What is Dopamine Doing?

Dopamine carries the brain's reward signal.
What is dopamine doing?
What is dopamine doing?
What is Dopamine Doing?
What is Dopamine Doing?

Dopamine carries the brain's reward signal.
What is Dopamine Doing?
Reinforcement learning and dopamine: prediction errors

Schultz, Satoh, Roedel, Zoghbi, Glimcher, Hyland, and many more

Negative PE: Positive PE:
Basic Data: VTA dopamine firing in Conditioning

Schultz, Montague & Dayan, 2007
Temporal Difference Learning: Equations

Value function, sum of discounted future rewards:

\[ V(t) = \langle \gamma^0 r(t) + \gamma^1 r(t + 1) + \gamma^2 r(t + 2) \ldots \rangle \] (1)
Temporal Difference Learning: Equations

Value function, sum of discounted future rewards:

\[ V(t) = \langle \gamma_0 r(t) + \gamma_1 r(t+1) + \gamma_2 r(t+2) + \ldots \rangle \]

(1)

Recursive definition:

\[ V(t) = \langle r(t) + \gamma V(t+1) \rangle \]

(2)

\[ (\cdot) + (1 + \cdot) + (\cdot) + (\cdot) + (\cdot) = (\cdot) \]

Recursive definition:

\[ \langle (1 + \cdot) + (\cdot) + (\cdot) + (\cdot) + (\cdot) \rangle = (\cdot) \]

Value function, sum of discounted future rewards:

Temproal Difference Learning: Equations
Temporal Difference Learning: Equations

1. **Value function, sum of discounted future rewards:**
   \[ V(t) = \left\langle \gamma^0 r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2) + \cdots \right\rangle \]

2. **Recursive definition:**
   \[ V(t) = \left\langle r(t) + \gamma V(t+1) \right\rangle = (t)\Lambda \]

3. **Error in predicted reward (from previous to next time-step):**
   \[ \delta(t) = (r(t) + \gamma \hat{V}(t+1)) - \hat{V}(t) \]

\[ (t) \Lambda - \left( (1 + t) \Lambda + (t)\nu \right) = (t)\rho \]

**Recursively:**
\[ \langle (1 + t) \Lambda + (t)\nu \rangle = (t)\Lambda \]

**Value function, sum of discounted future rewards:**
\[ \langle \cdots + (t)\nu z + (1 + t)\nu \hat{u} + (t)\nu o \rangle = (t)\Lambda \]

Temporal Difference Learning: Equations
Temporal Difference Learning: Equations

Value function, sum of discounted future rewards:
\[ V(t) = \langle \gamma^0 r(t) + \gamma^1 r(t+1) + \gamma^2 r(t+2) + \ldots \rangle \]  

(1)

Recursive definition:
\[ V(t) = \langle r(t) + \gamma V(t+1) \rangle \]  

(2)

Error in predicted reward (from previous to next time-step):
\[ \delta(t) = (r(t) + \gamma \hat{V}(t+1)) - \hat{V}(t) \]  

(3)

Update value estimate:
\[ \hat{V}(t) \leftarrow \hat{V}(t) + \alpha \delta(t) \]  

(4)

\( \alpha \) = learning rate

---

Let \( \Lambda \) denote the value function, sum of discounted future rewards:
\[ \langle \cdots (z + \gamma \Lambda + (1 + \gamma) \Lambda + (\gamma \Lambda) + \Lambda) \rangle = V(t) \]  

(5)

Temporal Difference Learning: Equations
Burst/Pause correlations with Rew Prediction Errors

Bayer et al. 2007 JNeurophys
How are dopamine-based RPE signals used to select actions?
What Do the Basal Ganglia Do?
What Do the Basal Ganglia Do?

• Hardly Anything: BG do not directly implement any cognitive (or motor) process.
What Do the Basal Ganglia Do?

- Hardly Anything: BG do not directly implement any (cognitive or motor) processes.
What Do the Basal Ganglia Do?

- Hardly Anything: BG do not directly implement any cognitive (or motor) process.
- Almost Everything: BG modulate activity in multiple cortical areas; affects motor, implicit learning, motivation, decision making and executive function processes.
- Parkinson’s disease (PD), ADHD: DA depletion in BG, resulting deficits in all above domains.
What Do the Basal Ganglia Do?

• Hardly Anything: BG do not directly implement any cognitive (or motor) process.

• Almost Everything: BG modulate activity in multiple cortical areas: affects motor, implicit learning, motivation, decision making and executive function processes.

• Parkinson’s disease (PD), ADHD: DA depletion in BG, resulting deficits in all above domains.

• Also: excess BC DA can induce impulsivity, e.g. pathological gambling, compulsive shopping (for review Dagher & Robbins, 2009).
Fronto-basal ganglia circuits in motivation, action, cognition
BG damage ⇒ deficits in motor, learning, motivation, working memory, cognitive control


Basal Ganglia Architecture: Cortically Based Loops
The Basal Ganglia as a Gate: Action Selection

- Each action, learned via dopamine...
- Gating occurs in proportion to relative probability of positive-negative outcomes for each action, learned via dopamine (Ivry & Spencer, 2004).
- BG selectively facilitates (gates) one action while suppressing others (Mink, 1996; Frank et al., 2001; Cunney et al., 2001; Brown et al., 2004).
Striato-Cortical Functional Circuitry
Striato-Cortical Functional Circuitry

Basal Ganglia

Thalamus

GPe

SNr/GPi

STN

Go/NoGo

SNc

Pre/motor Cortex

Excitatory –

Inhibitory •

Modulatory ▲
Neural Model of BG and dopamine (DA)

Integrates a wide range of physiological data into a single coherent framework

Frank, 2005, 2006
Striato-Cortical Functional Circuitry

- Thalamus
- SNr/GPi
- Pre/motor cortex

- Excitatory
- Inhibitory
- Modulatory
Striato-Cortical Functional Circuitry

Simulation

Thalamus

SNr/GPi

Corticostriatal

tonically active

Excitatory

Inhibitory

Pre/motor cortex

Striato-Cortical Functional Circuitry
Striato-Cortical Functional Circuitry

- Thalamus
- Striatum
- SNr/GPi
- GO
- Pre/motor cortex

Connections:
- Direct pathway
- "Disinhibition" pathway
- Modulatory pathways

Types of connections:
- Excitatory
- Inhibitory
- Modulatory
Striato-Cortical Functional Circuitry

- Excitatory
- Inhibitory
- Modulatory

- Striatum
- Thalamus
- Pre/motor cortex
- Go
- SNr/GPi

Simulation

Direct

"Disinhibition"
Disinhibition as a gating mechanism

Hikosaka and colleagues; Chevaller & Deniau, 90 etc.
Striato-Cortical Functional Circuitry

- Excitatory
- Inhibitory
- Modulatory

- GPe
- SNr/GPi
- Striatum
- Thalamus
- Direct
- Indirect
- Go
- Nogo
- Pre/motor
- Cortex

---

Striato-Cortical Functional Circuitry
Striato-Cortical Functional Circuitry

**Excitatory**

**Inhibitory**

**Modulatory**

- Thalamus
- Striatum
- Direct
- Indirect
- SNr/GPi
- Go
- Nogo
- Pre/motor cortex
- Simulation

Striato-Cortical Functional Circuitry
Samejima et al., 2005 Science

Separate striatal populations code for pos/neg action values
Evidence for go/no-go mechanism:
Optogenetic stimulation of direct and indirect pathways...and induces/inhibits movement

Kravitz et al, 2010, Nature

Go inhibits SNr; NoGo excites SNr
Dopamine effects on BG Learning: Positive PE

**SNc**

**GPe**

**D2**

**excitatory**

**inhibitory**

**modulatory**

**striatum**

**indirect**

**direct**

**thalamus**

**pre/motor cortex**

**D1**

**Go**

**NoGo**

**LTP**

**LTD**

**SNr/GPi**

**Dopamine effects on BG Learning: Positive PE**
Dopamine effects on BG Learning: Negative PE

- Excitatory
- Inhibitory
- Modulatory

Striatum

SNr/GPi

Dopamine
Simulating human learning and DA meds
Simulating human learning and DA meds

DA dips from inducing NoGo learning.

Medication ↓ DA levels, but tonic stimulation of D2 receptors prevents...

Cools et al, 2001; Frank, 2005

Simulated DA meds

Intact
Support for go/no go learning & choice mechanisms in rodents
Blocking neurotransmission in mouse Go/NoGo pathways

Hikida et al., 2010, Neuron
Intact nets extracted probabilistic structure by resolving differences in Go/NoGo representations.

PD nets were impaired due to reduced dynamic range of DA.

Frank, 2005, J Cog Neurosci
Reward prediction error and human functional imaging

O'Doherty et al., 2004; McClure et al., 2003; Daw et al., 2006; Caplin et al., 2010; Badre & Frank, 2011.
Human Probabilistic Reinforcement Learning

Avoid B?

Choose A?

Test

Train

A > CDEF

B > CDEF

E (60/40) F (40/60)
C (70/30) D (30/70)
A (80/20) B (20/80)
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, Bédil et al 09, Palminteri et al 09, Yoon et al 10)
BG model: DA modulates learning from pos/neg PE's

- Go learning to positive S-R requires sufficient phasic DA bursts
- NoGo learning to negative S-R requires sufficiently low DA during pauses

BG model: DA modulates learning from pos/neg PE's
Pause duration facilitates NoGo learning (P2)

Burst magnitude facilitates Go learning (D1)

Striatal Go/NoGo Activity

Test Condition

Go Pos NoGo Neg

50% Duration

50% Burst Mag

Sim PD

Sim DA Meds

BG Model Go/NoGo Associations

BG model: DA modulates learning from pos/neg PEs
DA stimulation vs. D2 blockade on go/nogo learning

Filled bars = medicated (D2 blockade or dopamine)
Open bars = unmedicated

See Palminteri et al., 2009 for model of D2 blockade effects on Nogo learning in rats.
Genetics of striatal dopamine function and model-based predictions
Genetics of striatal dopamine function and model-based predictions

- DARPP-32: protein concentrated in striatum, required for D1-dependent plasticity (Calabresi et al 00, Stipanovich et al 08)

Meyer-Lindenberg et al, 2007
Genetics of striatal dopamine function and model-based predictions.
Genetics of striatal dopamine function

Model: D1 = probabilistic Go learning

DARPP-32: protein concentrated in striatum, required for D1-dependent plasticity (Calabresi et al., 2000, Stipanovich et al., 2008)

Dylan quotes Aristotle quotes Plato on DARPP-32!

Meyer-Lindenburg et al., 2007

and model-based predictions

Genetics of striatal dopamine function
DRD2 gene: affects striatal D2 receptor function

Hirvonen et al., 2009
Drd2 gene affects striatal D2 receptor function

and here's what the red hot chili peppers have to say about this gene

Hirvonen et al. 2009
DRD2 gene affects striatal D2 receptor function

Model: D2 = probabilistic NoGo learning

Hirvonen et al., 2009

and here's what the red hot chili peppers have to say about this gene.
Frank et al. 07, PNAS

DA Genes and Probabilistic Learning
Infering learned values with RL model:

DARPP-32 effects

Frank et al. 07, PNAS
NoGo learning and D2 polymorphisms

Hirvonen et al. (2005); Mol Psychiatry.

Frank et al. (2007); Frank & Hutchison 2009; Klein et al. 2007.
Individual differences in go/no-go learning: striatal D1/D2 receptor binding

Preliminary data, with Sylvia Cox and Alain Dagher
Neurogenetic and pharmacological modulation of reinforcement learning parameters

Frank & Fossella, 2011
Neurogenetic and pharmacological modulation of reinforcement learning parameters

Frank & Fossella, 2011
(Influence of value on choice) in ventral striatum

Not just learning: DA modulates “incentive salience”

Salamone et al., 2003, …
OPponent Actor Learning (OpAL) model: Dissecting DA contributions to learning and choice incentive
DA modulates Go/NoGo activity states; DA depletion exaggerates the expression of prior NoGo learning (reward omission associated with high effort cue).
DA modulates Go/NoGo activity states; DA depletion exaggerates the expression of prior NoGo learning (reward omission associated with high effort cue) or benefits vs. costs. This can be optimized! (e.g., Niv et al. 2007)

DA acts as a knob on whether actions are selected primarily based on learned

tonic DA modulates Go/NoGo activity states, DA depletion exaggerates the expression of

<table>
<thead>
<tr>
<th>Test Condition</th>
<th>Striatal NoGo Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF1</td>
<td>0.0</td>
</tr>
<tr>
<td>FF5</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Negative value of effort
DA not needed for expression of low effort, but is for learning in the 1st place (Smith-Roe & Kelley 02)

DA modulates Go/NoGo activity states; DA depletion exaggerates the expression of prior NoGo learning (reward omissions associated with high effort cue)  

This can be optimized! (e.g., Niv et al 2007)

Tonic DA acts as a knob on whether actions are selected primarily based on learned benefits or costs.

Tonic DA not needed for expression of low effort, but is for learning in the 1st place (Smith-Roe & Kelley 02)

DA modulates Go/NoGo activity states; DA depletion exaggerates the expression of prior NoGo learning (reward omissions associated with high effort cue)  

This can be optimized! (e.g., Niv et al 2007)

Tonic DA acts as a knob on whether actions are selected primarily based on learned benefits or costs.

Tonic DA not needed for expression of low effort, but is for learning in the 1st place (Smith-Roe & Kelley 02)

DA modulates Go/NoGo activity states; DA depletion exaggerates the expression of prior NoGo learning (reward omissions associated with high effort cue)  

This can be optimized! (e.g., Niv et al 2007)

Tonic DA acts as a knob on whether actions are selected primarily based on learned benefits or costs.

Tonic DA not needed for expression of low effort, but is for learning in the 1st place (Smith-Roe & Kelley 02)

DA modulates Go/NoGo activity states; DA depletion exaggerates the expression of prior NoGo learning (reward omissions associated with high effort cue)  

This can be optimized! (e.g., Niv et al 2007)

Tonic DA acts as a knob on whether actions are selected primarily based on learned benefits or costs.
Evidence: Cost of effort encoded in "Nogo" pathway

- D2 antagonists potentiate NoGo activity/plasticity (e.g., Salamone et al., 2002)
- The adenosine A2A receptor is colocalized with D2 on indirect pathway cells and its stimulation has opposite effects on activity/plasticity (e.g., Shen et al., 2008)
- Pharmacological inhibition of VP (which should mimic striatal Nogo signals downstream) also increases cost of effort (Farrer et al., 2008)
- Pharmacological inhibition of VP (which should mimic striatal Nogo signals downstream) also increases cost of effort (Farrer et al., 2008)

- Drugs that target striatal A2A but not A1A reverse the effects of D2 antagonists on effort cost (Salamone et al., 2009; Nunes et al., 2010, Farrer et al., 2010).

- Disconnection procedures show that this striatopallidal pathway mediates the effect (Mingote et al., 2008)
- Pharmacological inhibition of VP (which should mimic striatal Nogo signals downstream) also increases cost of effort (Farrer et al., 2008)

- Evidence: Cost of effort encoded in "Nogo" pathway
Deep Brain Stimulation of the Subthalamic Nucleus (STN) for treatment of Parkinson's disease

Video #1: http://ski.cip.s.pirow.edu/dbs.mp4
Video #2: http://ski.cip.s.pirow.edu/dbs2.mp4
But not all is grand in the world of DBS...
But not all is grand in the world of DBS...
But not all is grand in the world of DBS...

Hi, I found your email address in an article I was reading about DBS surgery for Parkinson's. My dad had the surgery last May and we have a mess on our hands. Two months following the surgery we began to notice some personality changes. He became impulsive, cocky, oblivious to his surroundings, forgetful, has lied, he has no empathy, he uses foul language, and he is always on the move. Going somewhere and buying something... canceled his 2 follow up appointments, he was always very detail-oriented and now he is sloppy, and he is spending a lot of money. He has not gone one day without buying something. He can't sit still, he's always restless and now he is sloppy, and he is spending a lot of money. He has become impulsive, cocky, oblivious to his surroundings, forgetful, has lied, he has no empathy, he uses foul language, and he is always on the move. Going somewhere and buying something...

STN-DBS dramatically improves PD motor symptoms, but can induce impulsivity (Saint-Cyr et al. 06, Frank et al. 07; Wythe et al. 10; Halbach et al. 09)
From reinforcement learning...
...to reinforcement conflict-based decision making
Anatomy of BG gating: without STN
Anatomy of BG gating: with subthalamic nucleus (STN)

PFC-STN provides an override mechanism
Subthalamic Nucleus: Dynamic modulation of decision threshold

Conflict (entropy) in choice prob: ➞ Hold Your Horses!
STN and frontal cortex are directly connected via white matter.
Neural model and STN ephys: decision conflict

Wiecki & Frank, 2013 Psych Review
data from Isoda & Hikosaka, 2008

Spike rate:

Neural model and STN ephys: decision conflict
Neural model and STN ephys: decision conflict

Wiecki & Frank, 2013 Psych Review

Data from Isoda & Hikosaka, 2008

Behavior:

Spike rate:

Neural model and STN ephys: decision conflict
Human probabilistic reward/choice conflict

Low Conflict: e.g., 80 vs 30%

High Conflict: e.g., 80 vs 70%

\[ H(\text{softmax}) = 0.84 \]

\[ H(\text{softmax}) = 0.06 \]
Human probabilistic reward/choice conflict

Need STN to prevent impulsive responses

Low Conflict: e.g., 80 vs 70%

High Conflict: e.g., 80 vs 30%

$H(x_{softmax}) = 0.84$

$H(x_{softmax}) = 0.06$
Need STN to prevent impulsive responses

\[ H(p_{\text{softmax}}) = 0.6 \]

\[ H(p_{\text{softmax}}) = 0.84 \]

Low Conflict: e.g., 80 vs 30%

High Conflict: e.g., 80 vs 70%

Human probabilistic reward/choice conflict
human STN spiking, Zaghloul et al., 2012
STN-DBS reverses conflict RT adjustments

Seniors Off DBS On DBS

Patient Condition

-500
-400
-300
-200
-100
0
100
200
300
400

RT Diff (ms)

Within-Subject Conflict

Deep Brain Stimulation

Frank, Samanta, Moustafa & Sherman (2007)

see also Wylie et al 10; Halbig et al 09; Cavanaugh et al 11; Coulthard et al 12; Green et al 13
Interim Summary

DBS induces speeded responding in conflict conditions

Simulations: STN modulates decision threshold × cortical conflict
Interim Summary

• DBS induces speeded responding in conflict conditions

• Simulations: STN modulates decision threshold as a function of mediofrontal conflict

• More precise predictions to be tested:
  - Does DBS-DBS alter this relationship?
  - Does decision threshold vary as a function of mediofrontal conflict?
  - Does mediofrontal cortex and STN represent reinforcement conflict?
  - Does STN-DBS alter this relationship?

Simulations: STN modulates decision threshold as a function of mediofrontal conflict

DBS induces speeded responding in conflict conditions
Abstraction: the drift diffusion model
Abstraction: the drift diffusion model

- Provides quantitative fits to error rates and RT distributions in many tasks
- Allows estimation of decision threshold ($\eta'$) separately from other factors ($v, z, T_e$)

---

E.G. Ratcliff & McKoon, 2008
Abstraction: the drift diffusion model

- Allows estimation of decision threshold ($\theta$) separately from other factors
- Provides quantitative fits to error rates and RT distributions in many tasks
Contrasting drift rate vs threshold
Mechanism

Mediodorsal influences over decision threshold
Subthalamic nucleus stimulation reverses
Hierarchical interactions in BG-FC circuits:

PFC & cognitive control influences on learning

Collins & Frank 2013, Psych Rev; Frank & Badre 2012
Broader speculations:

Why does motor control develop so slowly in humans?

Fourth trimester

Standard story: Infants born early due to large head, small birth canal
Broader speculations:

Why does motor control develop so slowly in humans?

• Standard story: Infants born early due to large head, small birth canal

• "Fourth trimester"

• But 3 month old infants still pretty incompetent (from babycenter.com)
Broader speculations:

Why does motor control develop so slowly in humans?

Standard story: Infants born early due to large head, small birth canal.

But 3 month old infants still pretty incompetent (from babycenter.com):

You no longer need to support his head. When he's on his stomach, he can lift his head and chest. He can open and close his hands.

'Fourth trimester'
Broader speculations:

Why does motor control develop so slowly in humans?

Hypothesis: Human brain is wired to discover generalizable structure.

But 3 month old infants still pretty incompetent (from babycenter.com):

The "Fourth trimester" standard story: Infants born early due to large head, small birth canal.

You no longer need to support his head. When he's on his stomach, he can lift his head and chest. He can open and close his hands.

Infants born early due to large head, small birth canal.
which is initially inefficient.
Task-sets (TS)

C1

S1  S2  S3

A1  A2  A3
Latent task-set space

Abstracting Task-set rules

T5 as abstract rule objects

Woolgar et al. 2011
Reverberi et al. 2011
Abstracting Task-sets rules

? → T52 → S12 → A12

? → T51 → S11 → A11
Abstracting Task-sets Rules
Abstracting Task-sets rules
Abstracting Task-sets Rules
Unknown size
Latent task-set space:
Abstracting Task-sets Rules

TS
TS
TS
TS
TS
TS
TS
TS
TS
TS
see also Gershman et al 2010

\[
\frac{\Lambda}{(\Lambda|S \perp L)\partial^L \Lambda} = \left\{ \begin{array}{l}
\frac{(1+u|S \perp L = S, L)\partial^L \Lambda}{(1+u|\Lambda \perp S, L = S, L)\partial^L \Lambda}
\end{array} \right\} = (1+u|L = S, L)\partial^L \Lambda
\]

Prior prob on TS space given a new C:

Task-sets are clustered

C-TS model
Implementation
Neurobiologically plausible
Neurobiologically plausible implementation
Neuropathologically plausible implementation
Neurobiologically plausible implementation
Predicts positive, negative transfer.

The network learns efficiently unsupervised.

Neural Network - Results
Re-using and creating task-sets

MRI evidence: Badre & Frank 2012

Collins & Frank 2013, Psych Rev; Collins et al. 2014 J Neurosci; Collins & Frank, in Review
Model mimicry: C-TS and Hierarchical BG-PFC network

Sparseness of context-PFC connectivity matrix is linked to a clustering

Both models are approximations of the same process: building TS structure

fMRI evidence for Hierarchical PFC-BG mechanisms Badre & Frank 2012

Collins & Frank 2013

Psych Rev