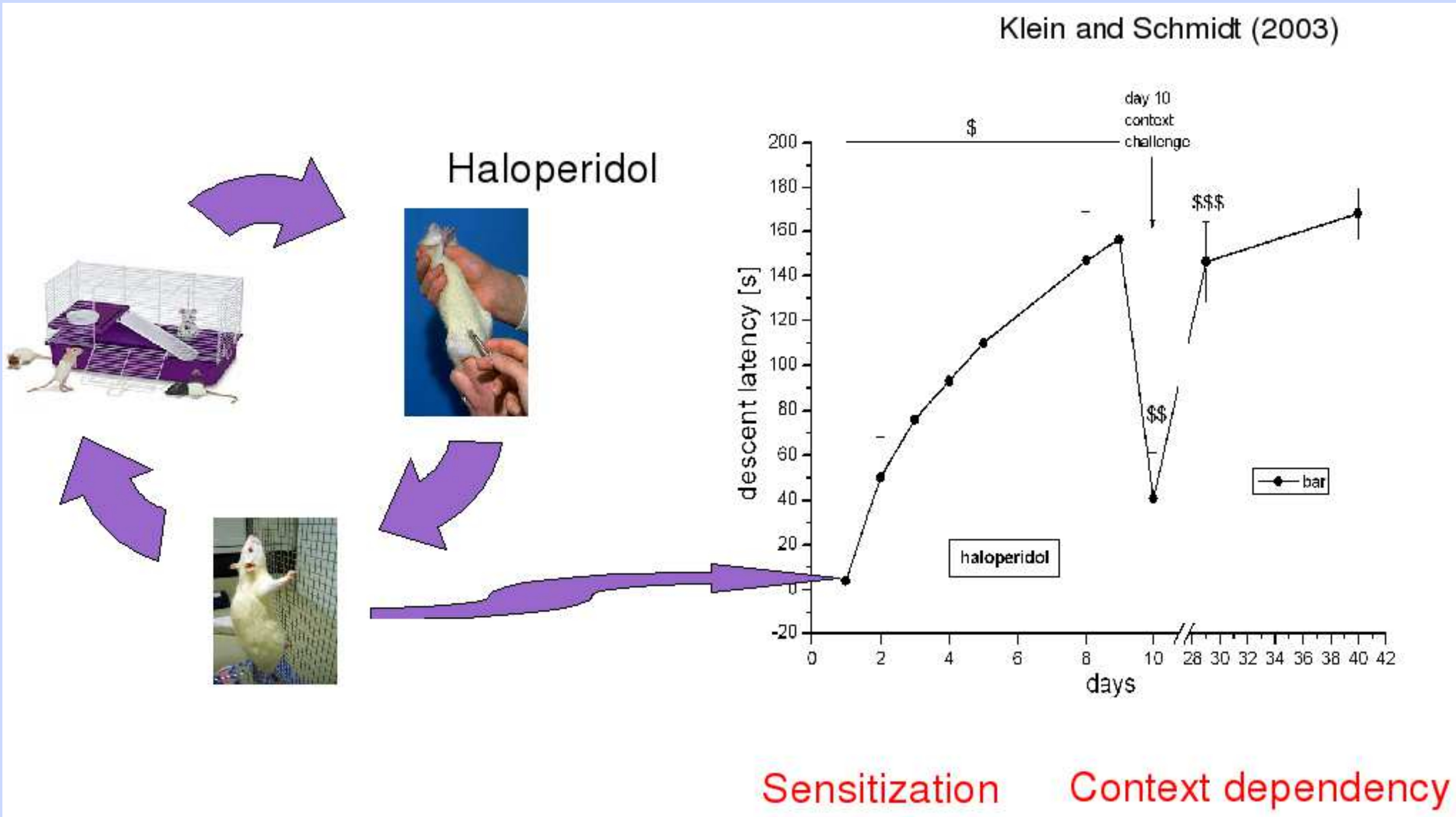
GABA Neurons of the VTA Drive Conditioned Place Aversion

Kelly R. Tan,[1] Cédric Yvon,[1,6] Marc Turiault,[1,6] Julie J. Mirzabekov,[2] Jana Doehner,[3] Gwenaël Labouèbe,[1] Karl Deisseroth,[2] Kay M. Tye,[2,4] and Christian Lüscher[1,5,*]

Dichotomous Dopaminergic Control of Striatal Synaptic Plasticity

Weixing Shen,[1] Marc Flajolet,[2] Paul Greengard,[2] D. James Surmeier[1]*

Can PD be learned?
Catalepsy sensitization by DA depletion or haloperidol



A case of exaggerated NoGo learning??

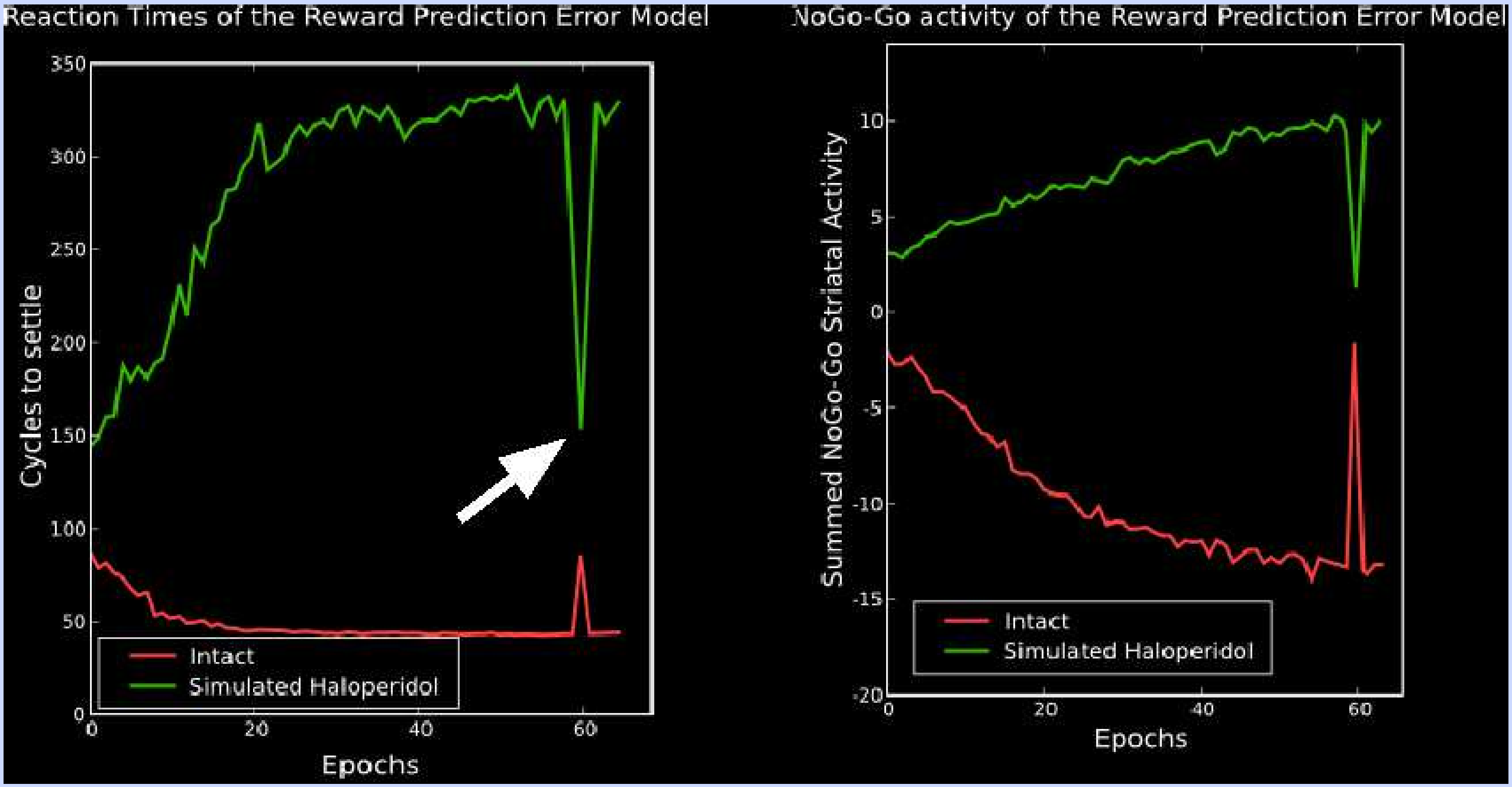
Haloperidol...

- potently blocks dopaminergic D2 receptors in the NoGo pathway:
  - $\uparrow$ NoGo activity (performance) Day et al 2008
  - $\uparrow$ NoGo plasticity (learning) Centonze et al 04; Hakansson et al 06; Shen et al 08

$\Rightarrow$Striatal D2 blockade induces synaptic potentiation of “NoGo” units:

Learned catalepsy (inaction)

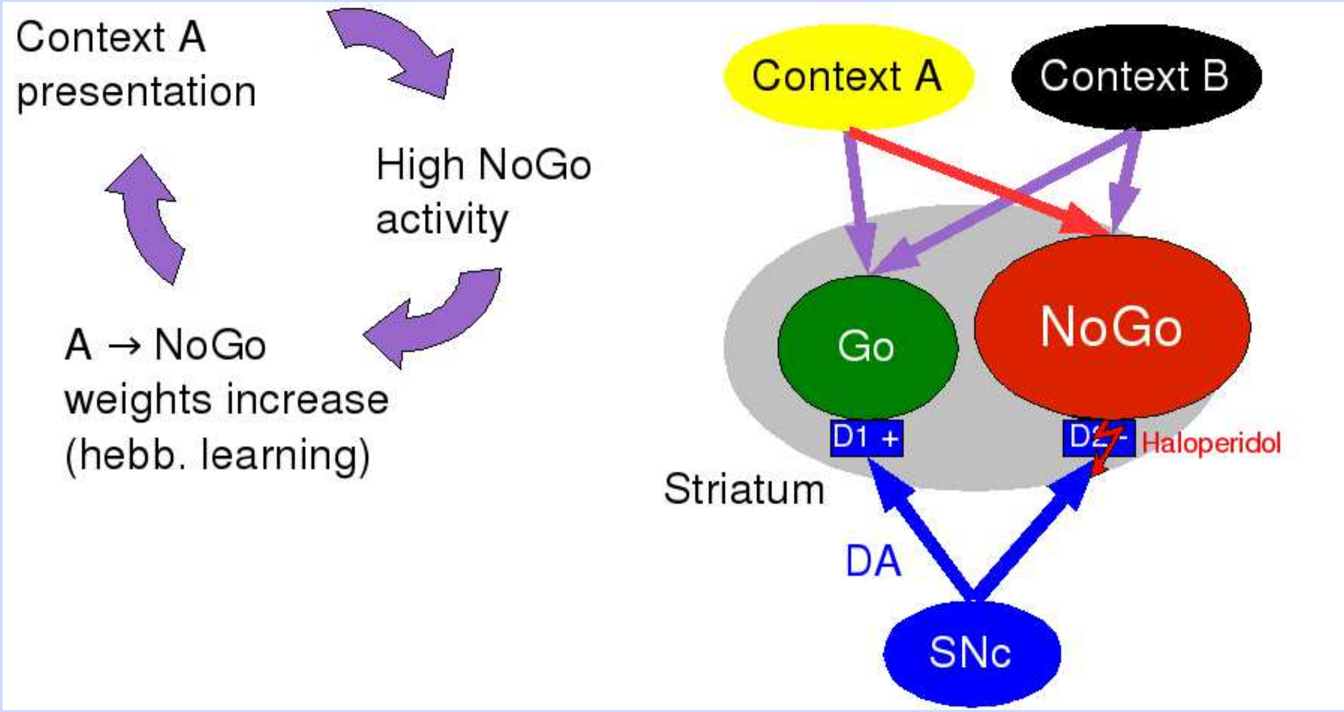
Model Results: Sensitization & Context-Dependency



Haloperidol simulated by postsynaptic striatal D2 blockade

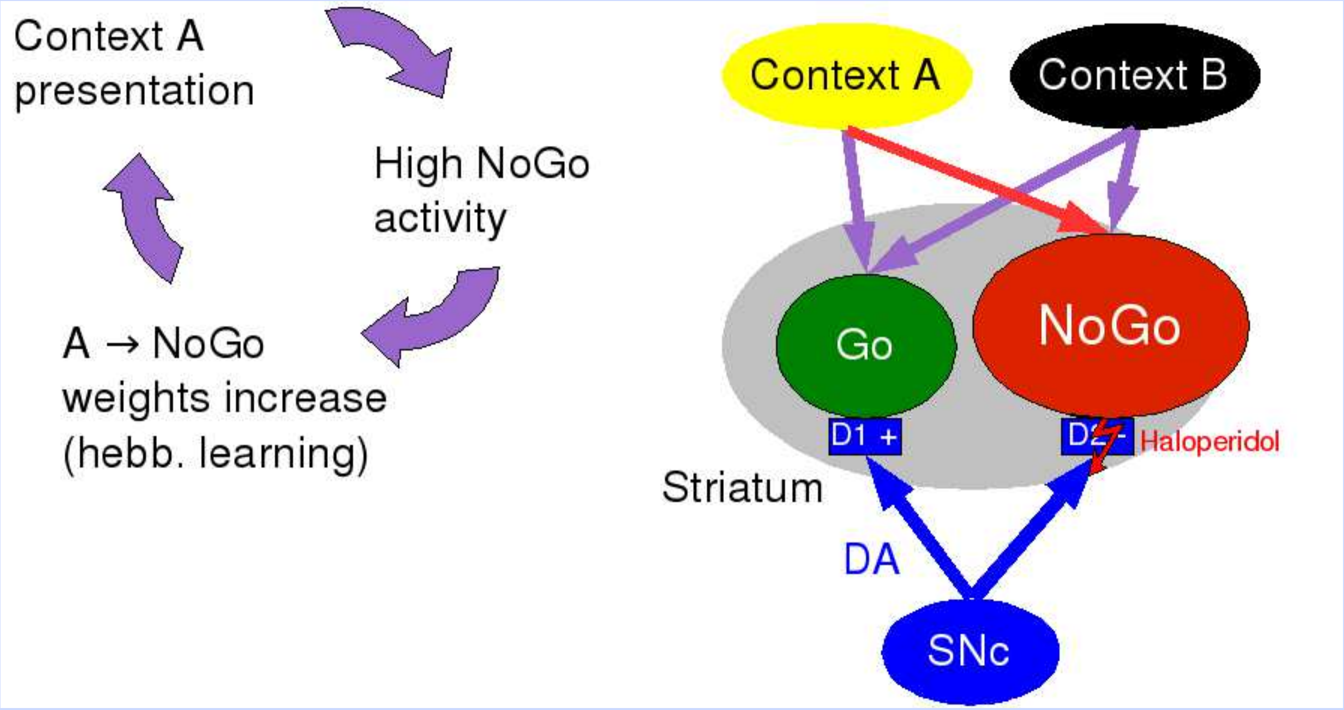
Wiecki et al 2009, *Psychopharmacology*

Mechanism



Context A
presentation
High NoGo
activity
A → NoGo
weights increase
(hebb. learning)
Context A
Context B
Go
NoGo
D1 +
D2 -
Haloperidol
Striatum
DA
SNc

Mechanism

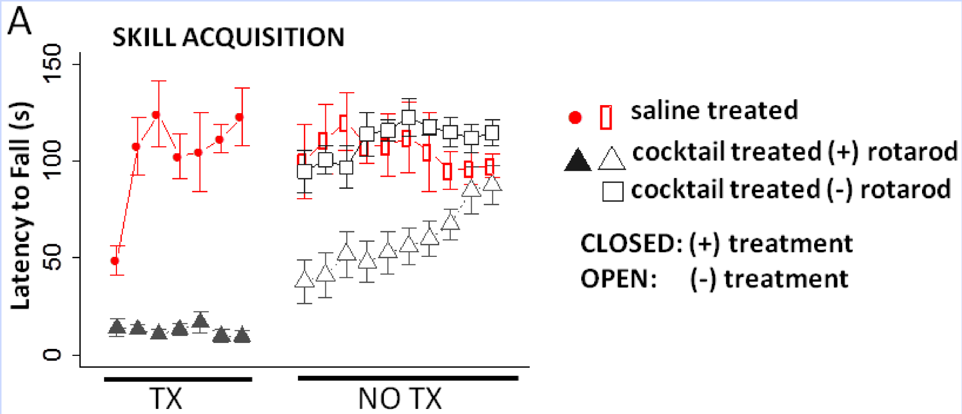


How does this mechanism affect learning / performance in active motor tasks?

Striatal-dependent motor learning task

(Accelerating Rotarod)

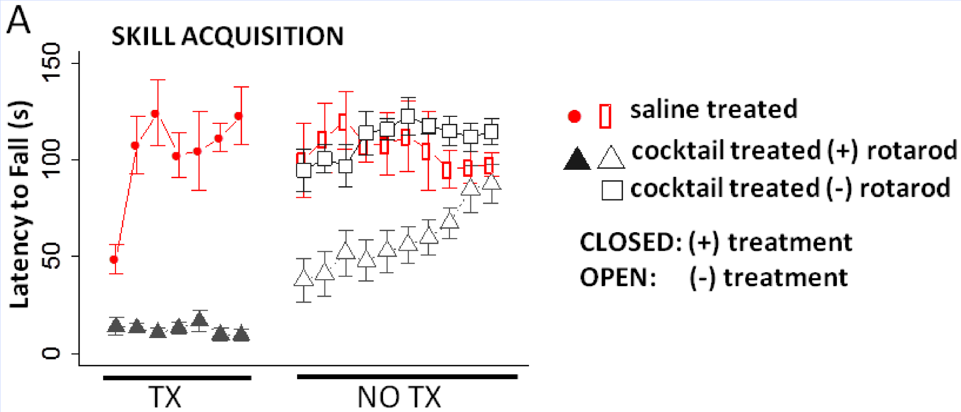
Data



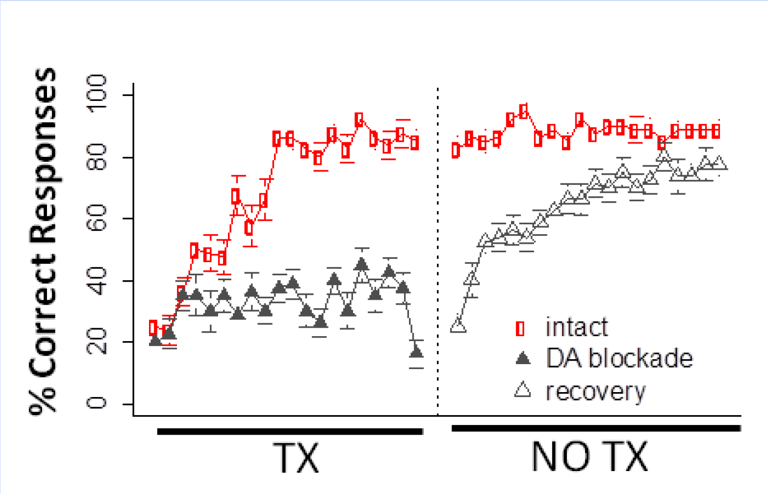
Striatal-dependent motor learning task

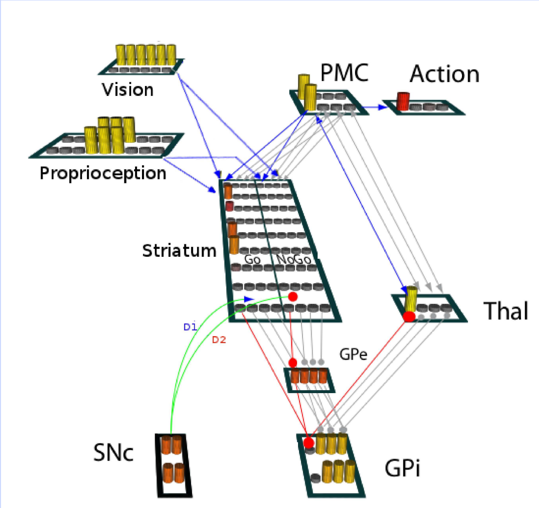
(Accelerating Rotarod)

Data



Model

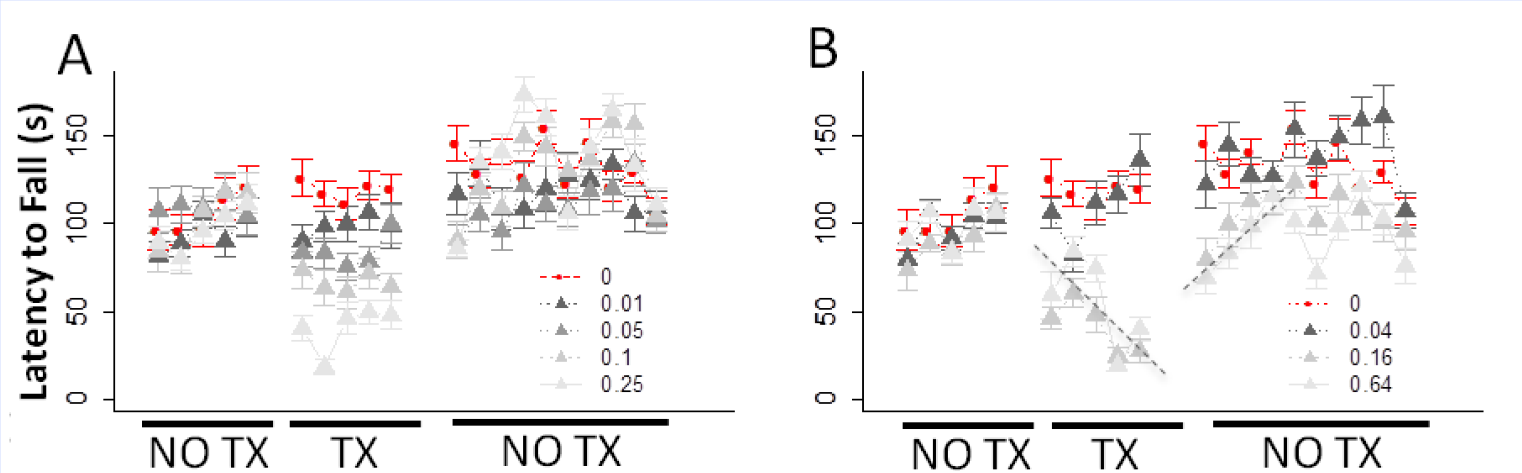




Beeler et al, submitted

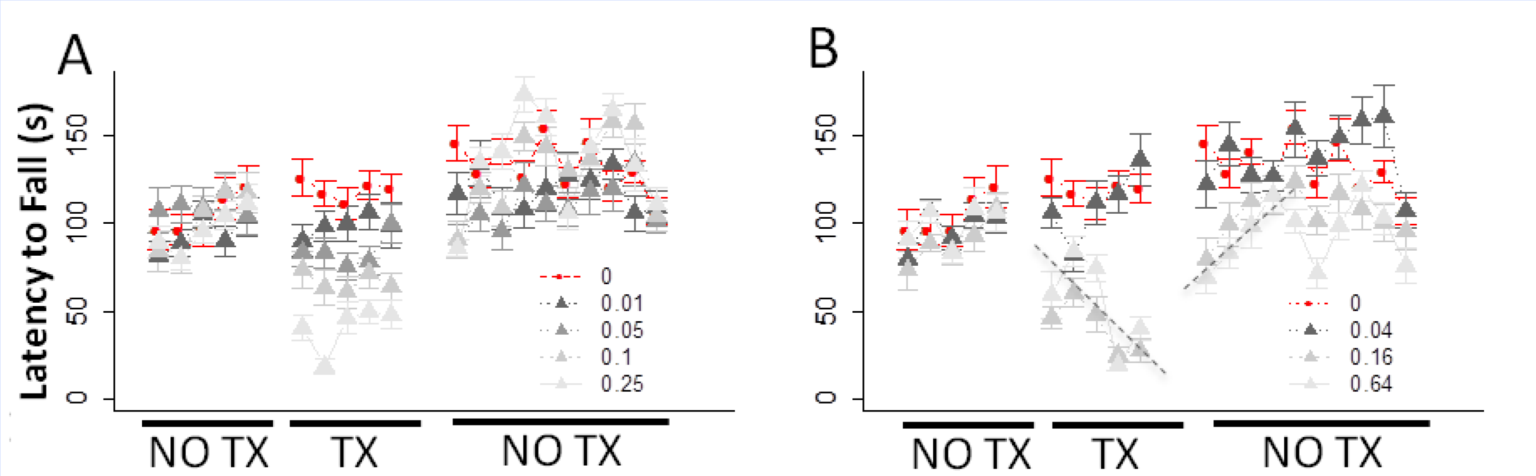
Selective D1 / D2 blockade during established skill

D1 block **D2 block**

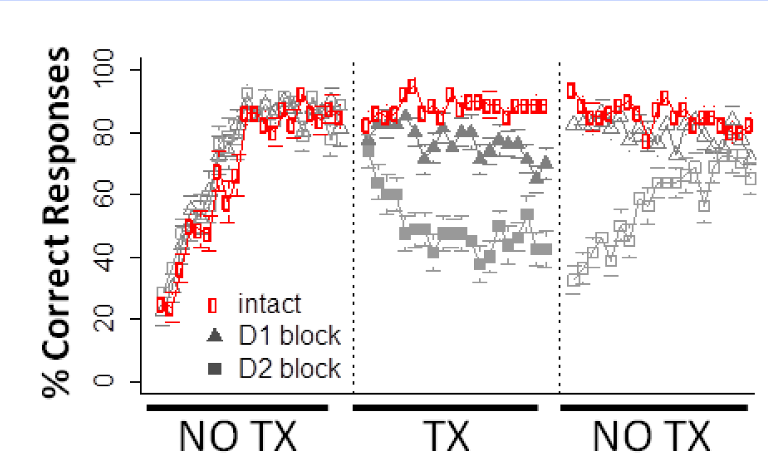


Selective D1 / D2 blockade during established skill

D1 block **D2 block**

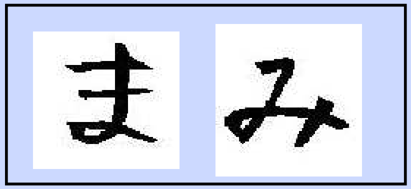


Model

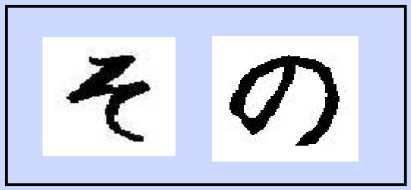


Human probabilistic reinforcement learning

Train



A (80/20) B (20/80)



C (70/30) D (30/70)

E (60/40) F (40/60)

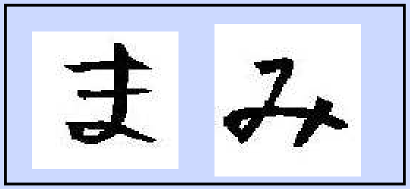
Test

A > CDEF **Choose A?**

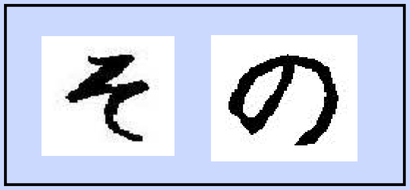
B < CDEF **Avoid B?**

Human probabilistic reinforcement learning

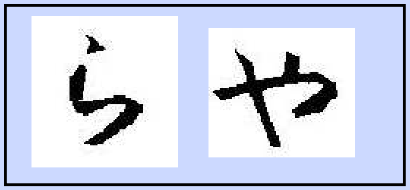
Train



A (80/20) B (20/80)



C (70/30) D (30/70)

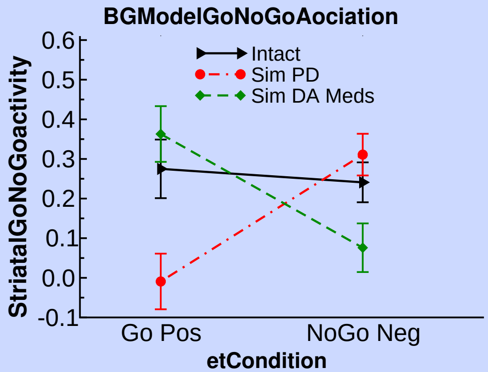


E (60/40) F (40/60)

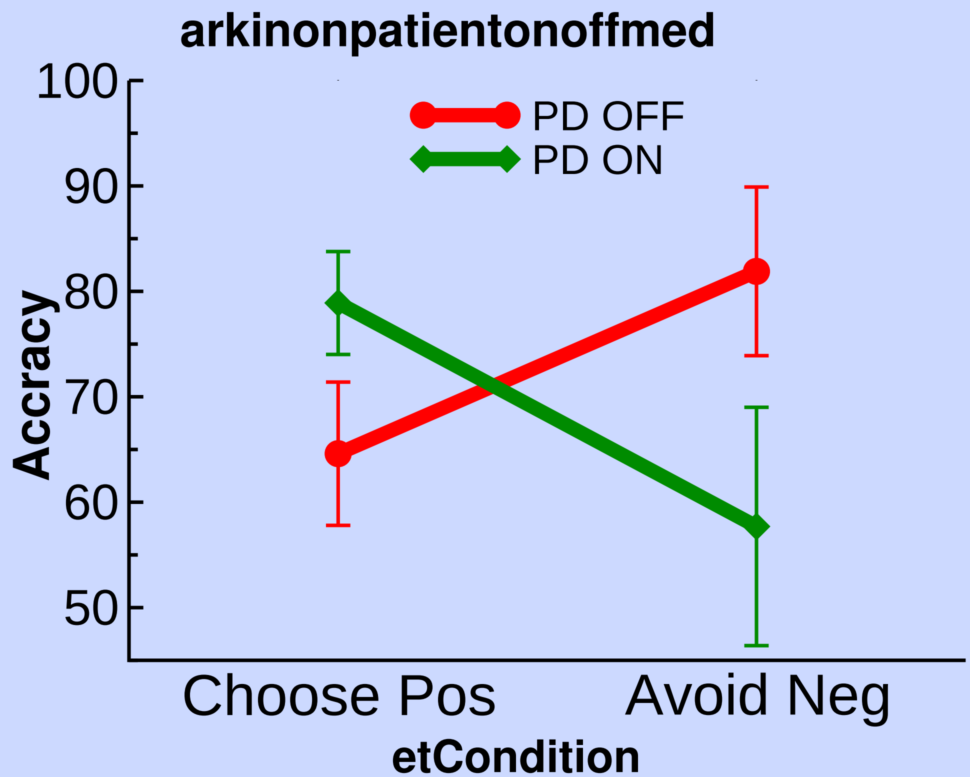
Test

A > CDEF **Choose A?**

B < CDEF **Avoid B?**



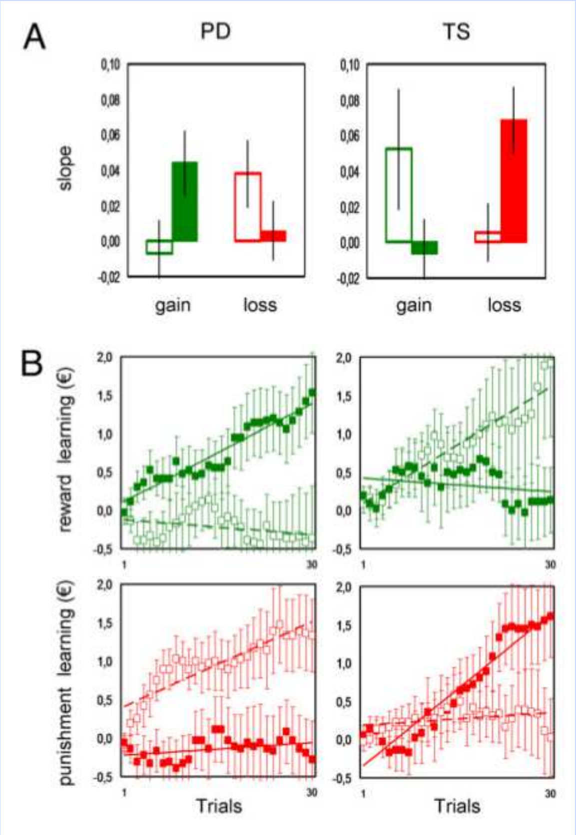
Testing the model:
Parkinson’s and medication effects



Frank, Seeberger & O’Reilly (2004)

(See also: Cools et al, 06, Frank et al 07, Moustafa et al 08, Bódi et al 09, Palminteri et al, 09, Voon et al 10)

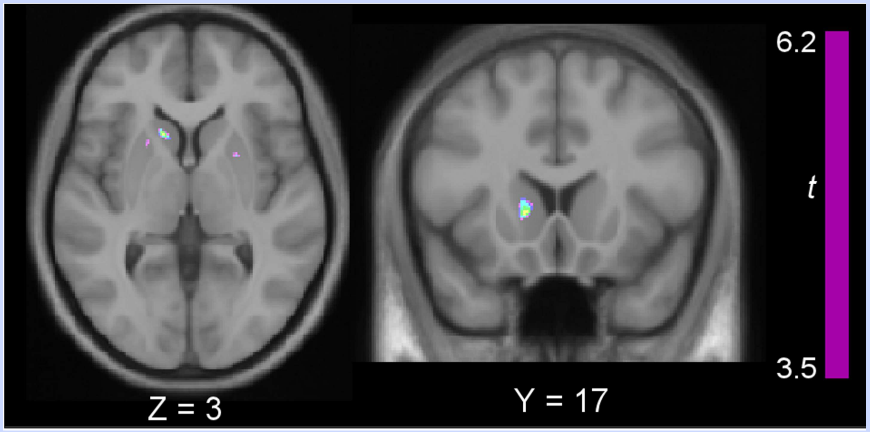
Converging evidence:
DA stimulation vs. D2 blockade on go/nogo learning

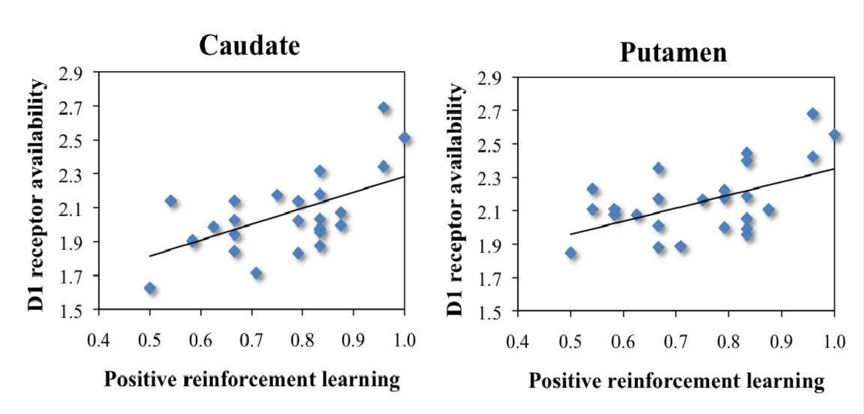


BG model: D2 blockade → ↑ NoGo activity, learning

Palminteri et al, 2009

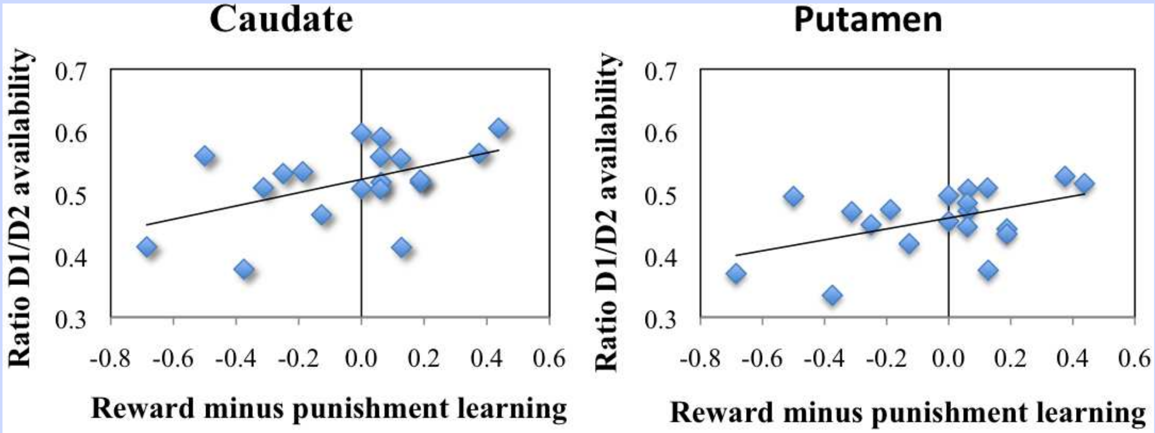
Individual differences in go learning: striatal D1 receptor binding





preliminary data, with Sylvia Cox and Alain Dagher

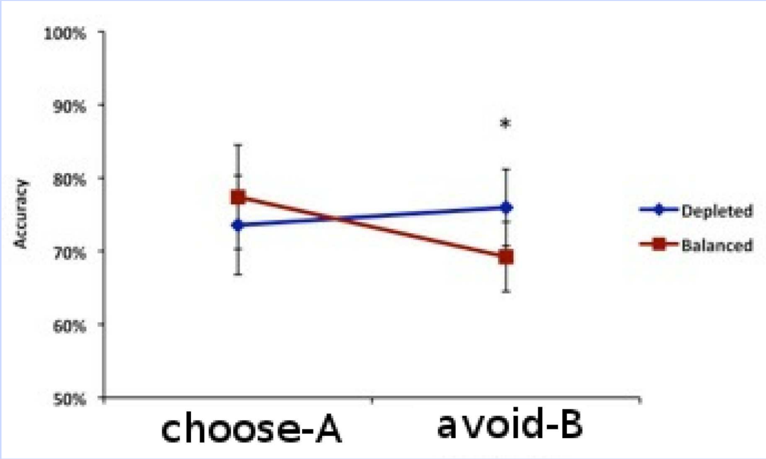
Individual differences in go/nogo learning: striatal D1/D2 receptor binding

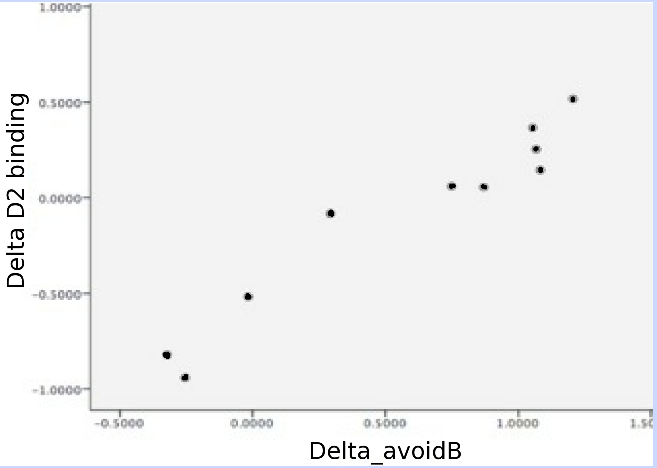


The ratio [^{11}C]SCH-23390 BP_{ND} (D1) to [^{11}C]raclopride BP_{ND} (D2) in the caudate (r=0.49,p=0.033) and putamen (r=0.49, p=0.035) predicts relative learning from positive versus negative outcomes on the PST

preliminary data, with Sylvia Cox and Alain Dagher

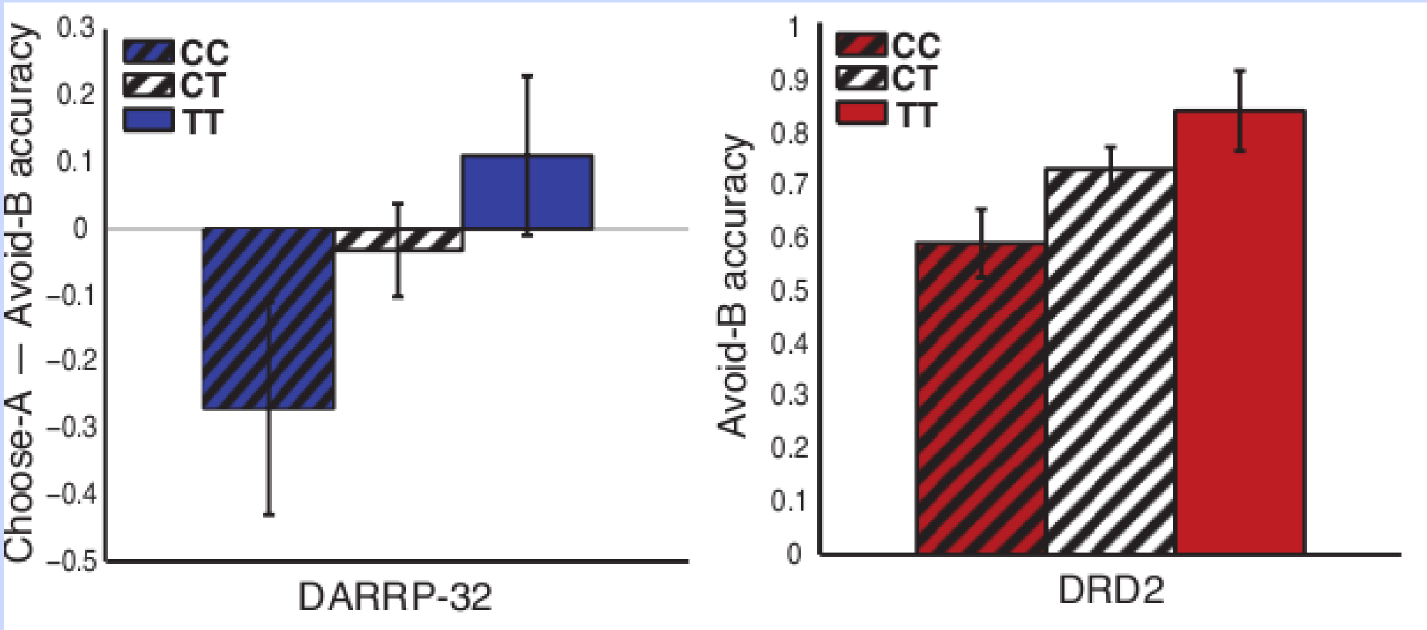
Tyrosine depletion: reducing DA levels improves avoidance/learning





preliminary data, with Alain Dagher

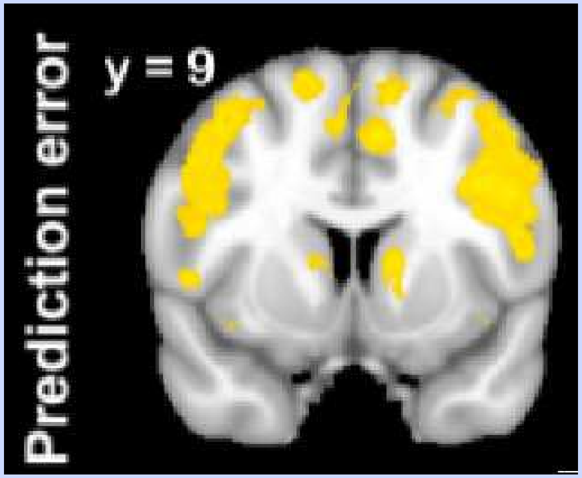
Genetic effects: Striatal dopaminergic variants

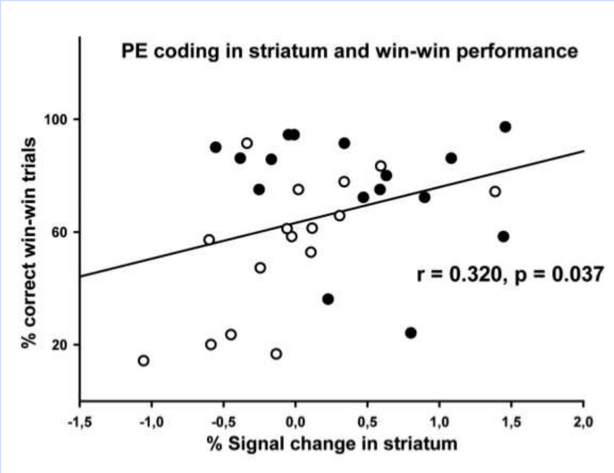


DARPP-32: required for corticostriatal D1-LTP and reward learning
(e.g. Stipanovich et al 08)

Doll et al, 11; Frank et al, 07, Klein et al 07; Frank & Hutchison 09; Frank & Fosella, 11..

DA drugs: Learning and motivational effects





- DA drug amplifies striatal RPE's during learning, increasing reward-based choice
- But: DA meds can also have motivational effects unrelated to learning per se. Meds modestly increase choose-A without RL (Smittenaar et al, 12; Shiner et al 12).

Jocham, Klein & Ullsperger 2011; see also Ott et al 2011

Formalizing BG/DA computations: Q learning, take one

Learning: Update "Q" values of actions in proportion to prediction errors δ:

$$Q_i(t+1) = Q_i(t) + \alpha_{\mathbf{W}}[\delta(t)]_+ + \alpha_{\mathbf{L}}[\delta(t)]_-$$

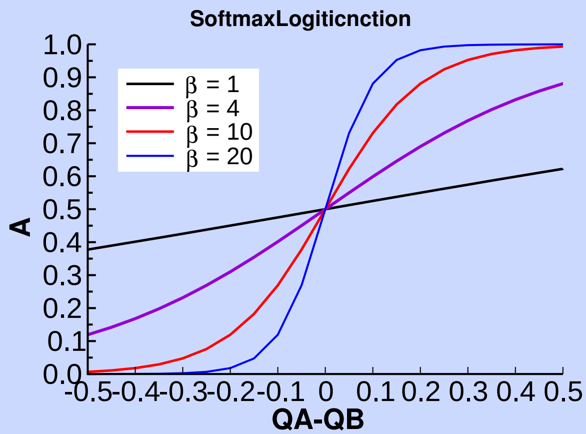
Formalizing BG/DA computations: Q learning, take one

Learning: Update "Q" values of actions in proportion to prediction errors δ:

$$Q_i(t+1) = Q_i(t) + \alpha_{\mathbf{W}}[\delta(t)]_+ + \alpha_{\mathbf{L}}[\delta(t)]_-$$

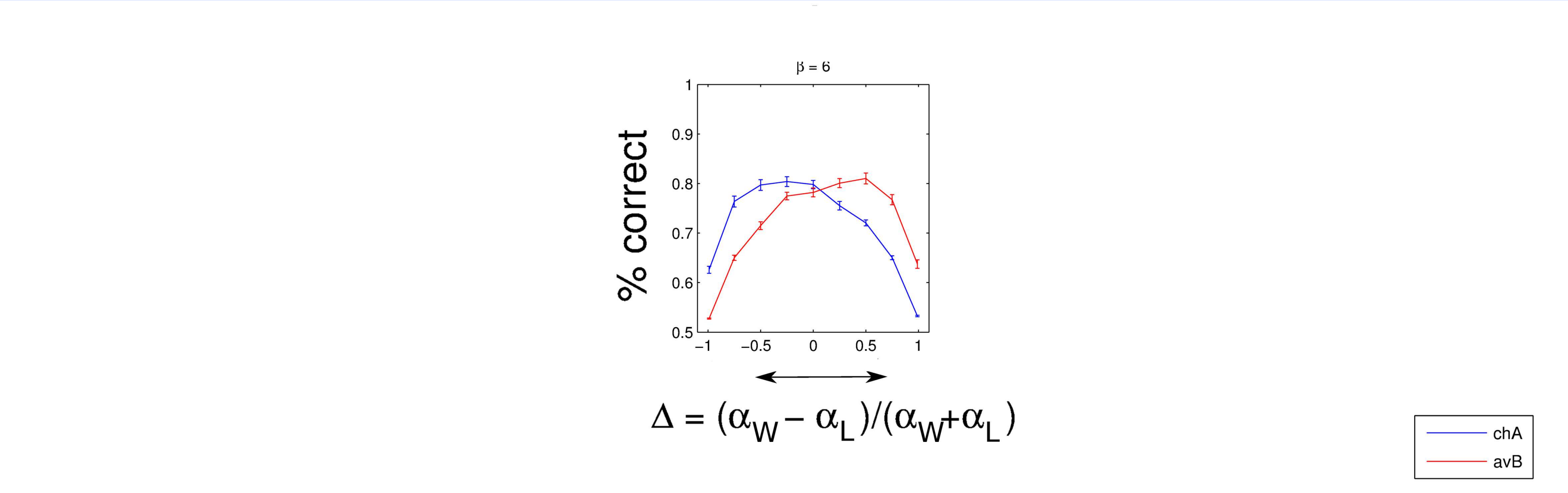
Choice: Standard "softmax" function for modeling striatal choice among Q values:

$$P_A = \frac{1}{1 + e^{-\beta(Q_A - Q_B)}}$$



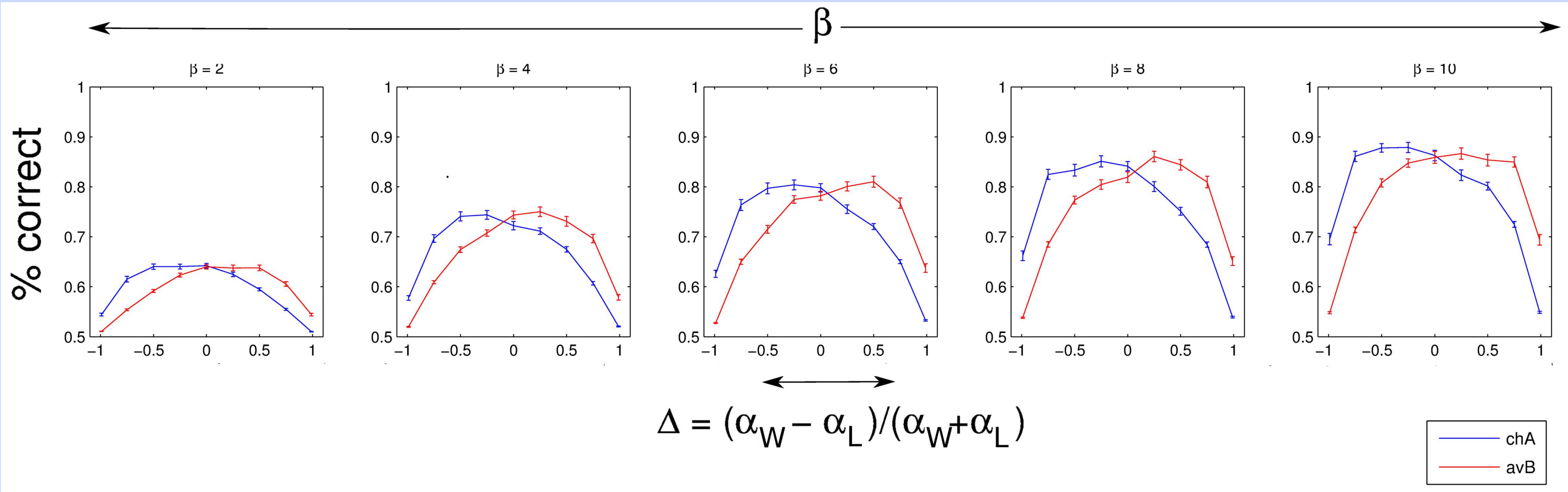
Can this basic model account for DA effects on learning / choice?

Standard Q model



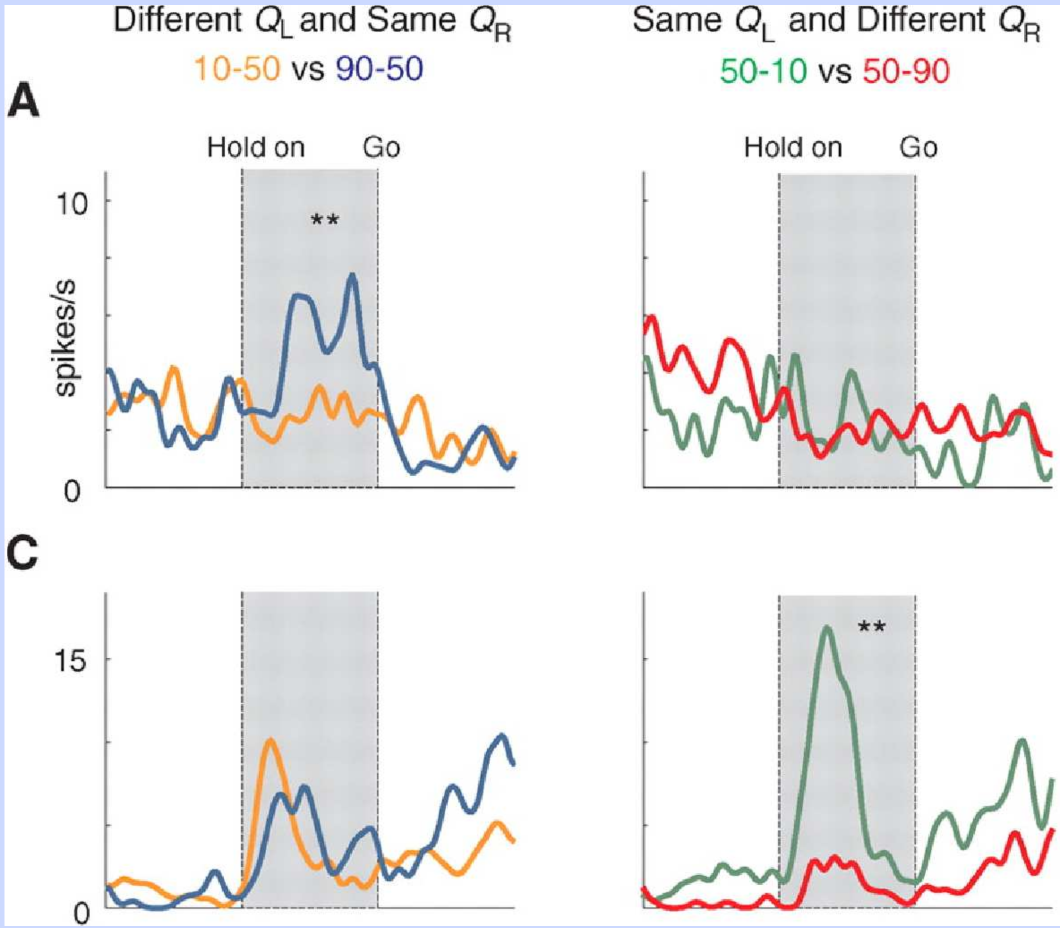
- α asymmetry leads to choose-A $\neq$ avoid-B (cf gene effects Frank et al 07)

Standard Q model



- α asymmetry leads to choose-A $\neq$ avoid-B (cf gene effects Frank et al 07)
- but no mechanism for motivational effects:
  increasing $\beta \rightarrow$ better overall accuracy, no pos/neg difference

Neural model and monkey data: separate populations coding for QGo and QNoGo



Samejima et al 2005

Q-GN model: separate QG from QN

Learning:

$$QG_i(t+1) = QG_i(t) + \alpha_{\mathbf{G}}[\delta(t)]$$

$$QN_i(t+1) = QN_i(t) + \alpha_{\mathbf{N}}[-\delta(t)]$$

Choice:

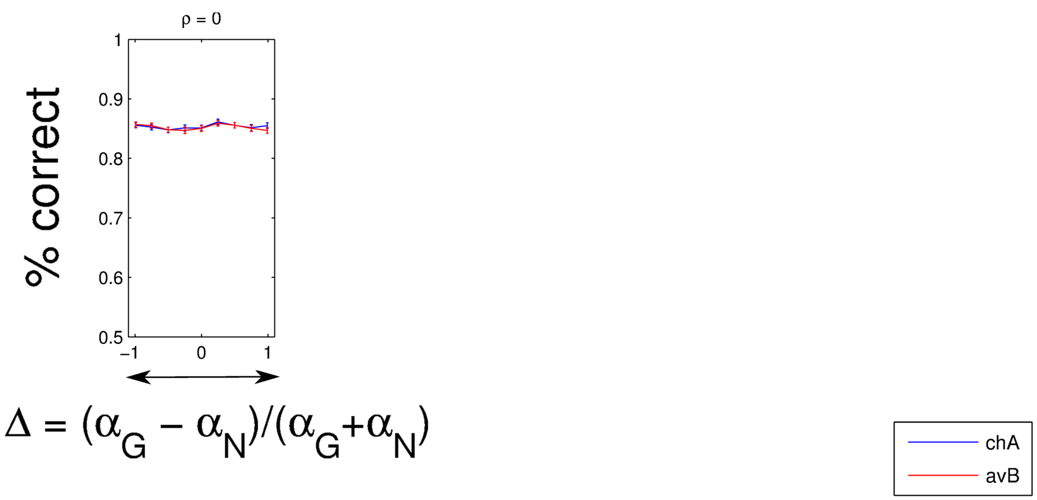
$$P_A = \frac{1}{1 + e^{-[\beta_{\mathbf{G}}(QG_A - QG_B) - \beta_{\mathbf{N}}(QN_A - QN_B)]}}$$

Dopamine modulates:

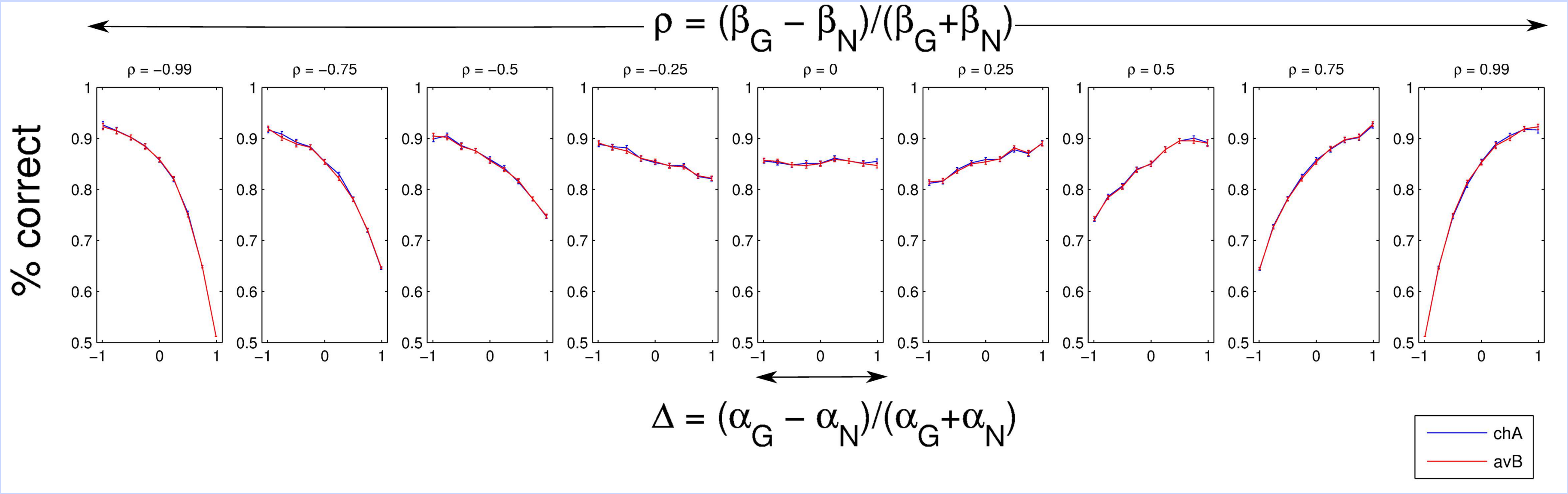
- learning (α asymmetry)
- incentive: QG vs QN during choice (β asymmetry)

choice function formally equivalent to prospect theory

QGN model

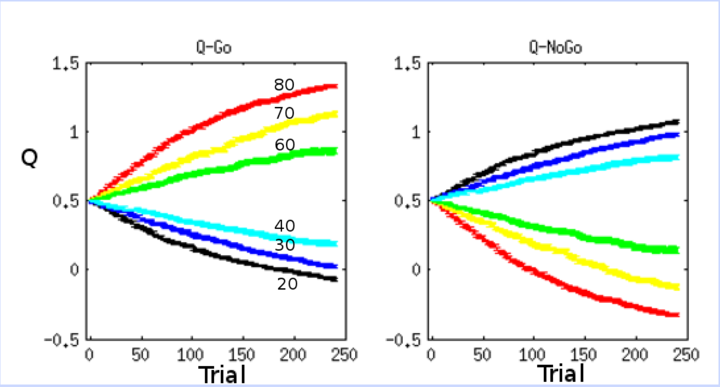


QGN model



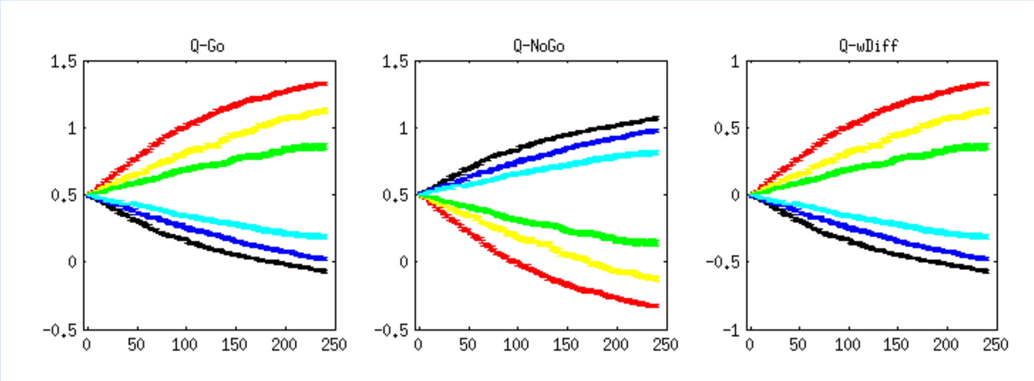
- no difference between choose-A / avoid-B for any param combination!
- better overall performance when learning and performing in the same dopaminergic state (α and β asymmetry in the same direction)

QGN model



QG **QN**

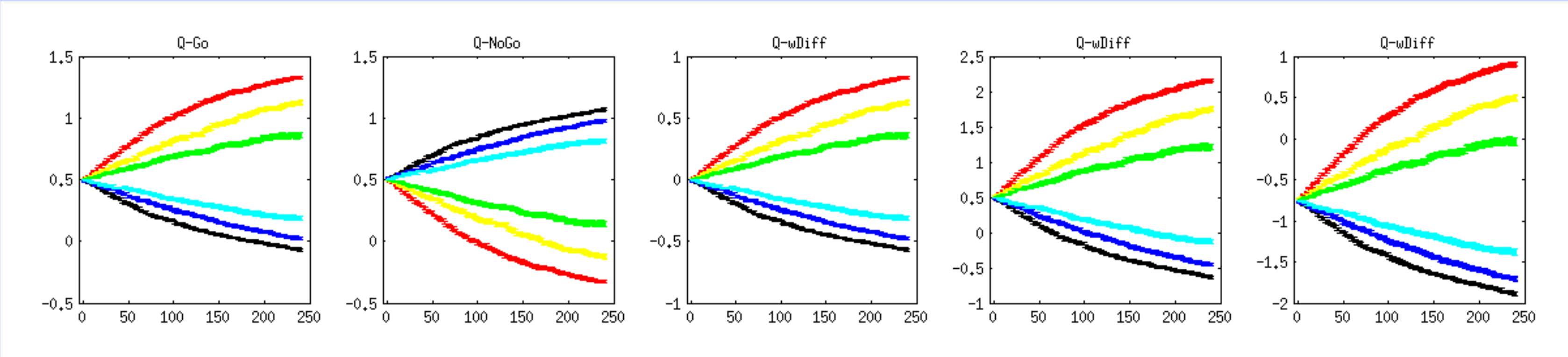
QGN model



QG **QN** **Q-wDiff**

$(\beta_G = \beta_N)$

QGN model



QG **QN** **Q-wDiff** ($\beta_G = \beta_N$) **Q-wDiff** ($\beta_G > \beta_N$) **Q-wDiff** ($\beta_G < \beta_N$)

Q-GN model with activity-modulated learning

Learning:

$$QG_i(t+1) = QG_i(t) + \alpha_G \mathbf{QG_i(t)}[\delta(\mathbf{t})]$$

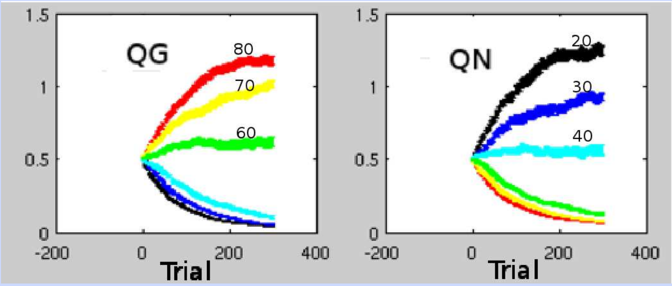
$$QN_i(t+1) = QN_i(t) + \alpha_N \mathbf{QN_i(t)}[-\delta(\mathbf{t})]$$

Q-GN model with activity-modulated learning

Learning:

$$QG_i(t+1) = QG_i(t) + \alpha_{\mathbf{G}} \mathbf{QG_i(t)}[\delta(\mathbf{t})]$$

$$QN_i(t+1) = QN_i(t) + \alpha_{\mathbf{N}} \mathbf{QN_i(t)}[-\delta(\mathbf{t})]$$

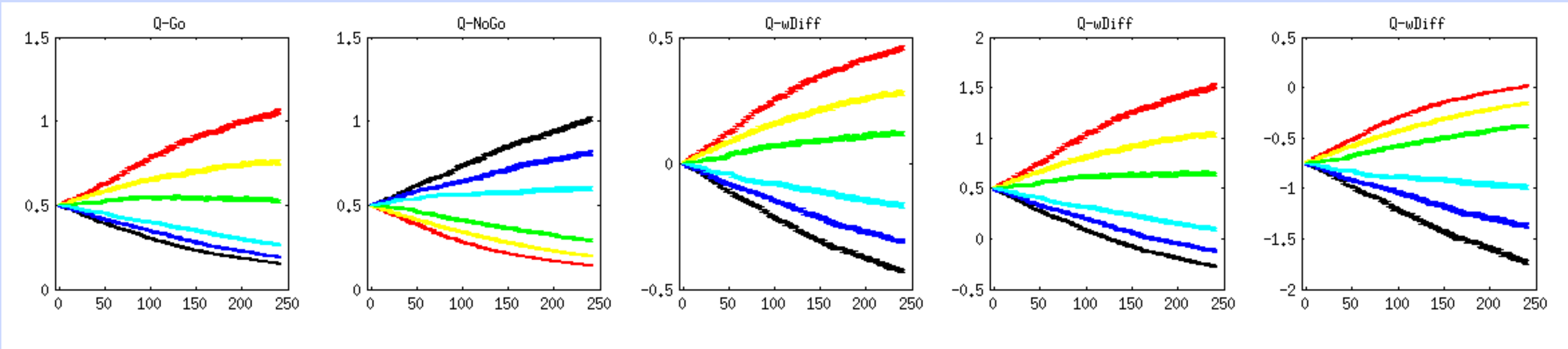


Choice:

$$P_A = \frac{1}{1 + e^{-[\beta_{\mathbf{G}}(QG_A - QG_B)\ -\ \beta_{\mathbf{N}}(QN_A - QN_B)]}}$$

with Anne Collins

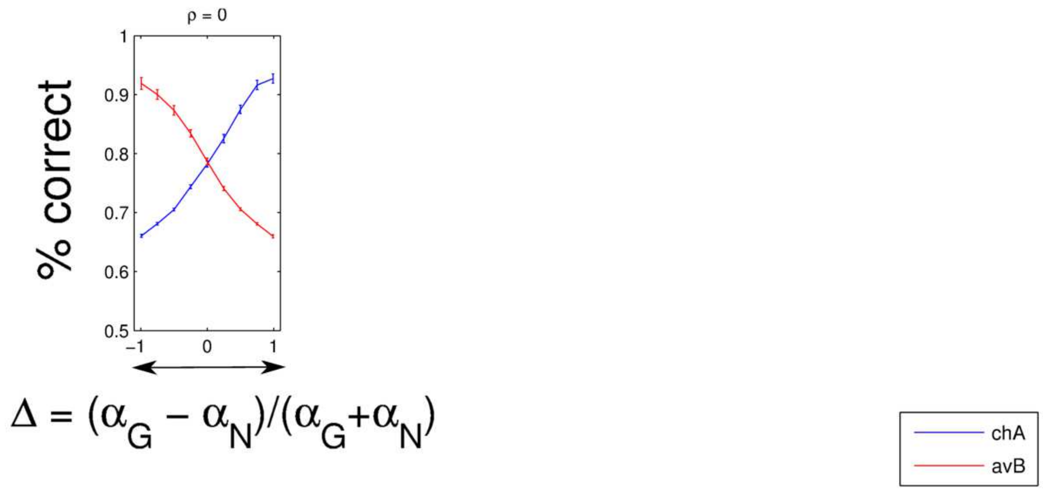
QGN model with activity-modulated learning



QG	QN	Q-wDiff	Q-wDiff	Q-wDiff
		$(\beta_G = \beta_N)$	$(\beta_G > \beta_N)$	$(\beta_G < \beta_N)$

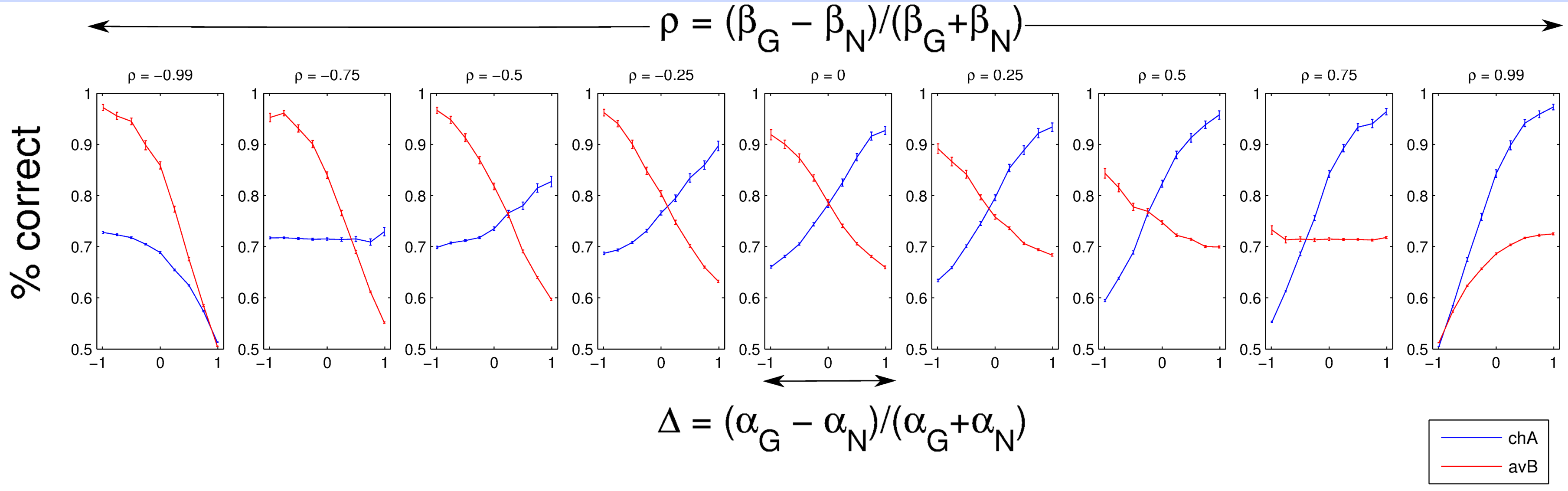
with Anne Collins

QGN model: learning and motivational effects

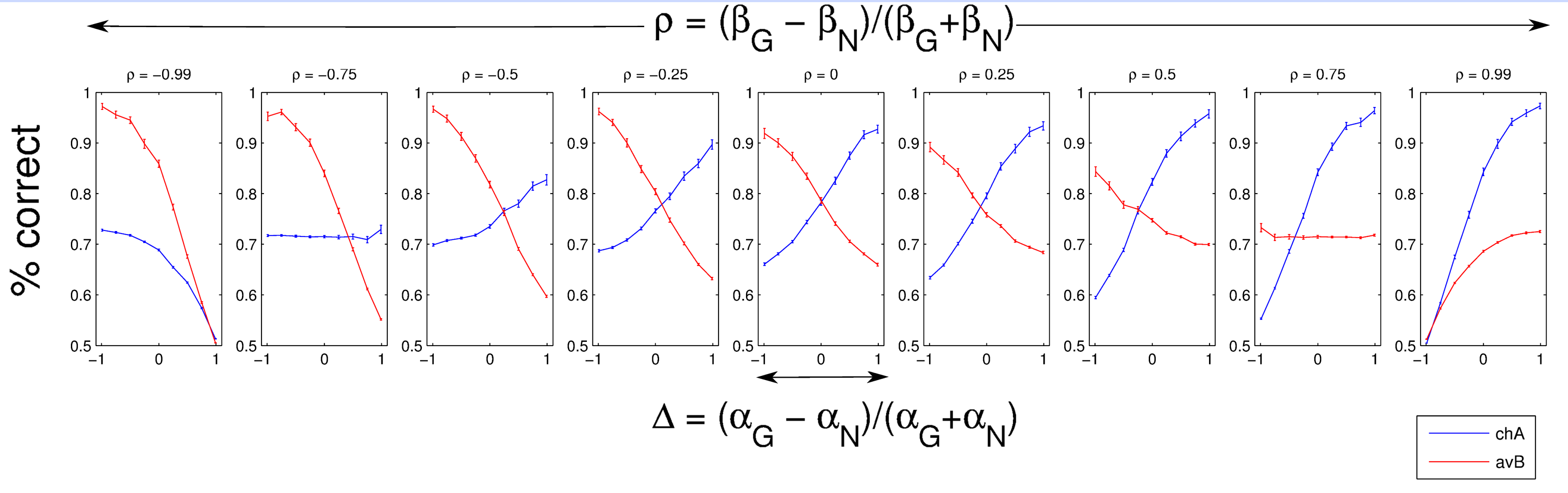


with Anne Collins

QGN model: learning and motivational effects



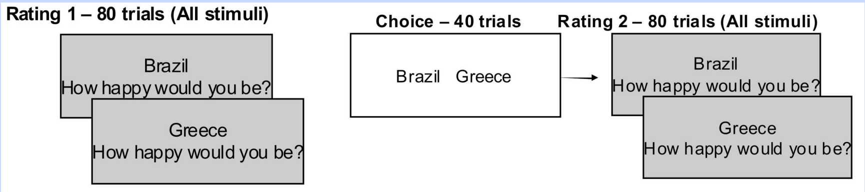
QGN model: learning and motivational effects

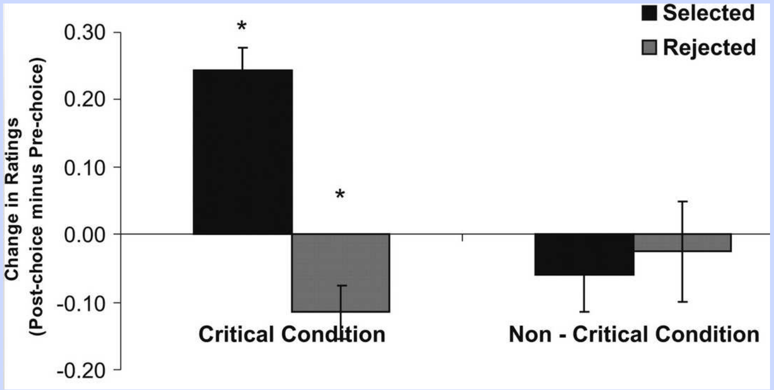


⇒ *Learning and motivational effects of striatal DA arise only if the two processes interact!*

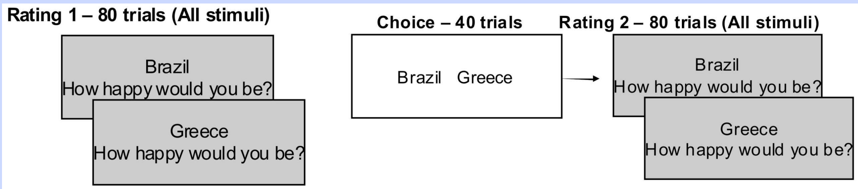
with Anne Collins

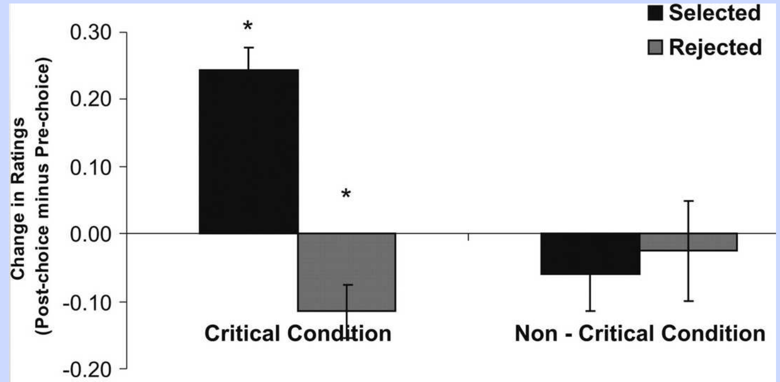
Valuation increases after choice…



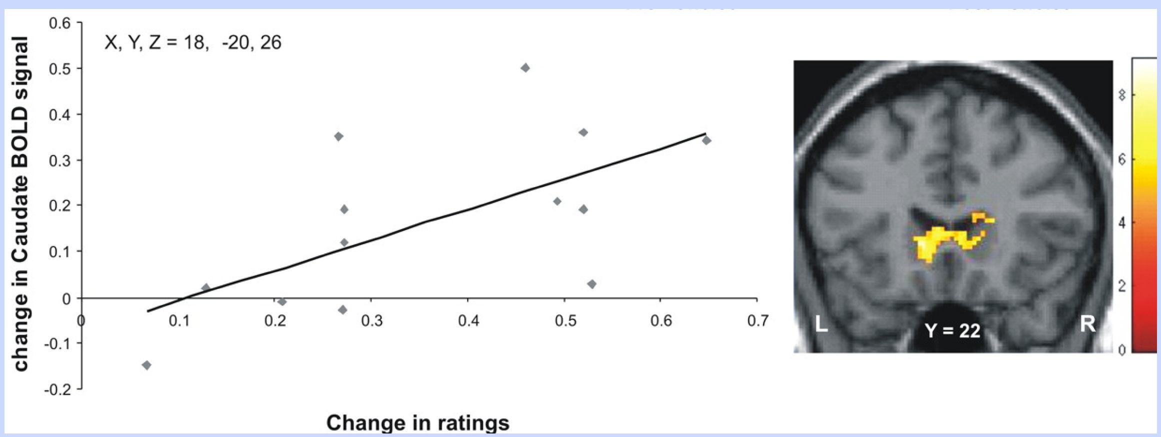


Valuation increases after choice...



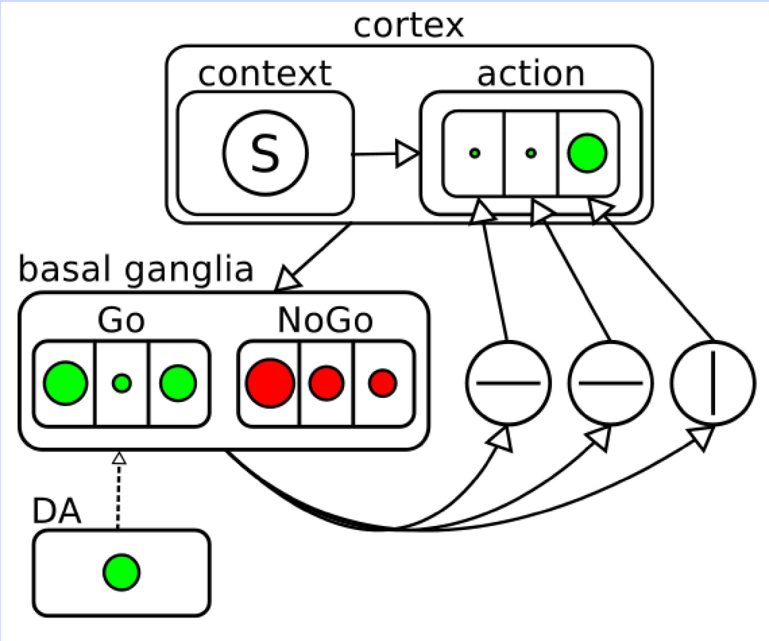


...this enhanced valuation correlates with BG activity

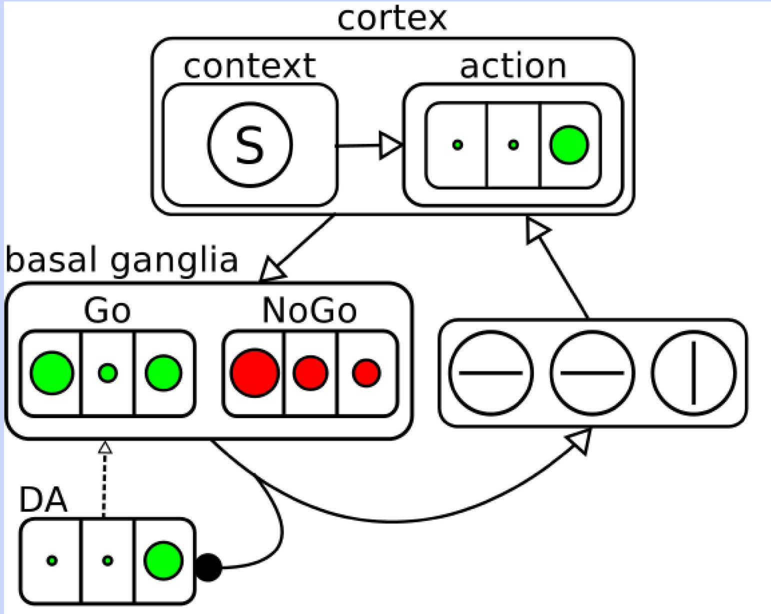


Sharot et al, 2009a

Why might this happen:
The (Structural) Credit Assignment Problem

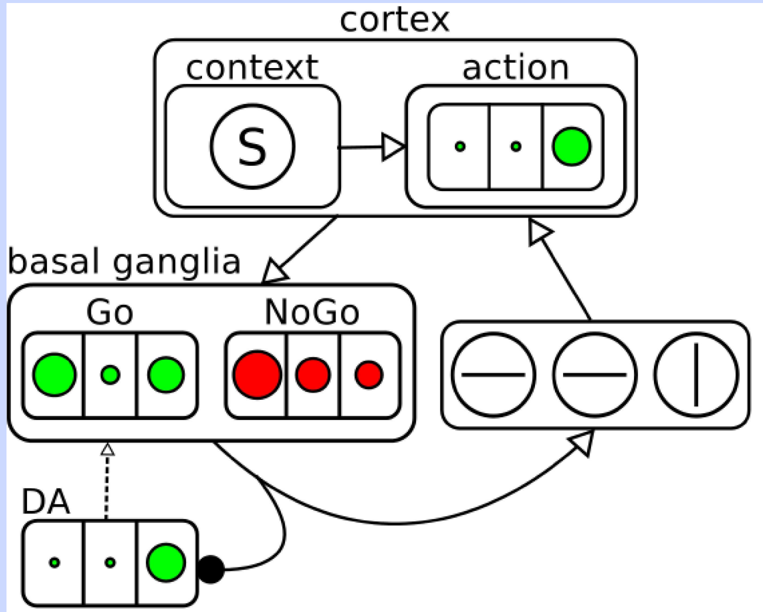


A solution



This mechanism substantially improves learning in advanced tasks; O'Reilly & Frank 2006

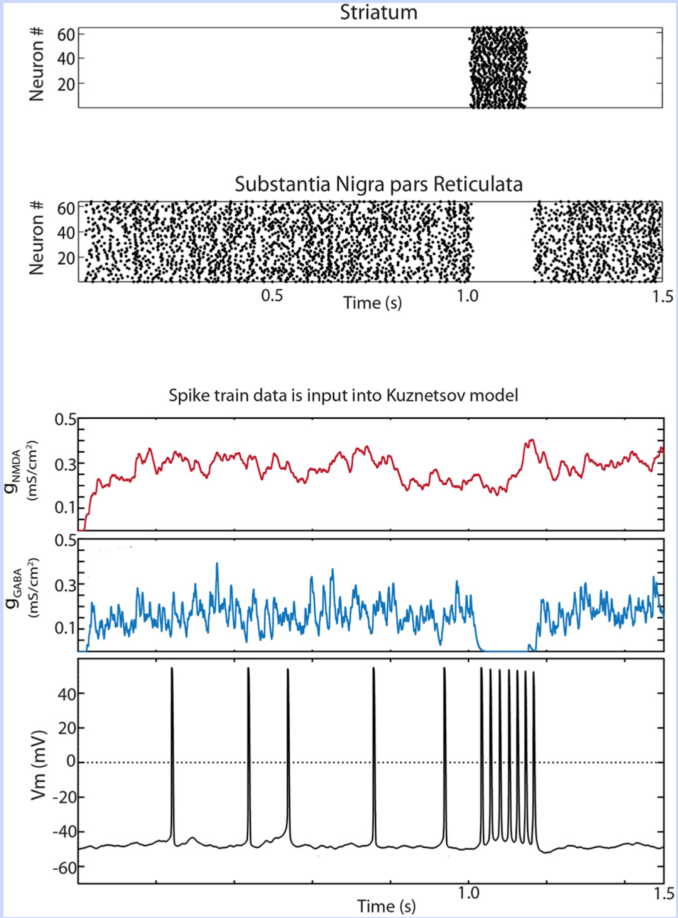
A solution



This mechanism substantially improves learning in advanced tasks; O'Reilly & Frank 2006

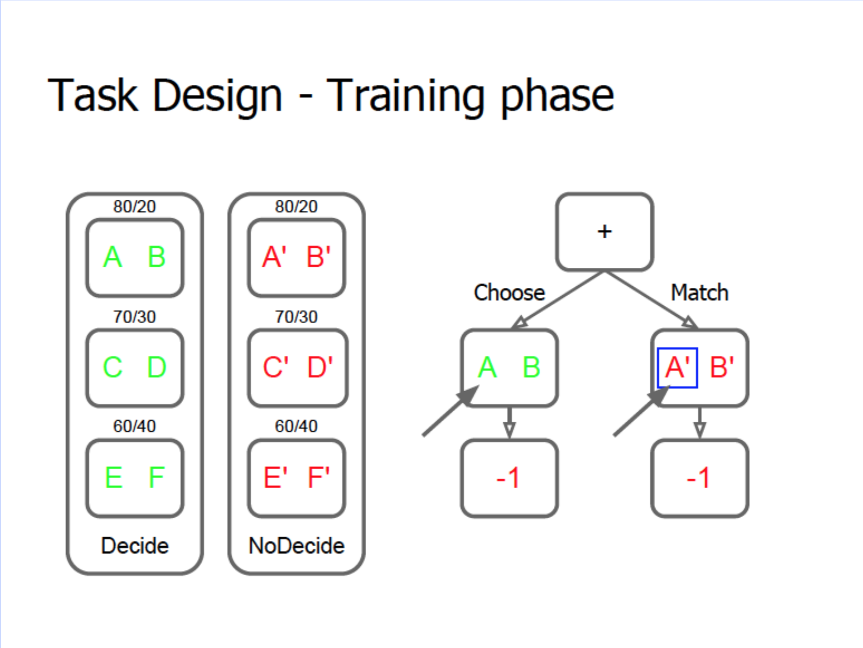
Some evidence:
Independent and simultaneous RPE signals in BG for actions with independent outcomes (Gershman et al 09)

The Mechanism: Disinhibition

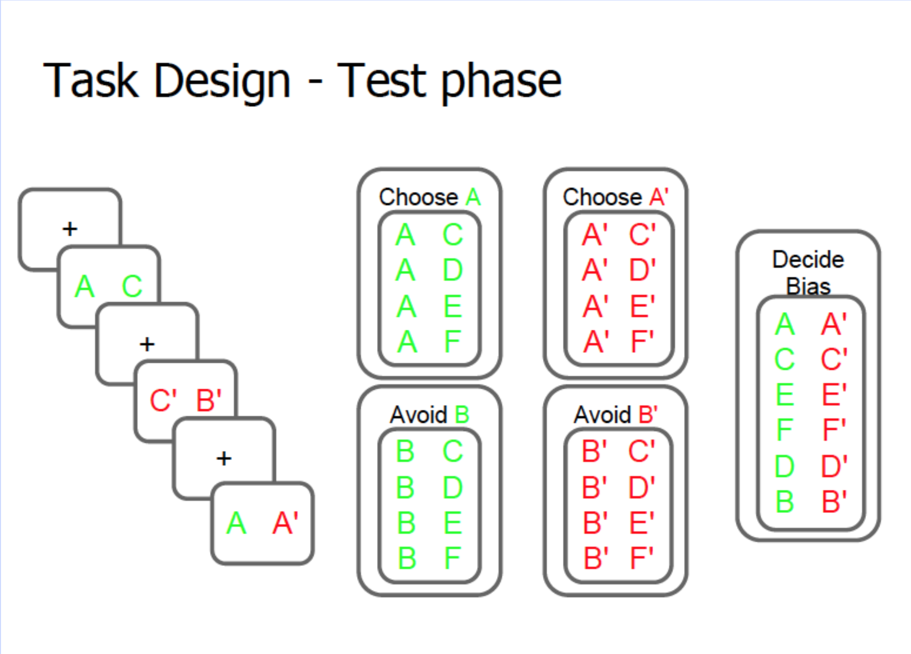


Joel & Weiner, ’00; O’Reilly & Frank ’06; Lobb et al, 2011

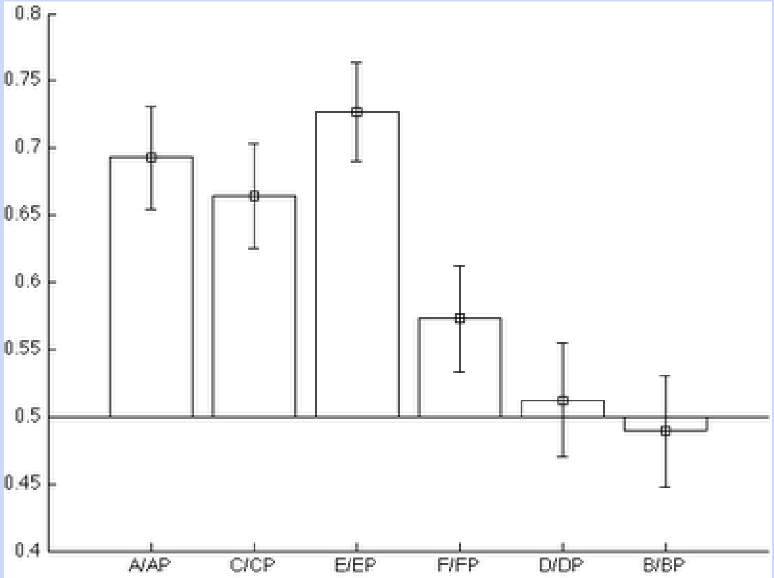
Quantifying Choice Bias in RL



Quantifying Choice Bias in RL

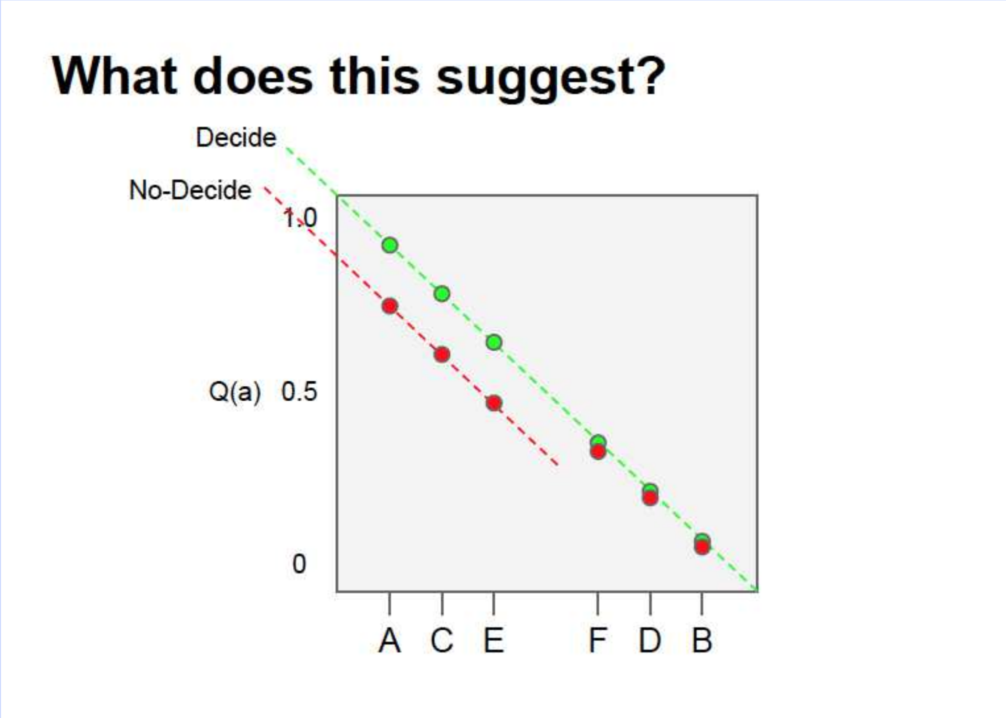


Choice bias results



- choice-bonus only for options that yielded positive outcomes
- consistent with actions disinhibiting DA bursts

In Q value terms...

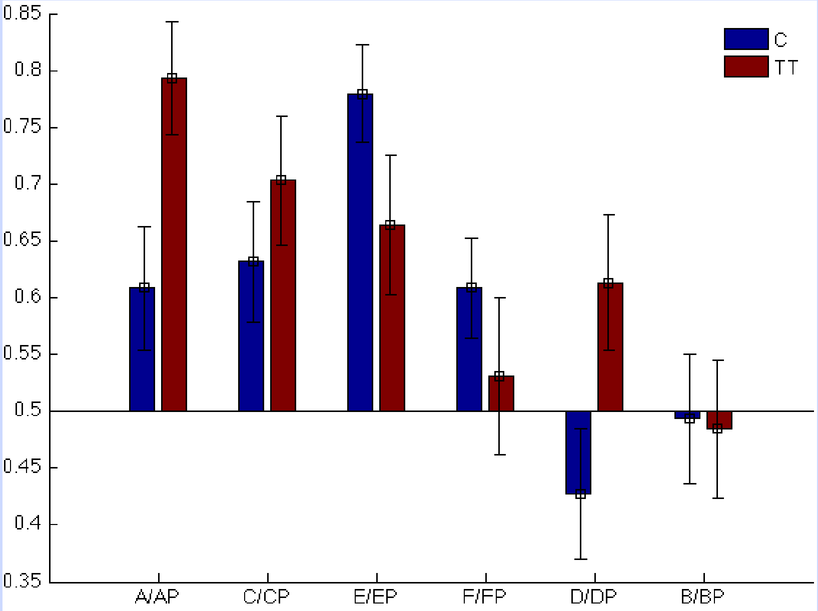


Individual differences are informative! DARPP-32 gene

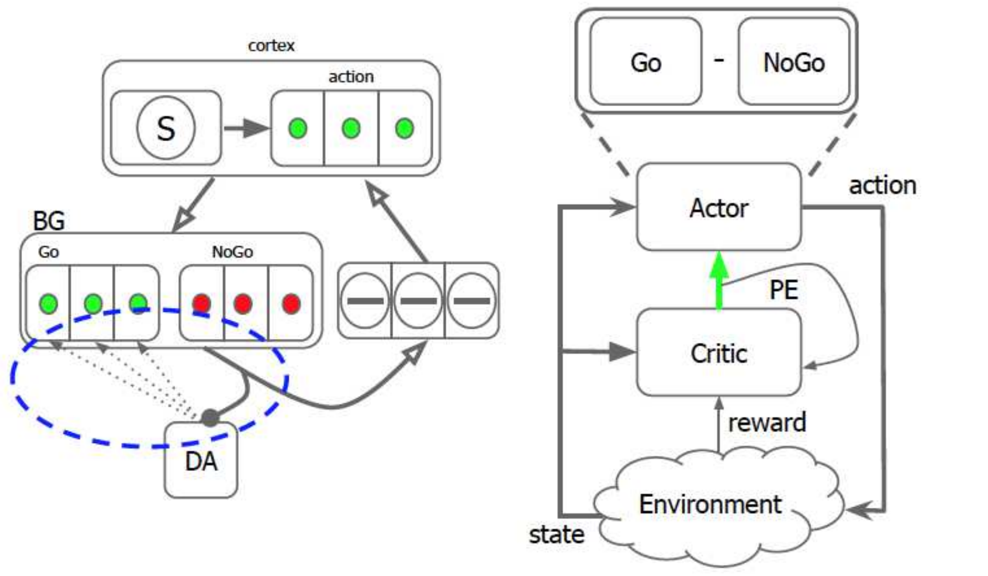
If bias arises from PE-bonus & Go learning, should be enhanced in DARPP-32 TT carriers...

Individual differences are informative! DARPP-32 gene

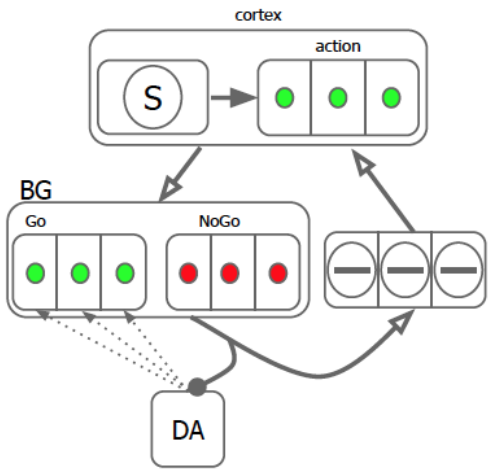
If bias arises from PE-bonus & Go learning, should be enhanced in DARPP-32 TT carriers...



Modelling the Mechanism



Modelling the Mechanism



Learning

$PE = r - V(s)$

Decide & $PE > 0$
$PE = PE * boost$

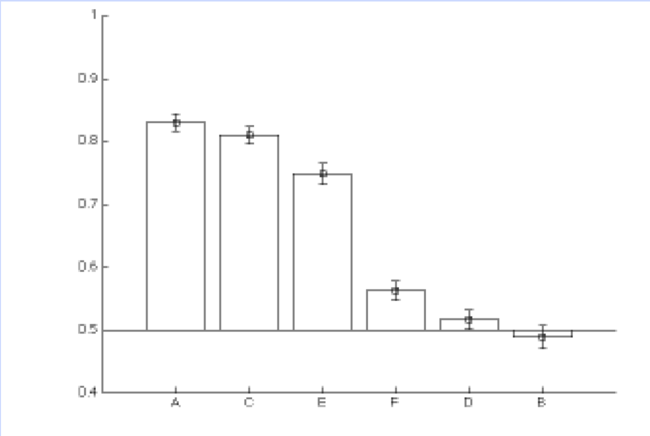
$V(s) = V(s) + a * PE$
$Q_G(a) = Q_G(a) + a_G * PE * Q_G(a)$
$Q_{NG}(a) = Q_{NG}(a) + a_{NG} * PE * Q_{NG}(a)$

Action Selection

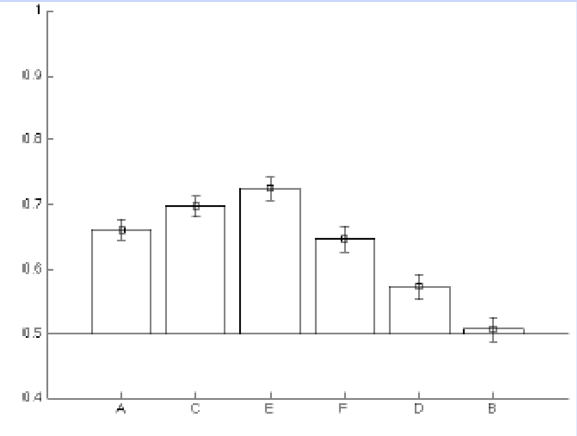
$Q = Q_G * B_G - Q_{NG} * B_{NG}$

$\mathrm{Softmax}(Q(a_1, a_2))$

Generative QGN: choice bias varies by learning asymmetry

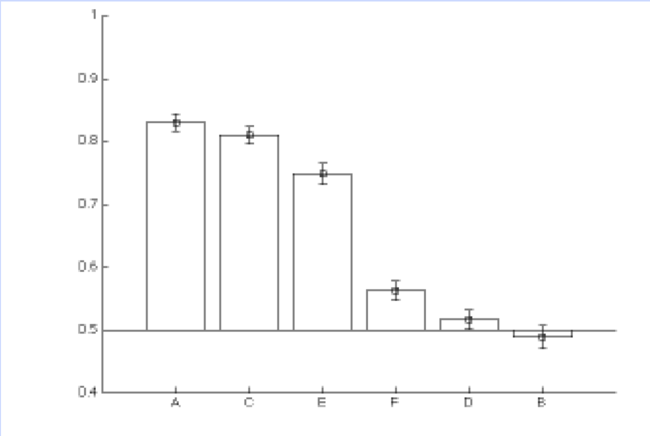


$\alpha_G > \alpha_N$

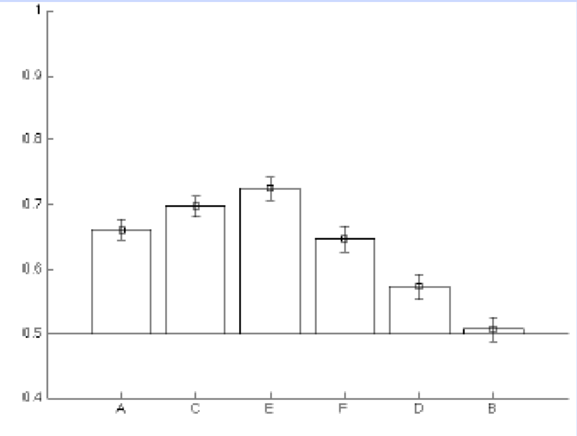


$\alpha_G < \alpha_N$

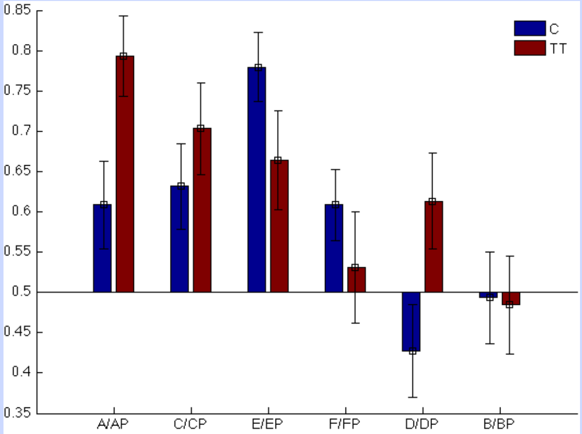
Generative QGN: choice bias varies by learning asymmetry



$\alpha_G > \alpha_N$

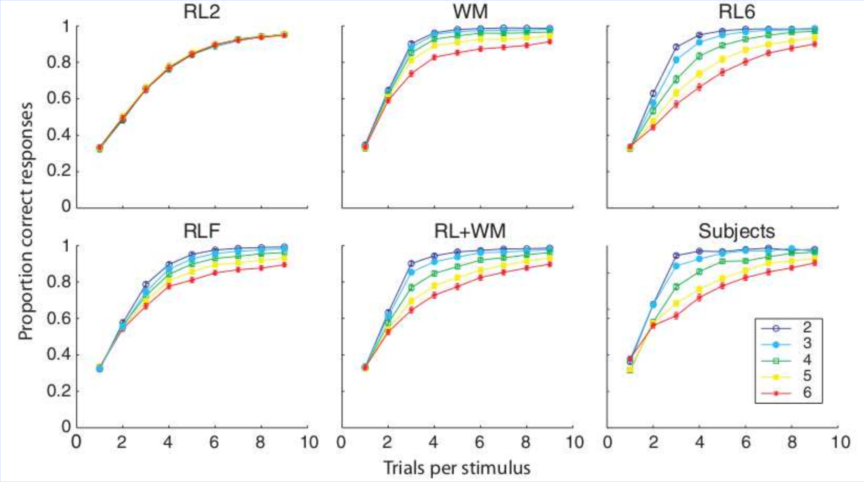


$\alpha_G < \alpha_N$

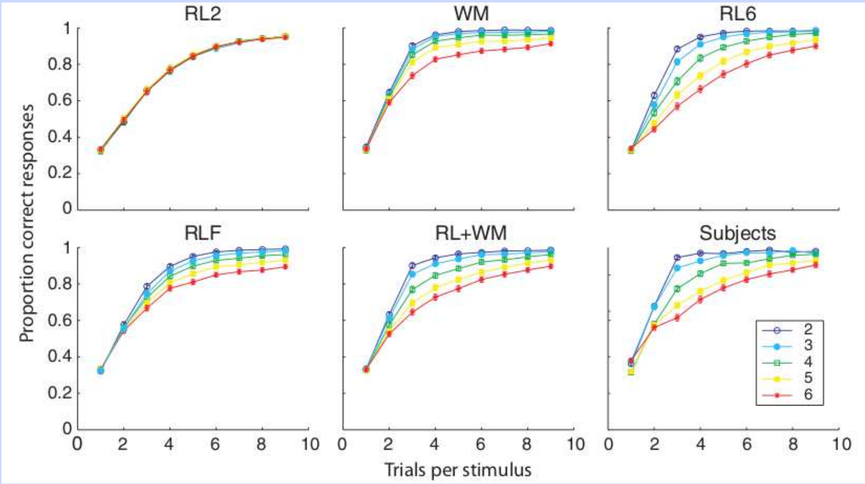


"Adaptive" mechanism for credit assignment can lead to odd behavior...

Working memory effects on RL



Working memory effects on RL

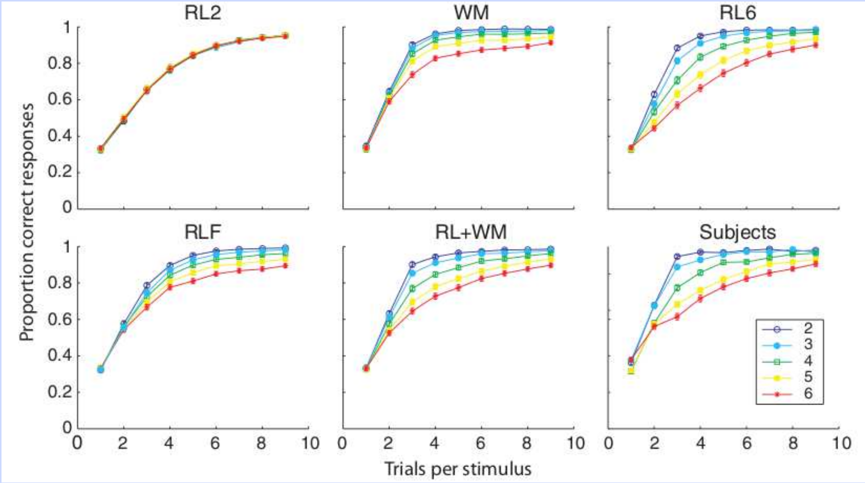


RL+WM model:

$$p(a) = (1 - w(t)) * p_{RL}(a) + w(t) * p_{WM}(a)$$

$$w_{n_S}(t+1, s) = \frac{p_{WMC}(r_t|s_t, a_t) w_{n_S}(t, s)}{p_{WMC}(r_t|s_t, a_t) w_{n_S}(t, s) + p_{RL}(r_t|s_t, a_t)(1 - w_{n_S}(t, s))}.$$

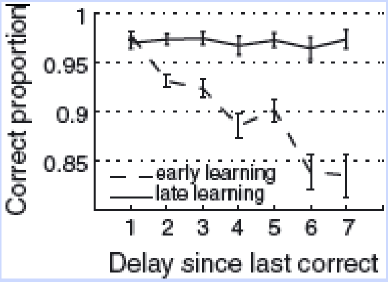
Working memory effects on RL



RL+WM model:

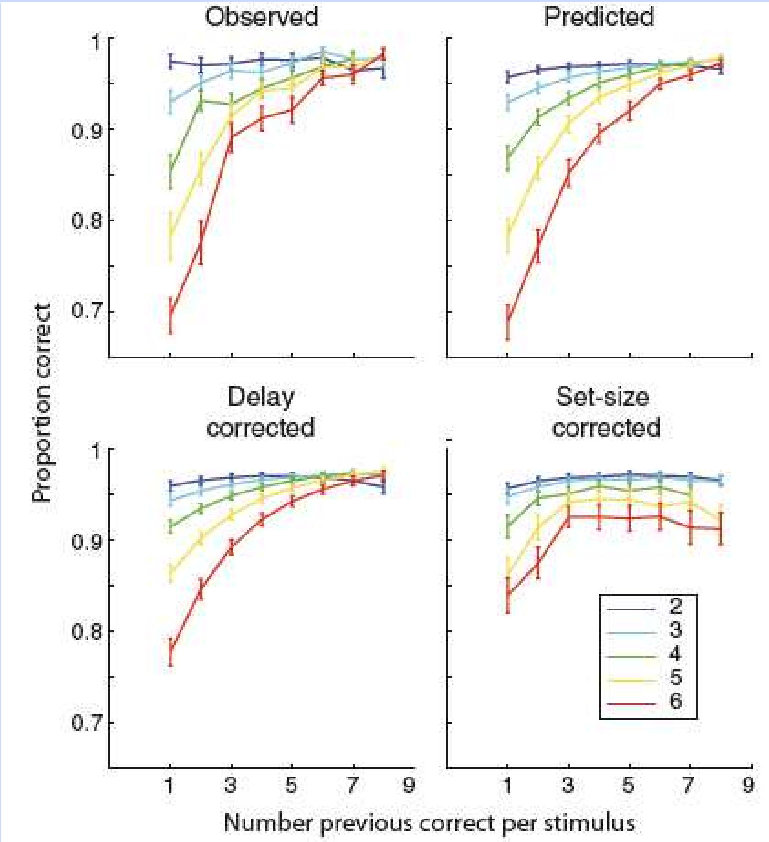
$$p(a) = (1 - w(t)) * p_{RL}(a) + w(t) * p_{WM}(a)$$

$$w_{n_S}(t+1, s) = \frac{p_{WMC}(r_t|s_t, a_t) w_{n_S}(t, s)}{p_{WMC}(r_t|s_t, a_t) w_{n_S}(t, s) + p_{RL}(r_t|s_t, a_t)(1 - w_{n_S}(t, s))}.$$



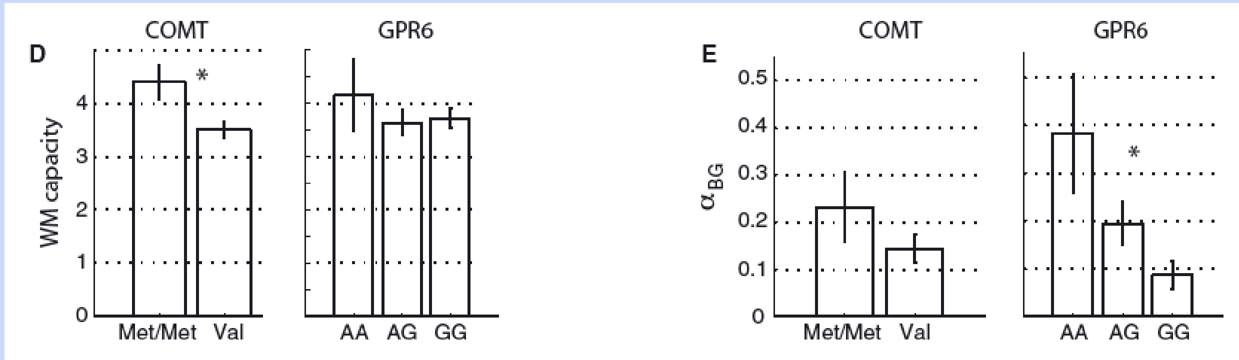
Collins & Frank 2012 EJN

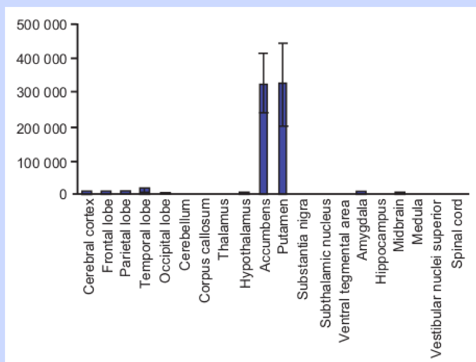
Two features of WM contributions: delay and capacity limits

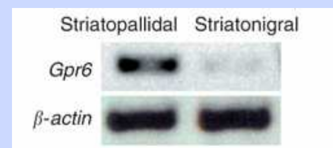


Collins & Frank 2012 EJN

Dissociable PFC vs BG genetic predictors of WM vs RL



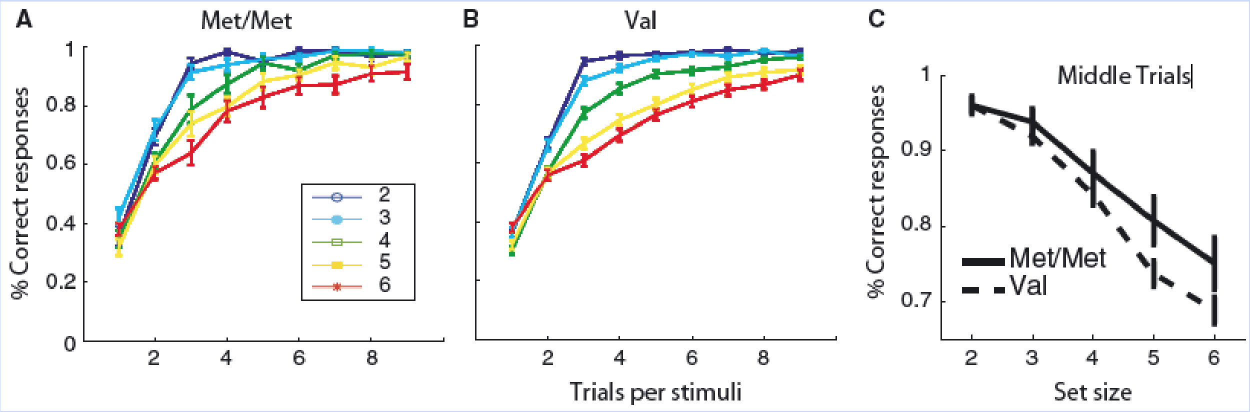




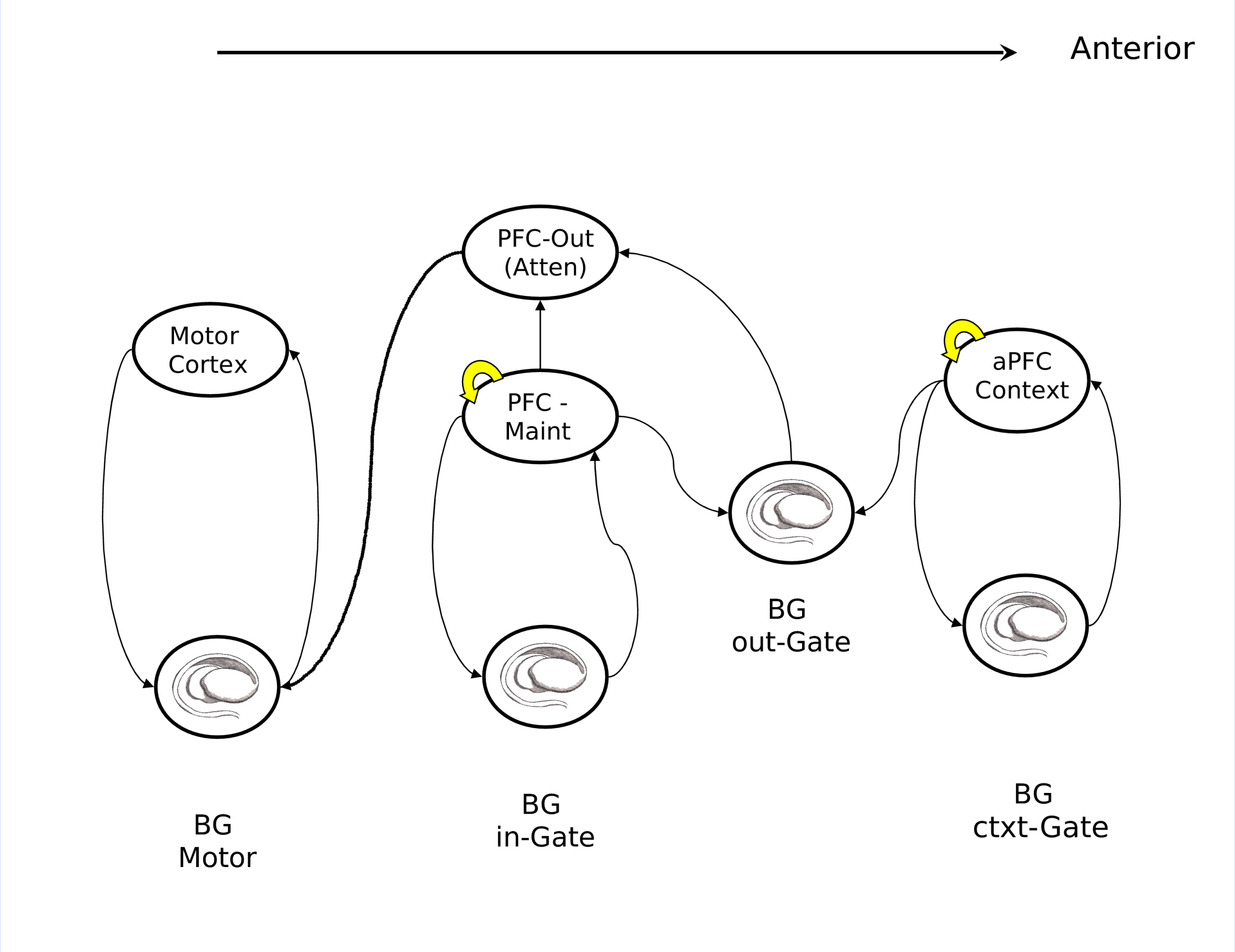
GPR6: postmortem human (Roth et al 06)

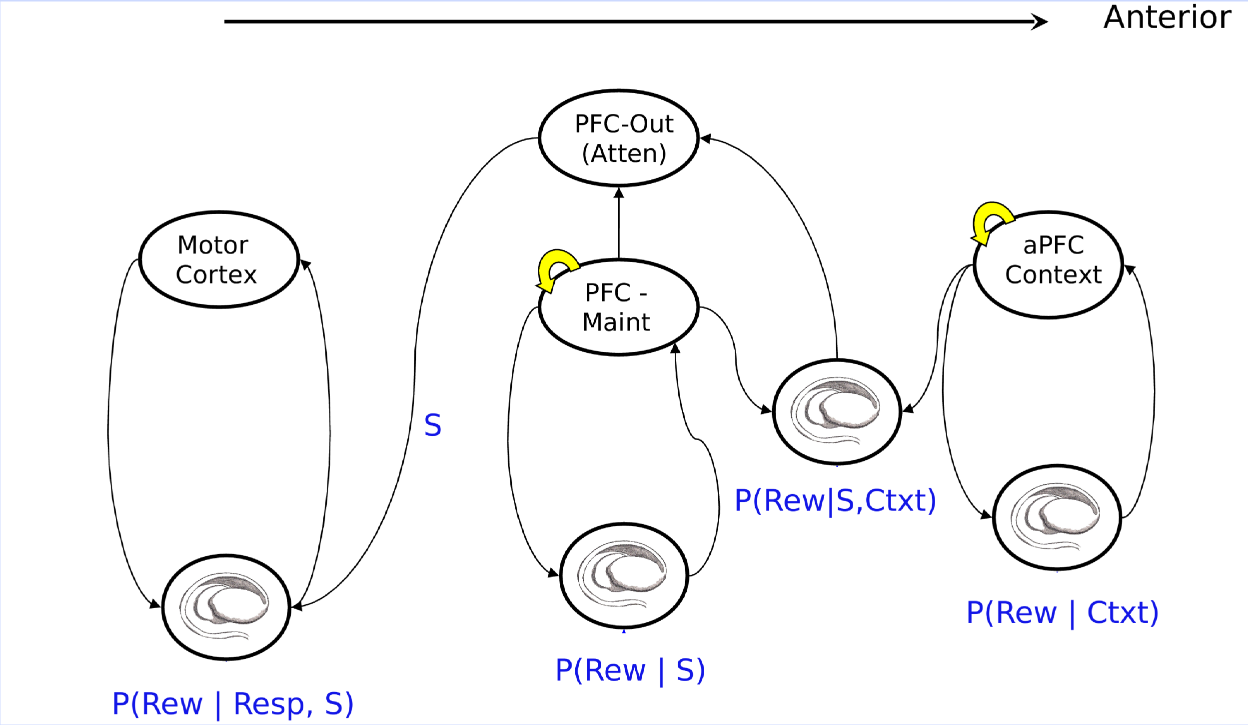
mouse (Lobo et al 07)

Closer look at COMT WM effects on learning



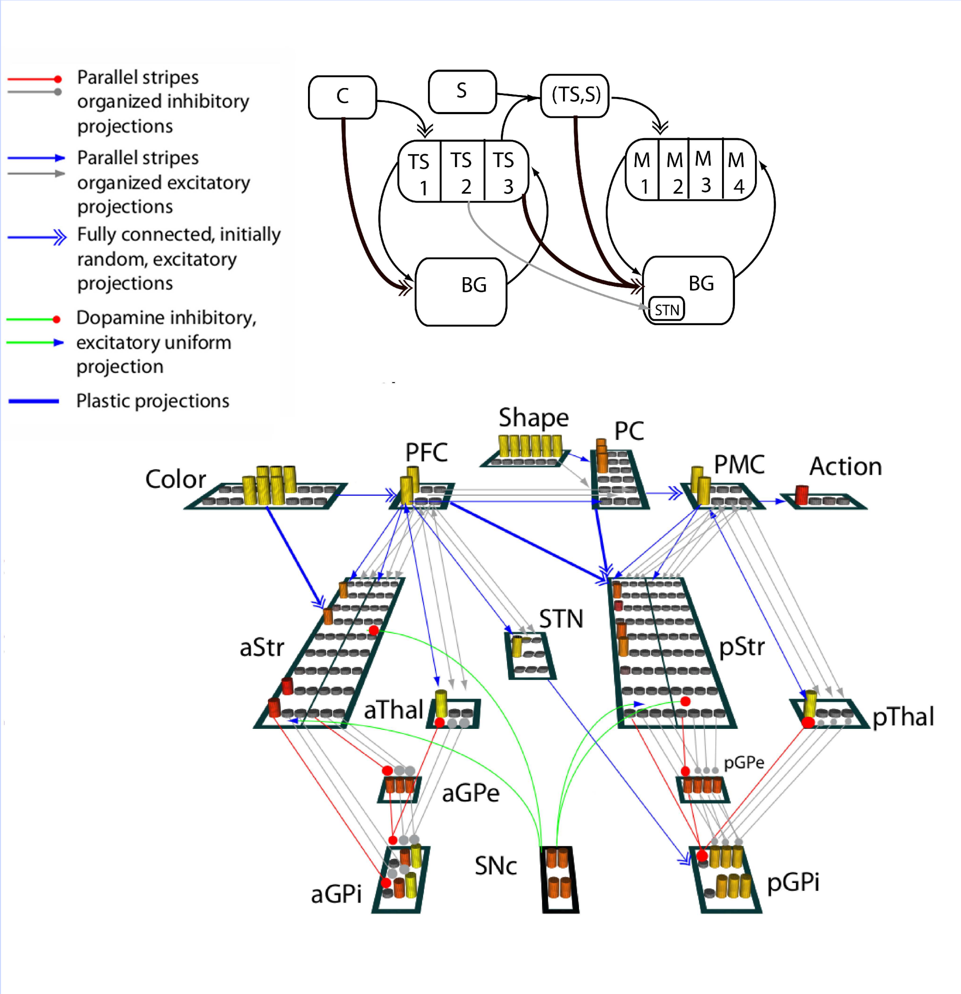
Hierarchical interactions among BG-FC circuits





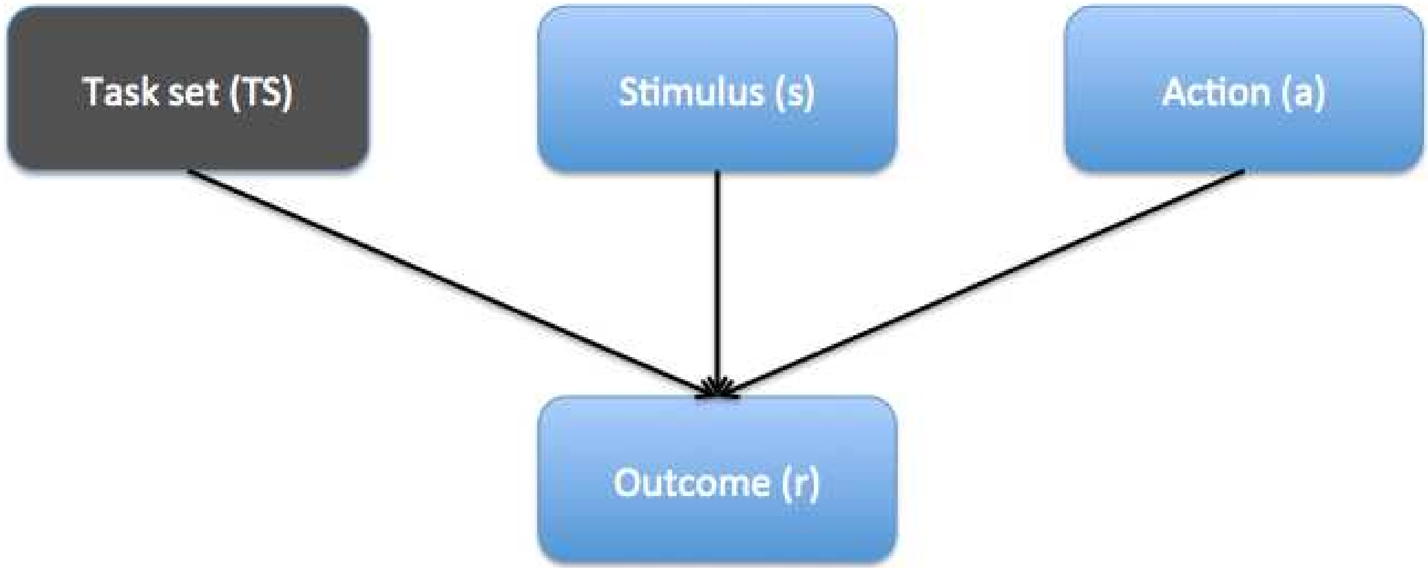
Frank & Badre, 2011; Badre & Frank 2011

Hierarchical interactions among BG-FC circuits



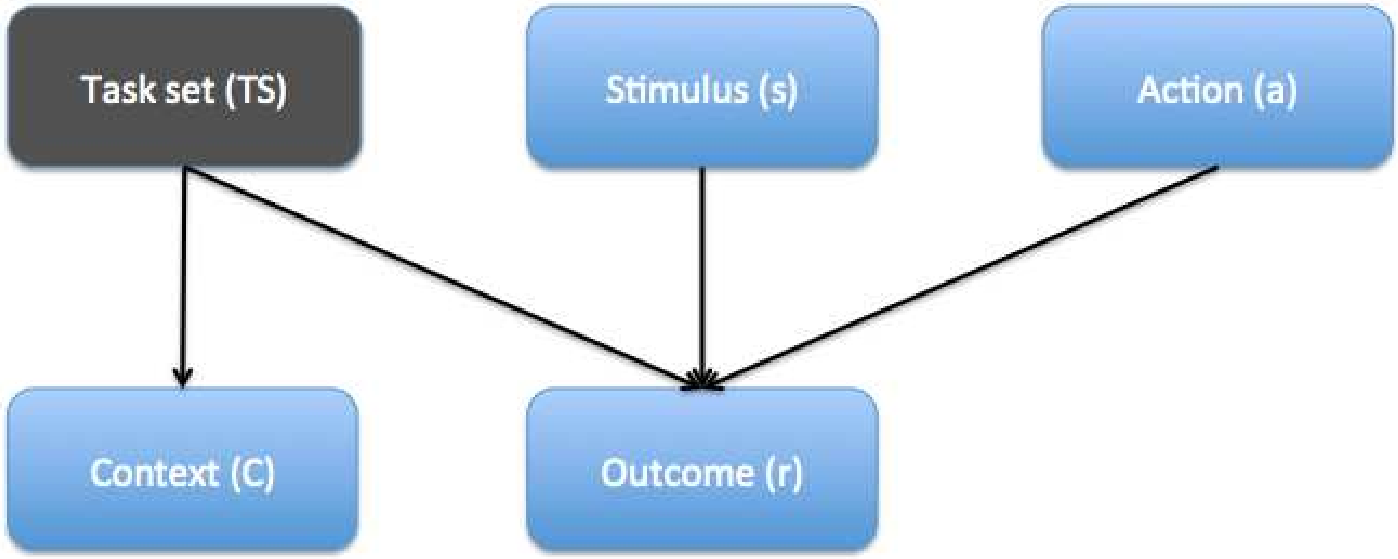
Collins & Frank submitted; Frank & Badre, 2011

Framework/Questions

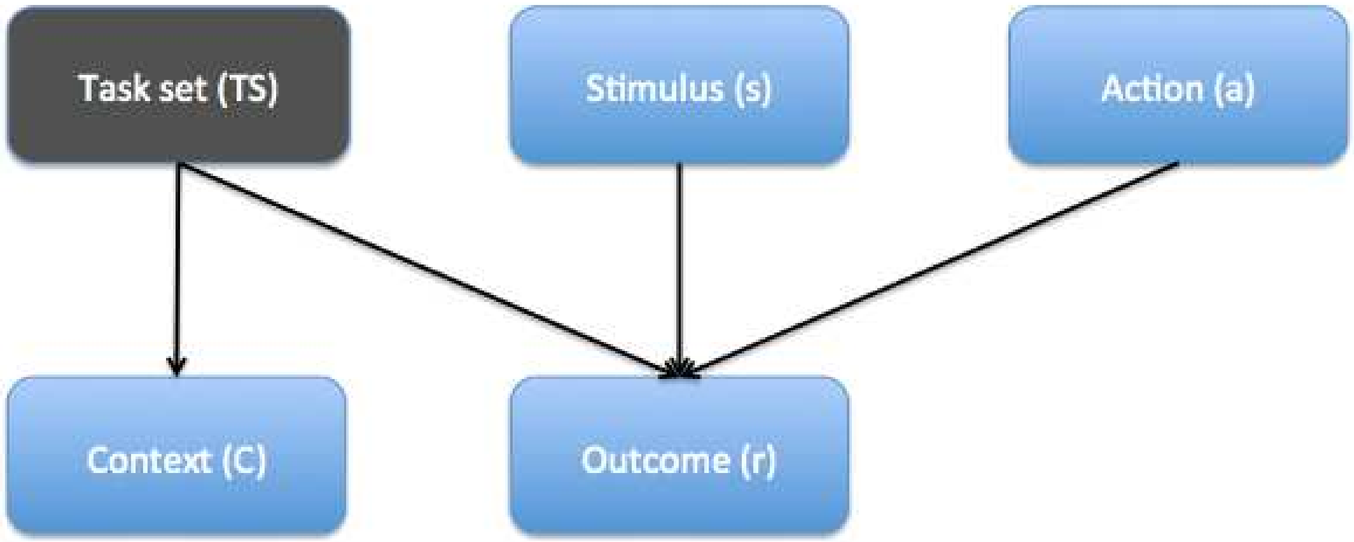


How are task sets learned?

Framework/Questions



Model goals

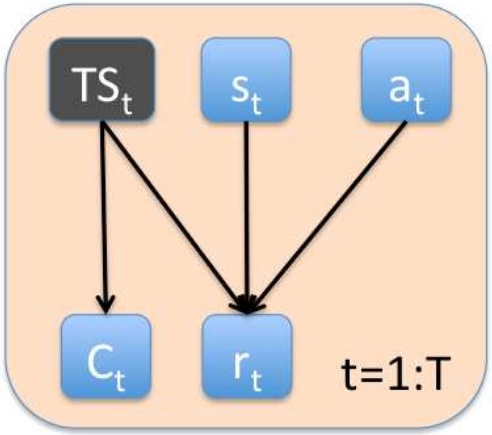


How is task set space built?

How are the task sets learned?

How is the current task set inferred?

C-TS model

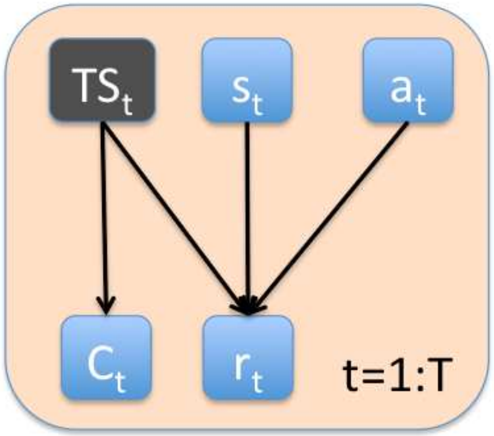


Goal: learn to infer TS_t

Problem: what is the task set space?

Exact Bayesian inference
Intractable

C-TS model

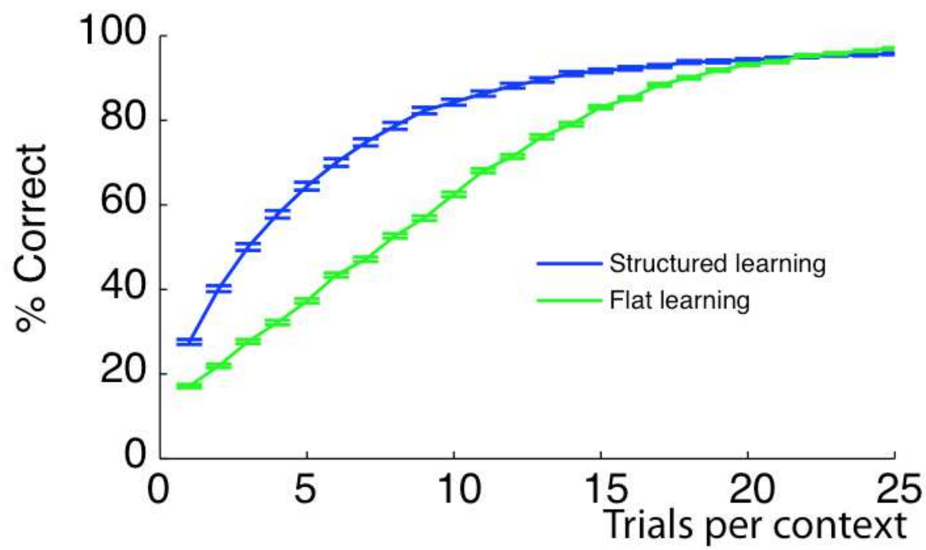


Simplifying assumptions:
Maximum a posteriori learning

$$P(TS|C_t,s_t,a_t,r_t) \propto P(r_t|s_t, a_t, TS)P(TS|C_t)$$

Learning of $P(r_t|s_t,a_t,TS)$ occurs
only for the most likely TS *(a post.)*

C-TS clustering: results



As expected, model predicts gains from structure

2 loop PFC-BG network maps onto C-TS abstraction (Collins & Frank, pending)

Summary

- Dynamic interaction between performance and learning effects of striatal dopamine: rodents, humans, models
- Learning from internally/externally generated action dissociates BG gating from action. Boosted values assoc with rewarding actions
- Dissociable contributions of PFC and BG to RL / WM
- Negative symptoms in schizophrenia: degraded PFC value representations for choice and exploration
- PFC-STN pathway regulates decision threshold
- Hierarchical interactions between multiple loops enable structured model learning

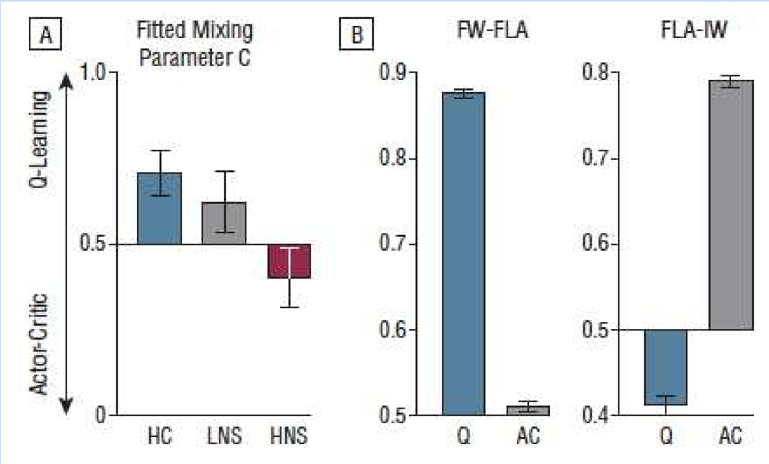
Thanks To...

Jeff Cockburn
Jeff Beeler
Anne Collins
Bradley Doll
Thomas Wiecki
Jim Cavanagh
Christina Figueroa
Sean Masters
Scott Sherman, MD
Kent Hutchison
Xiaoxi Zhuang

PD patients

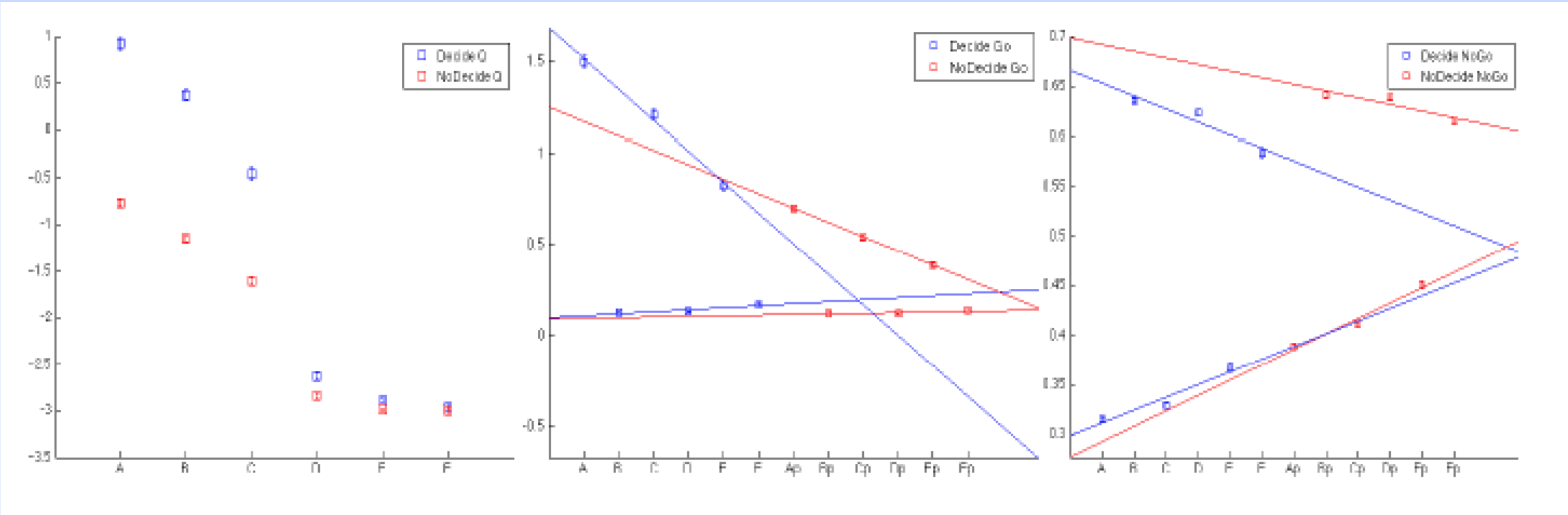
Lab for Neural Computation and Cognition

Negative symptoms in schizophrenia: actor critic vs Q learning



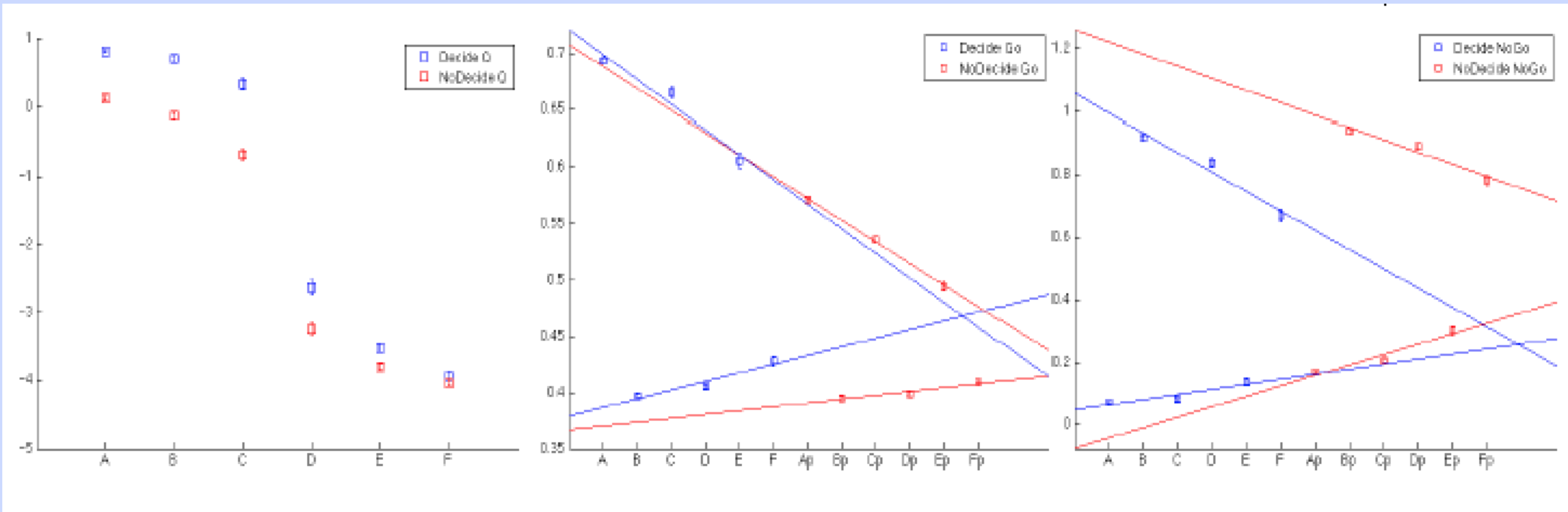
Gold et al, 2012 Archives of Gen Psych

QG and QN values: Decide and No-Decide ($\alpha_G > \alpha_N$)



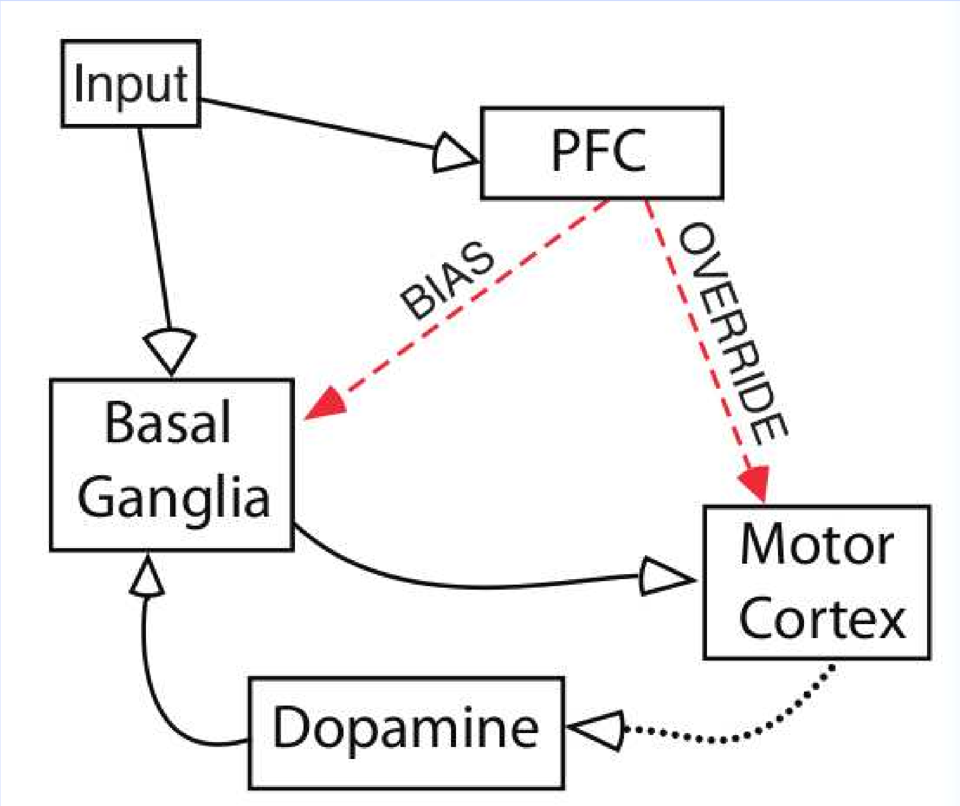
Q_{net} Q_G Q_N

QG and QN values: Decide and No-Decide ($\alpha_G < \alpha_N$)



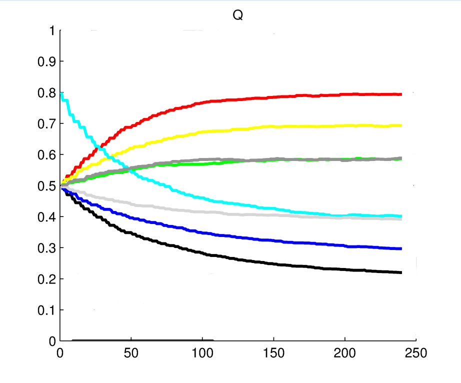
Q_{net} Q_G Q_N

Top-down activity can also influence learning: PFC instructional influence on BG

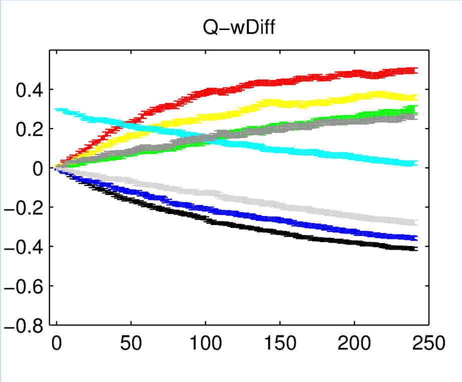


Doll et al, 2009

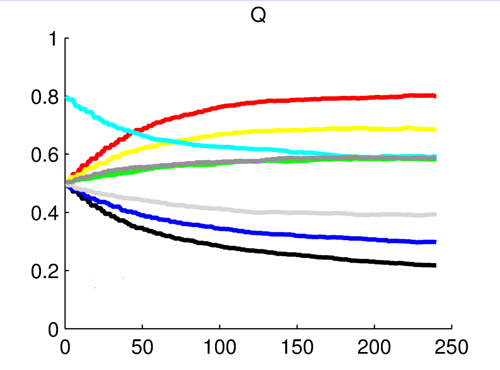
QGN model predicts distorted learning: confirmation bias



standard Q

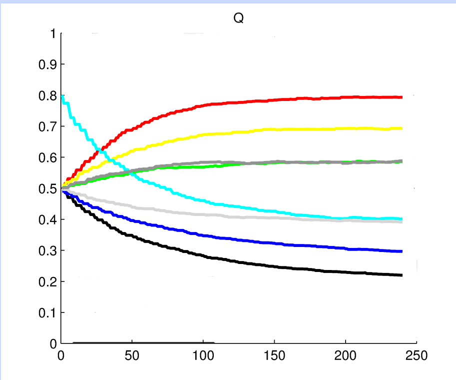


QGN

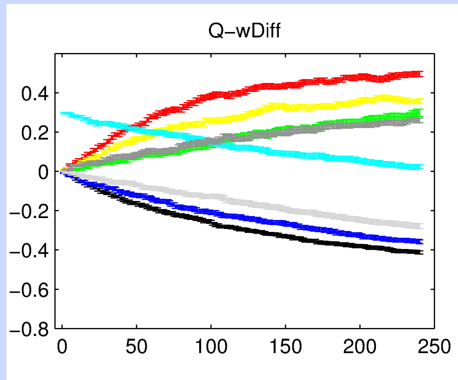


Q + conf bias

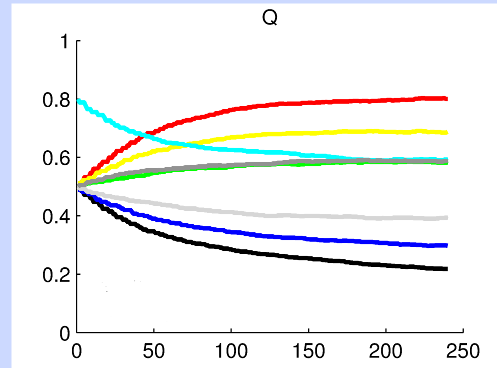
QGN model predicts distorted learning: confirmation bias



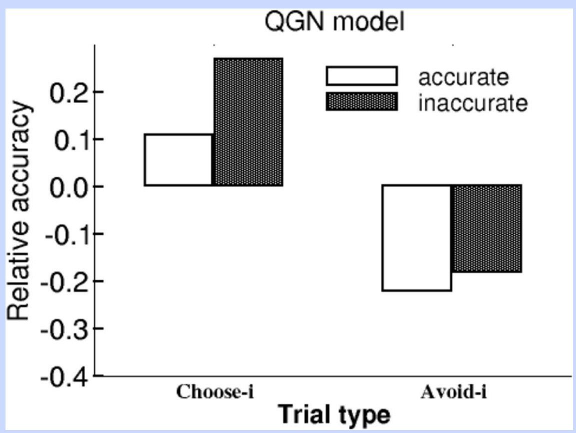
standard Q



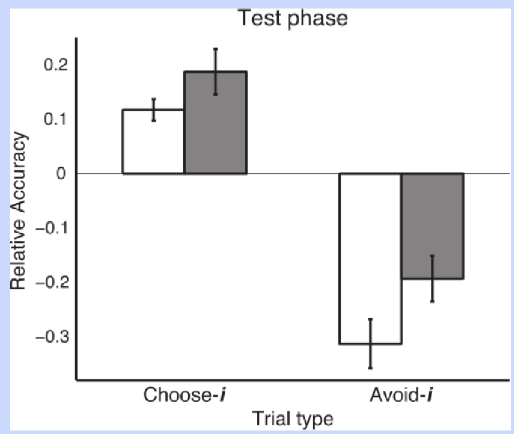
QGN



Q + conf bias

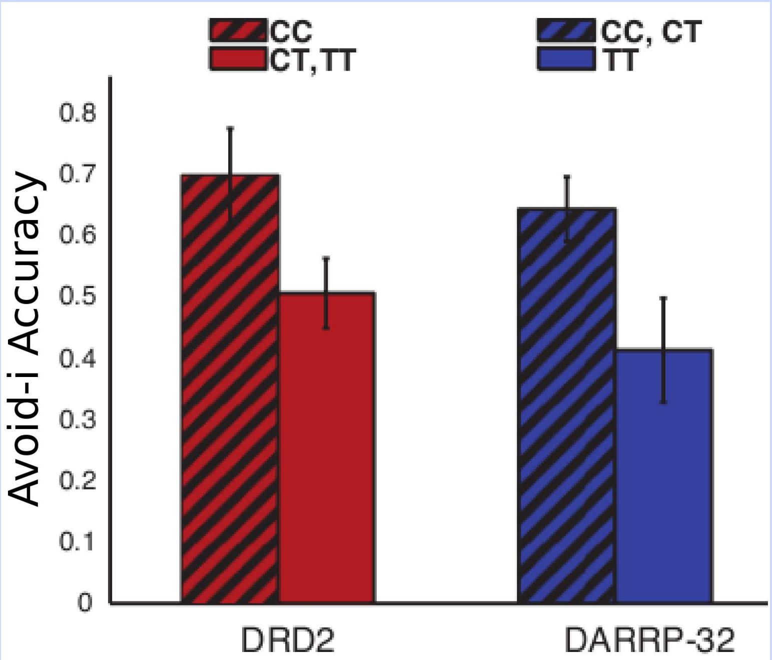


model



data

Genetic modulation of instruction effects

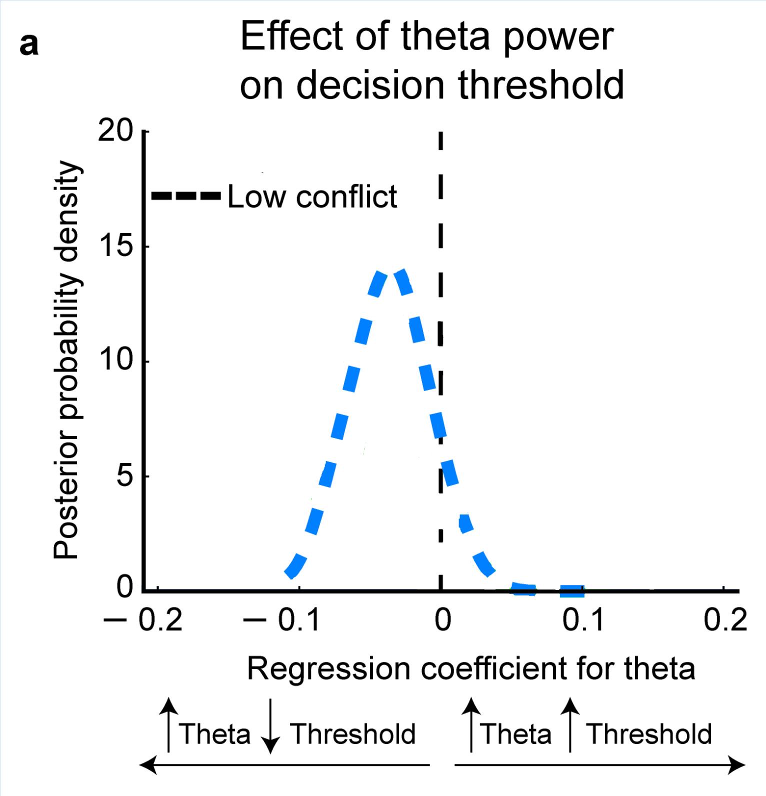


Genetic variants associated with better uninstructed learning
⇒ *more* value distortion by priors

⇒ Supports PFC-BG bias model: activity drives learning.

Doll et al, 2011, *JNeurosci*

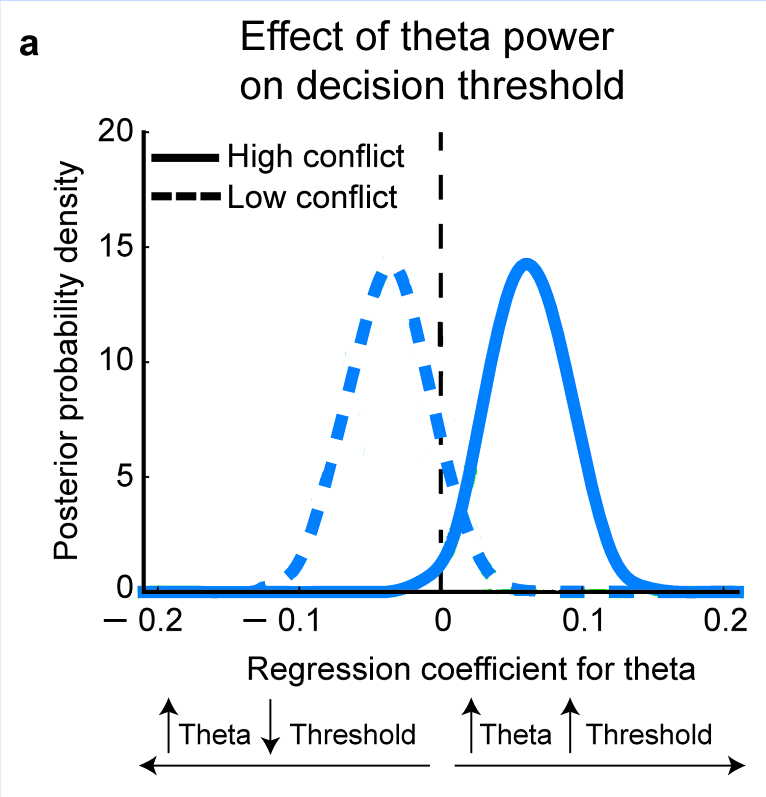
diffusion model fits to PD choices/EEG



OFF DBS

hierarchical-Bayes param estimation tool for DDM: `http://ski.cog.brown.edu/hddm_docs`

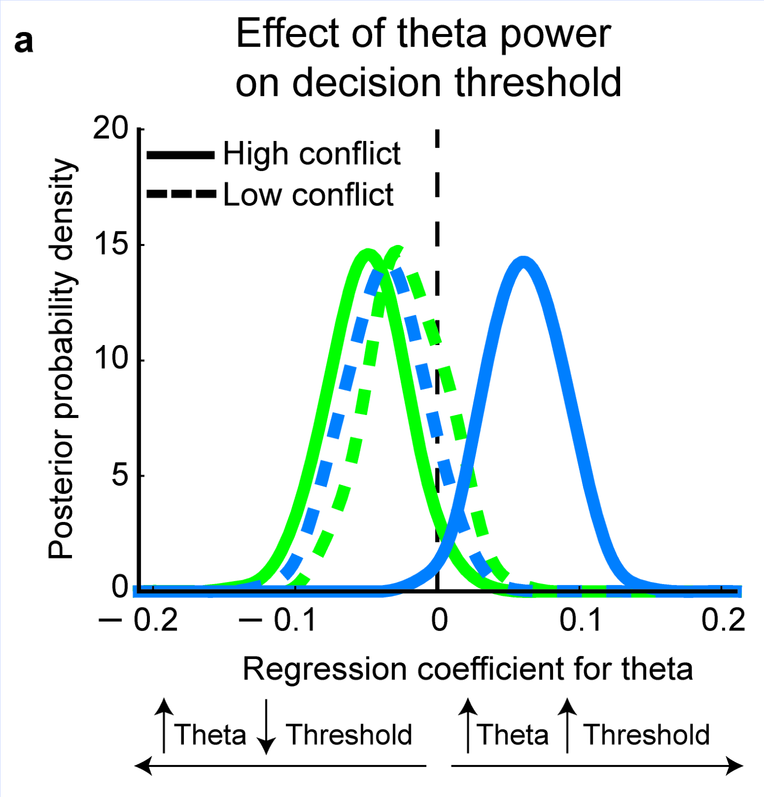
diffusion model fits to PD choices/EEG



OFF DBS

$\Rightarrow$ *threshold increases with theta in high conflict trials*

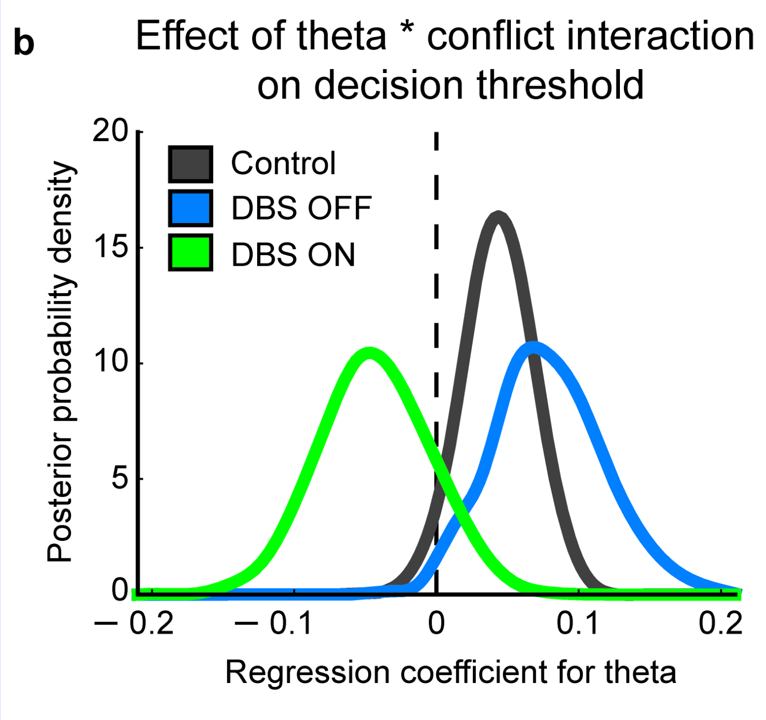
DBS reverses mPFC influence over decision threshold



ON DBS

Cavanagh et al, 2011 Nature Neuroscience

STN-DBS reverses mPFC influence over decision threshold



Cavanagh et al, 2011 Nature Neuroscience