One critic for two actors

Nicolas P. Rougier

ONE CRITIC FOR TWO ACTORS

NICOLAS ROUGIER, INRIA
ROBOTICS & NEUROSCIENCES, BORDEAUX, 2016

STARRING CEREBRAL CORTEX, BASAL GANGLIA & THALAMUS
SPECIAL GUEST HEBBIAN LEARNING, REINFORCEMENT LEARNING & COVERT LEARNING
IN COOPERATION WITH THOMAS BORAUD, DAISUKE KASE, CAMILLE PIRON & MEROPI TOPALIDOU
HTTP://WWW.LABRI.FR/PERSO/NROUGIER/
PLAYING NOVEMBER 17TH, 2016

ONE CRITIC FOR TWO ACTORS

NICOLAS ROUGIER, INRIA

ROBOTICS & NEUROSCIENCES, BORDEAUX, 2016

STARRING CEREBRAL CORTEX, BASAL GANGLIA & THALAMUS

SPECIAL GUEST HEBBIAN LEARNING, REINFORCEMENT LEARNING & COVERT LEARNING

IN COOPERATION WITH THOMAS BORAUD, DAISUKE KASE, CAMILLE PIRON & MEROPI TOPALIDOU

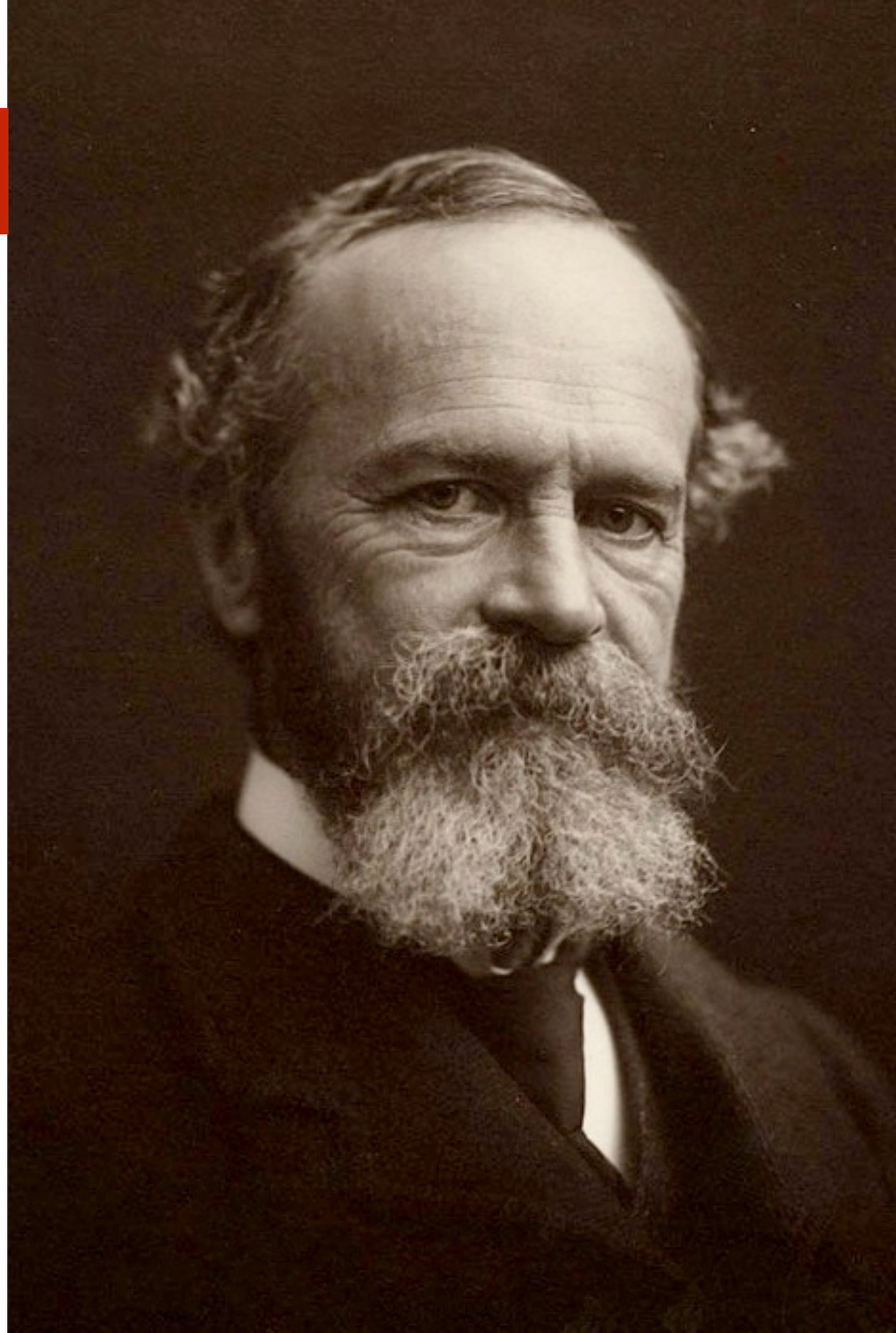
HTTP://WWW.LABRI.FR/PERSO/NROUGIER/

PLAYING NOVEMBER 17TH, 2016

Decision making

(1) the reasonable sort, whereby we accede to rational arguments (2) the sort that is triggered by external circumstances, such as overhearing a rumor; **(3) the sort that is prompted by our submission to something within ourselves, such as a habit formed by past actions;** (4) the sort that results from a sudden change of mood such as might be caused by a feeling of grief; and (5) the rare sort that is a consequence of our own voluntary choice, which will be identified as the “will to believe.”

William James, 1890



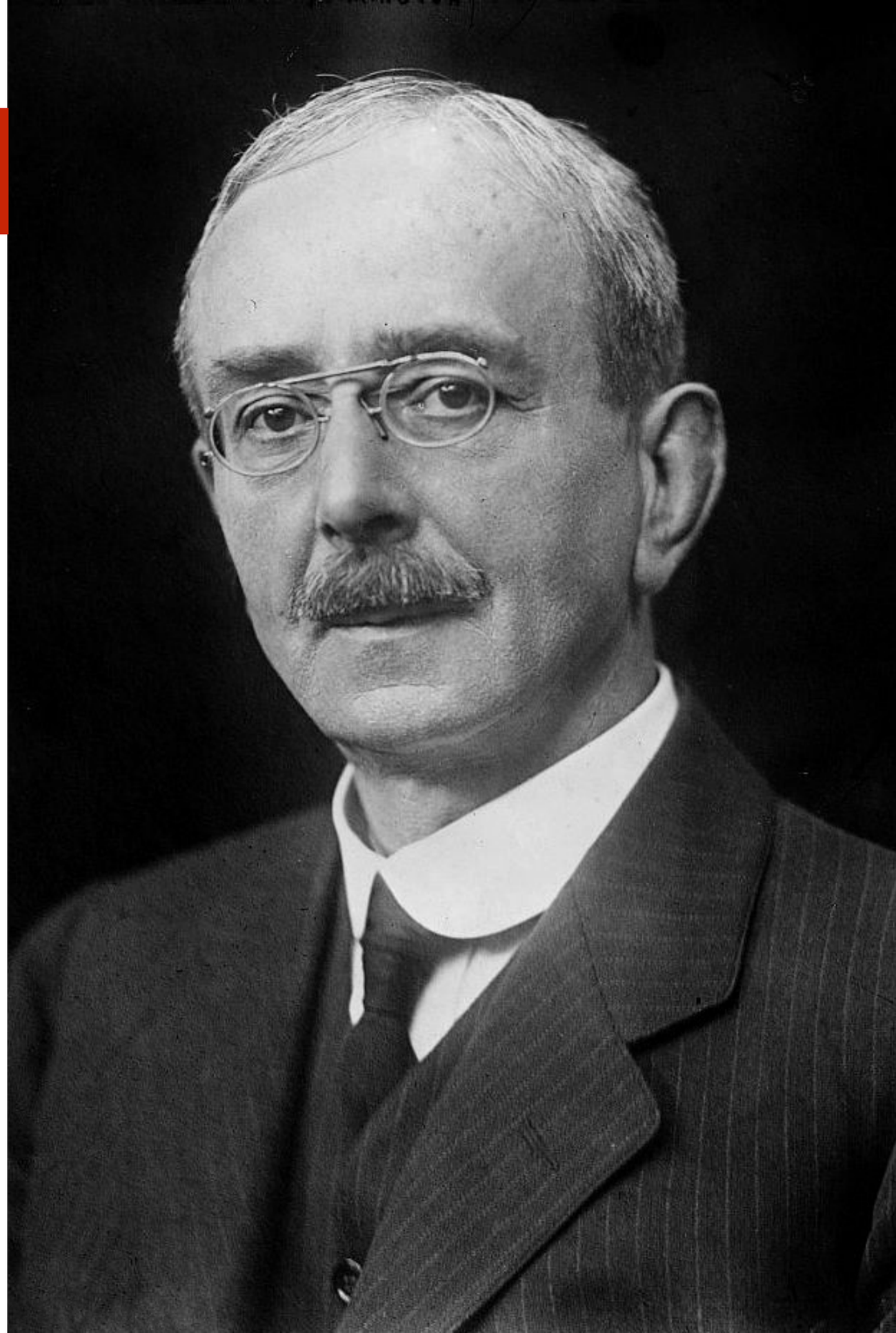
Decision making

(1) the reasonable sort, whereby we accede to rational arguments (2) the sort that is triggered by external circumstances, such as overhearing a rumor; **(3) the sort that is prompted by our submission to something within ourselves, such as a habit formed by past actions;** (4) the sort that results from a sudden change of mood such as might be caused by a feeling of grief; and (5) the rare sort that is a consequence of our own voluntary choice, which will be identified as the “will to believe.”

William James, 1890

The transition from reflex action to volitional is not abrupt and sharp. Familiar instances of individual acquisition of motor coordination are furnished by the cases in which short, simple movements, whether reflex or not, are by practice under volition **combined into new sequences and become in time habitual** in the sense that though able to be directed they no longer require concentration of attention upon them for their execution.

Charles Sherrington, 1906



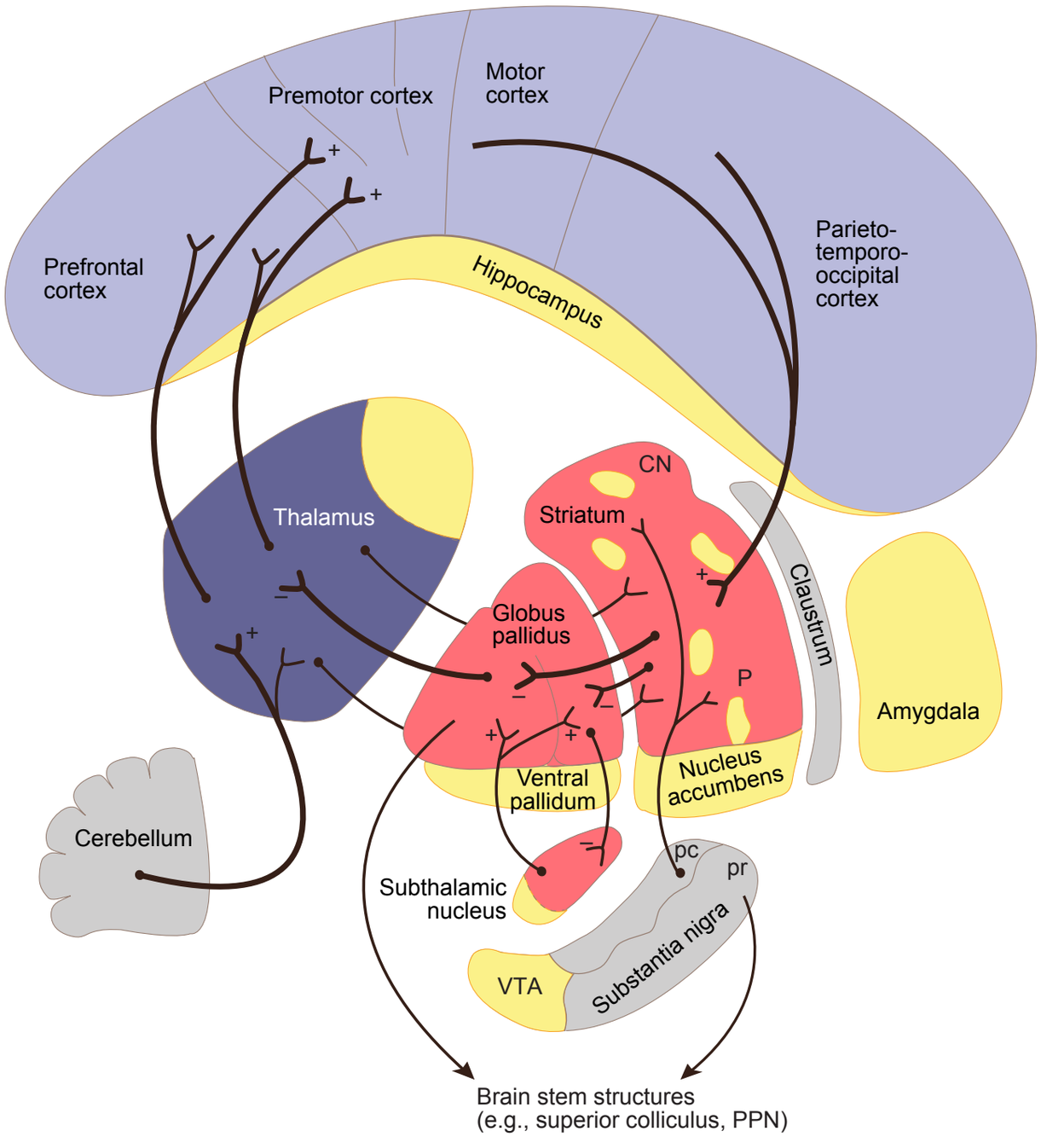
Decision making

(1) the reasonable sort, whereby we accede to rational arguments (2) the sort that is triggered by external circumstances, such as overhearing a rumor; **(3) the sort that is prompted by our submission to something within ourselves, such as a habit formed by past actions;** (4) the sort that results from a sudden change of mood such as might be caused by a feeling of grief; and (5) the rare sort that is a consequence of our own voluntary choice, which will be identified as the “will to believe.”

William James, 1890

The transition from reflex action to volitional is not abrupt and sharp. Familiar instances of individual acquisition of motor coordination are furnished by the cases in which short, simple movements, whether reflex or not, are by practice under volition **combined into new sequences and become in time habitual** in the sense that though able to be directed they no longer require concentration of attention upon them for their execution.

Charles Sherrington, 1906



Decision making

Brain is optional

- **Decision-making without a brain: how an amoeboid organism solves the two-armed bandit** (2016)
  Chris R. Reid, Hannelore MacDonald, Richard P. Mann, James A. R. Marshall, Tanya Latty, Simon Garnier

- **Habituation in non-neural organisms: Evidence from slime moulds** (2016)
  Romain P. Boisseau, David Vogel & Audrey Dussutour

- **A two-neuron system for adaptive goal-directed decision-making in Lymnaea** (2016)
  Crossley M., Staras K., Kemenes G.

- **Functional organization and adaptability of a decision-making network in Aplysia** (2012)
  Nargeot R., Simmers J.

- **Neuronal microcircuits for decision making in C. Elegans** (2012)
  Faumont S., Lindsay T.H., Lockery, S. R.

- **Decision-making in soccer game: a developmental perspective** (2005)
  Rulence-Pâquesa P., Frucharta E., Drub V., Mullet E.

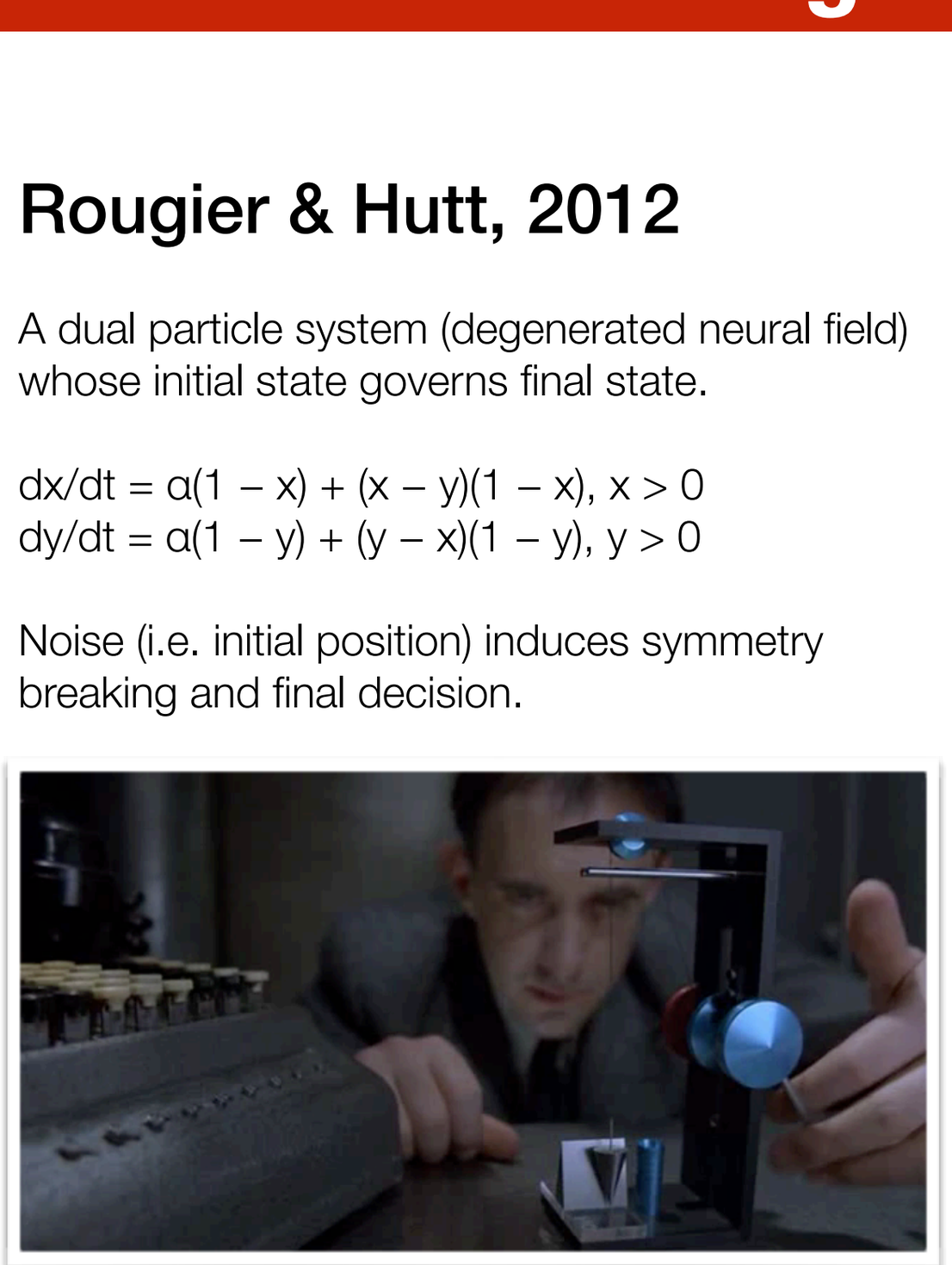
Decision making

Rougier & Hutt, 2012

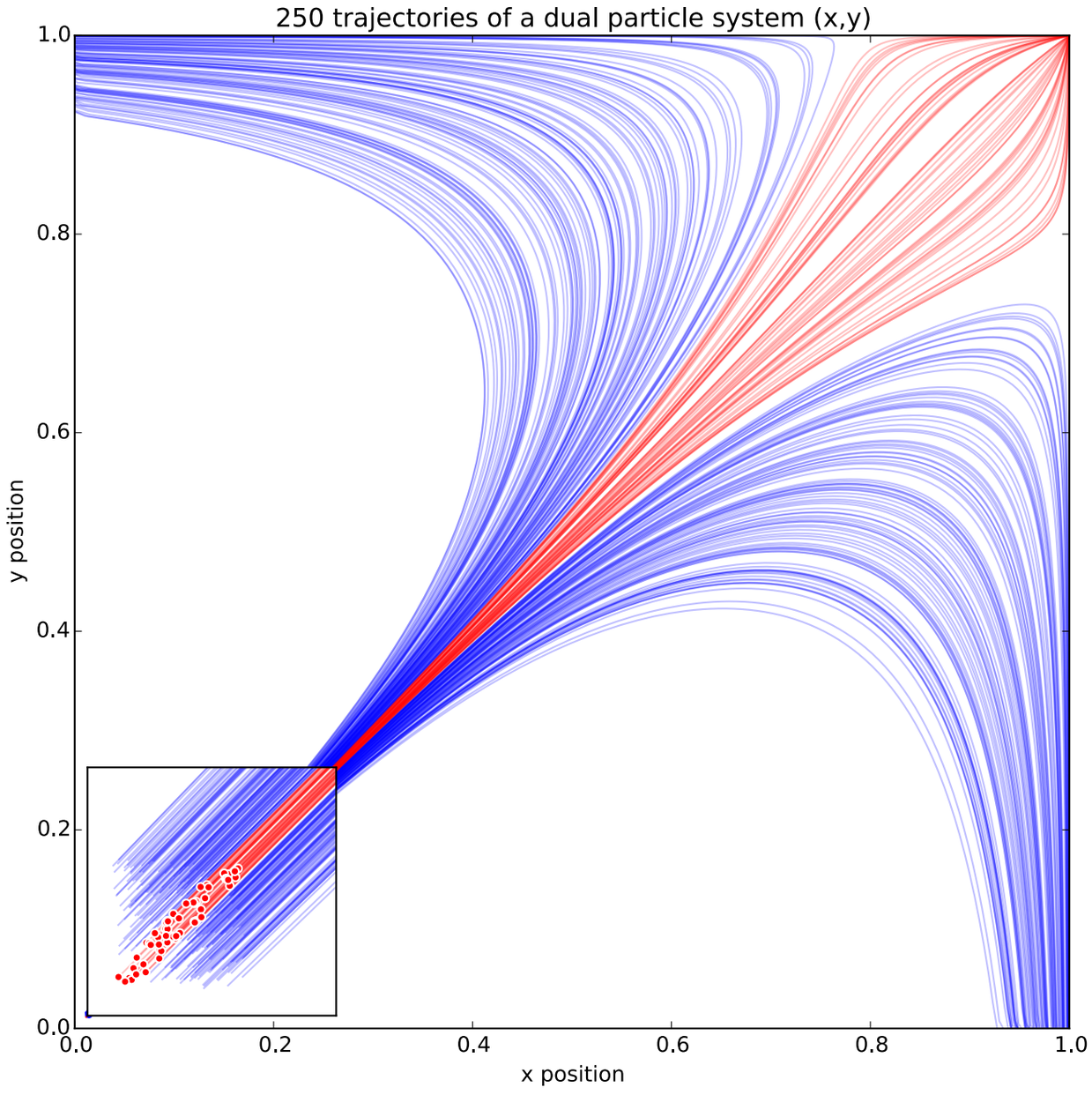
A dual particle system (degenerated neural field) whose initial state governs final state.

$dx/dt = \alpha(1 - x) + (x - y)(1 - x), x > 0$
$dy/dt = \alpha(1 - y) + (y - x)(1 - y), y > 0$

Noise (i.e. initial position) induces symmetry breaking and final decision.



The executive decision maker
Brazil, Terry Gilliam, 1985



The habit factor

The habit factor

A tentative definition

Yin and Knowlton (2006), Graybiel (2008), Seger and Spiering (2011)

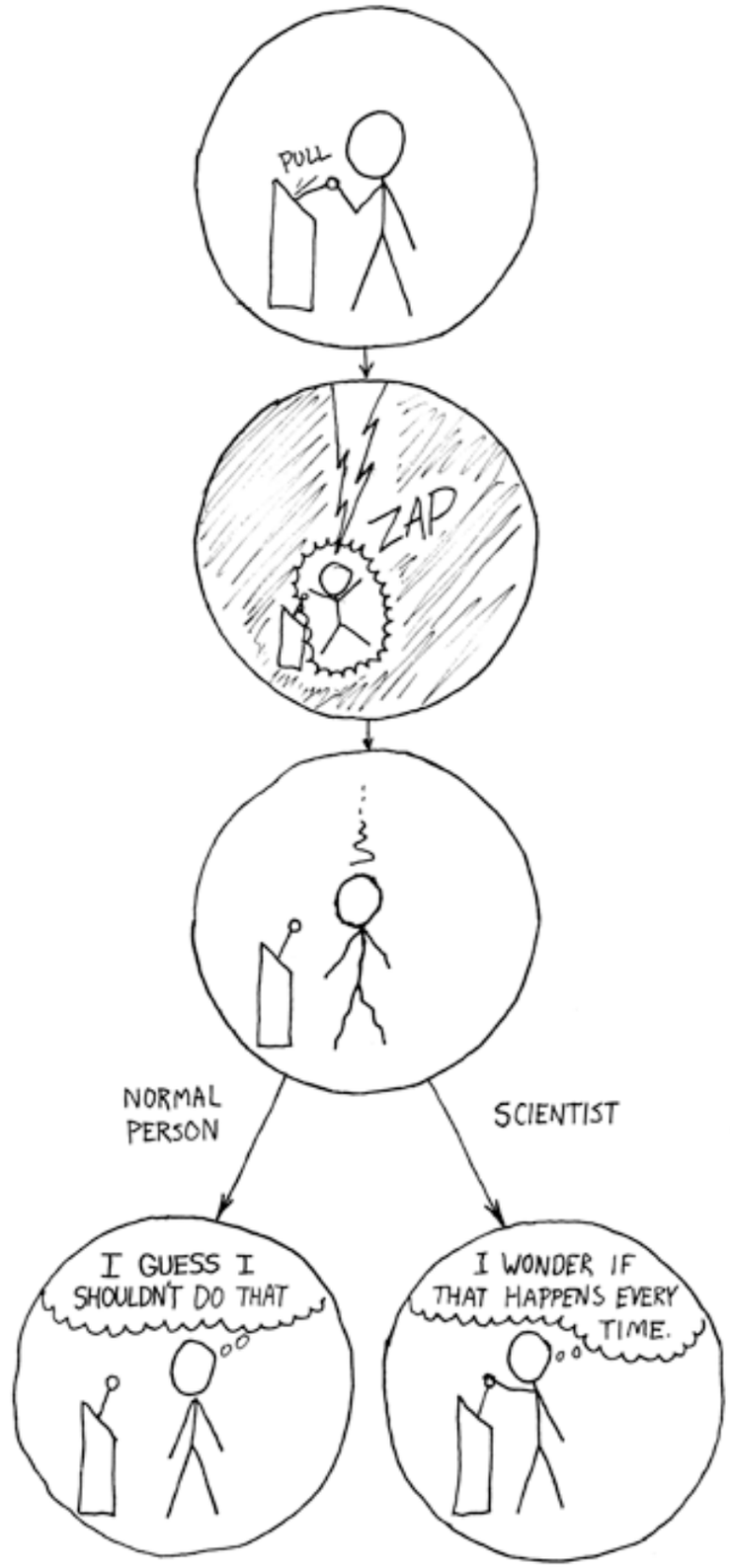
- Elicited by a particular context or stimulus
  → **stimulus-response** as opposed to **action-outcome**

- Acquired via experience
  → require extensive training or repetition

- Performed automatically
  → the mere presence of the stimulus induces the response

- Resistant to outcome devaluation
  → disengagement from the goal

- Performed unconsciously
  → without “thinking” about it



XKCD #242

A simple question

Action-outcome **then** stimulus-response?

Action-outcome comes first until being transformed into stimulus-response

Action-outcome **versus** stimulus-response?

Both processes are present and compete for expression

Action-outcome **with** stimulus-response?

Final decision is a mix of both processes

Action-outcome **and** stimulus-response?

Processes cooperate and influence each other, always

Main structures

Cortex (CTX)
- Posterior
- Motor / Premotor
- Prefrontal

Thalamus (THL)

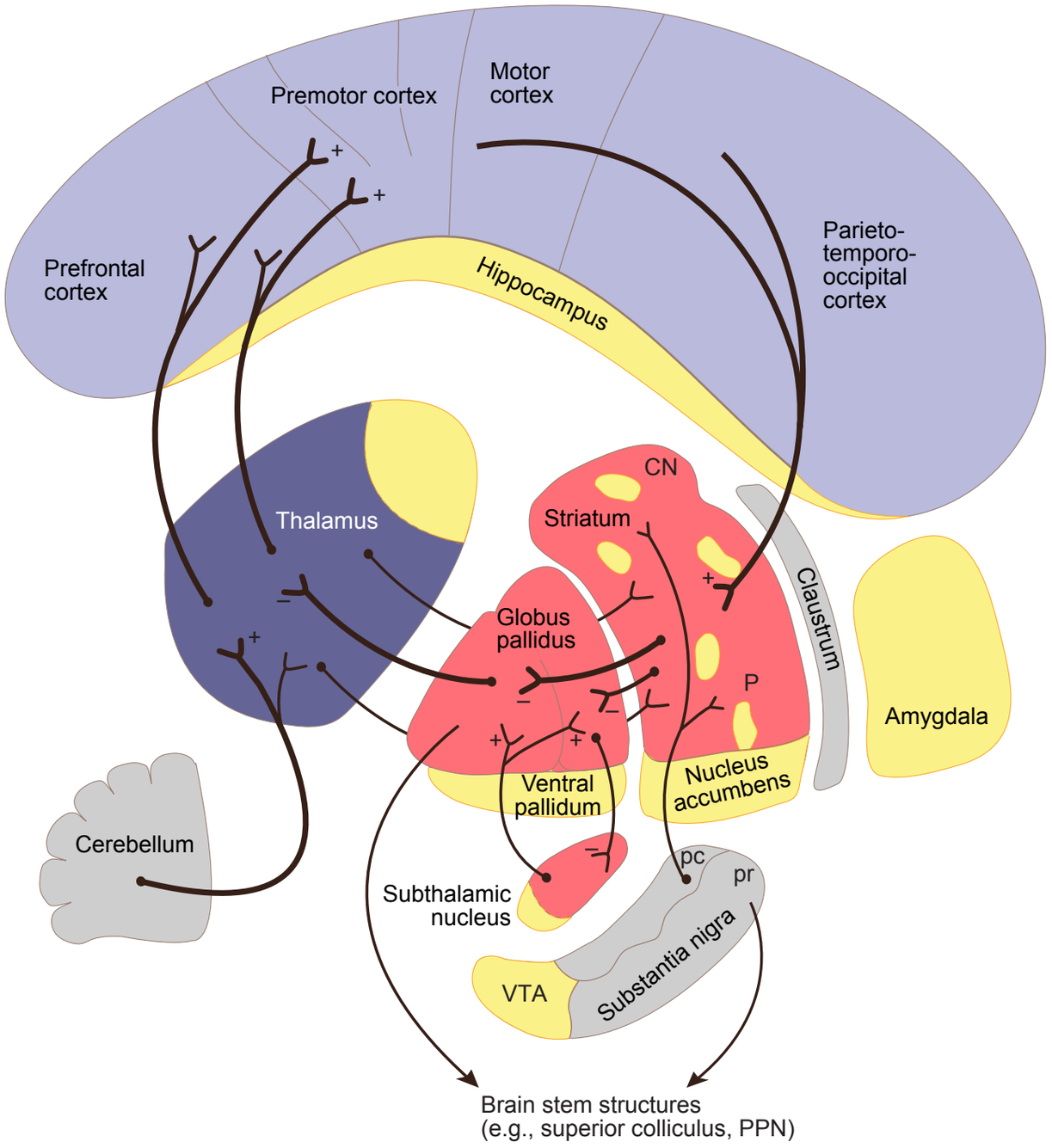
Amygdala (AMY)

Striatum (STR)
- Caudate
- Putamen
- Nucleus Accumbens

Subthalamic Nucleus (STN)

Globus Pallidus
- Internal (GPi)
- External (GPe)

Subtantia Nigra
- pars Compacta (SNc)
- pars Reticulata (SNr)

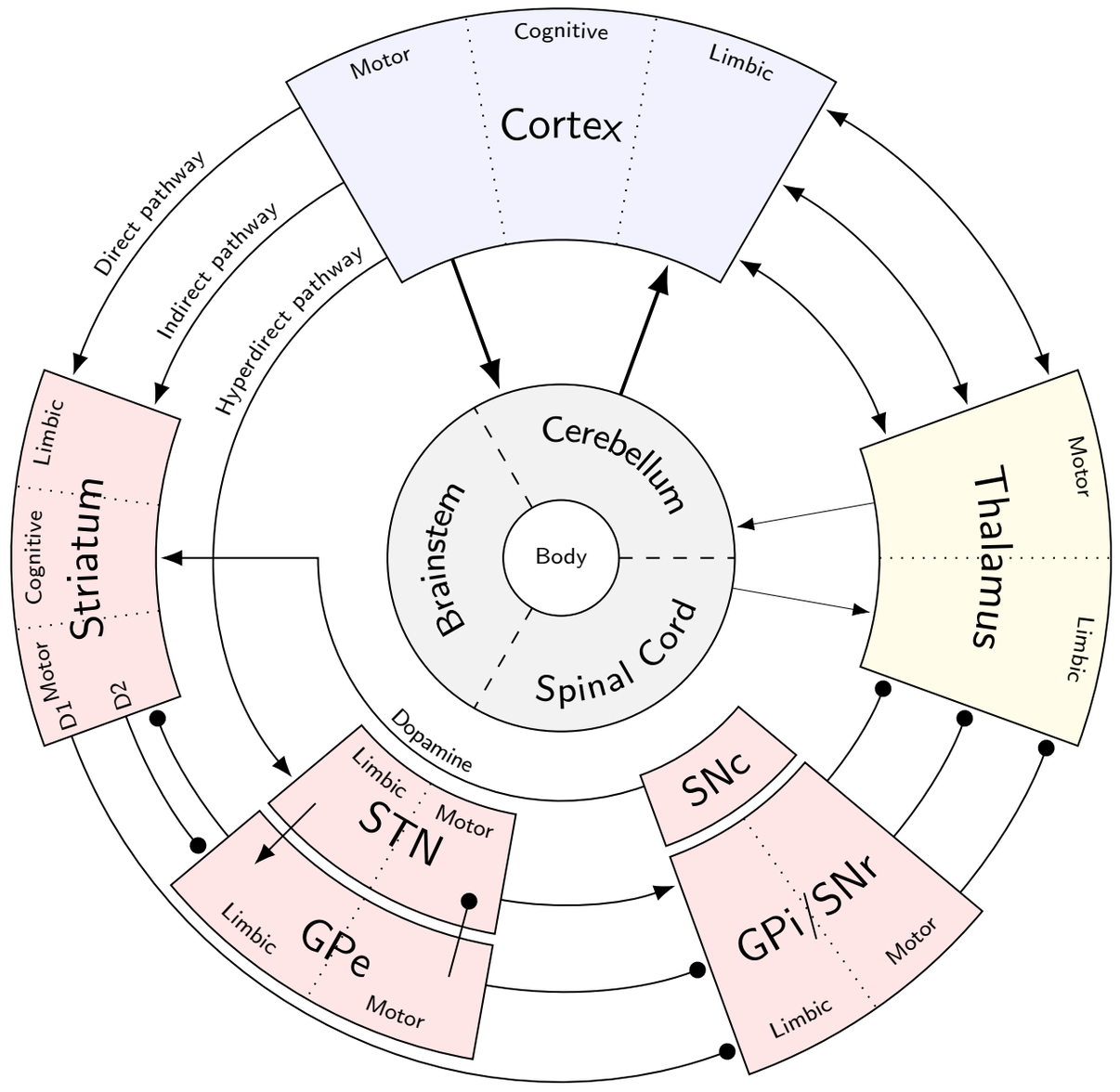


Functional pathways

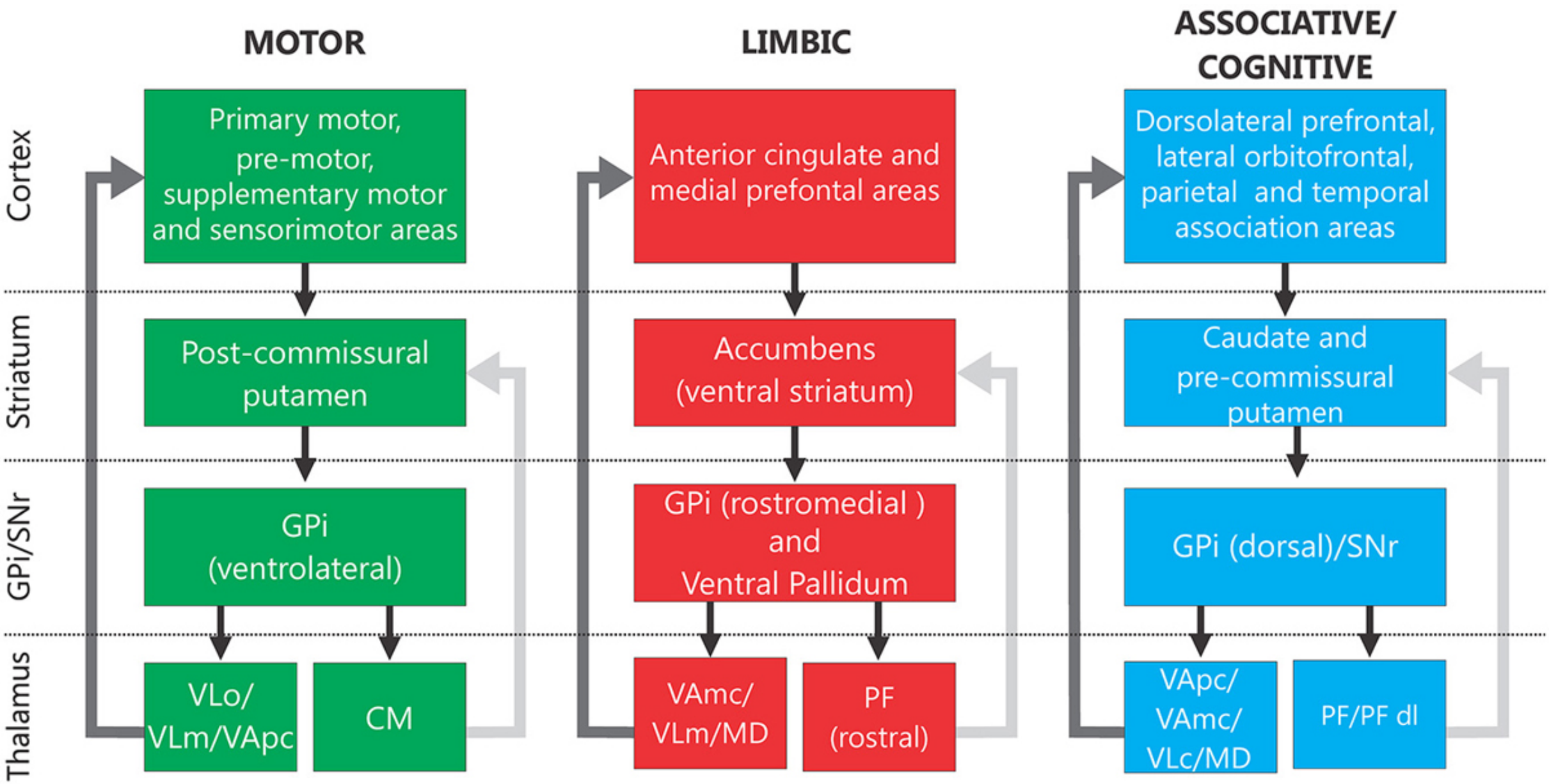
Direct pathway (go pathway)
CTX → STR → GPi/SNr → THL → CTX

Indirect pathway (no go pathway)
CTX → STR → GPe → GPi/SNr → THL → CTX
→ STN → THL → CTX

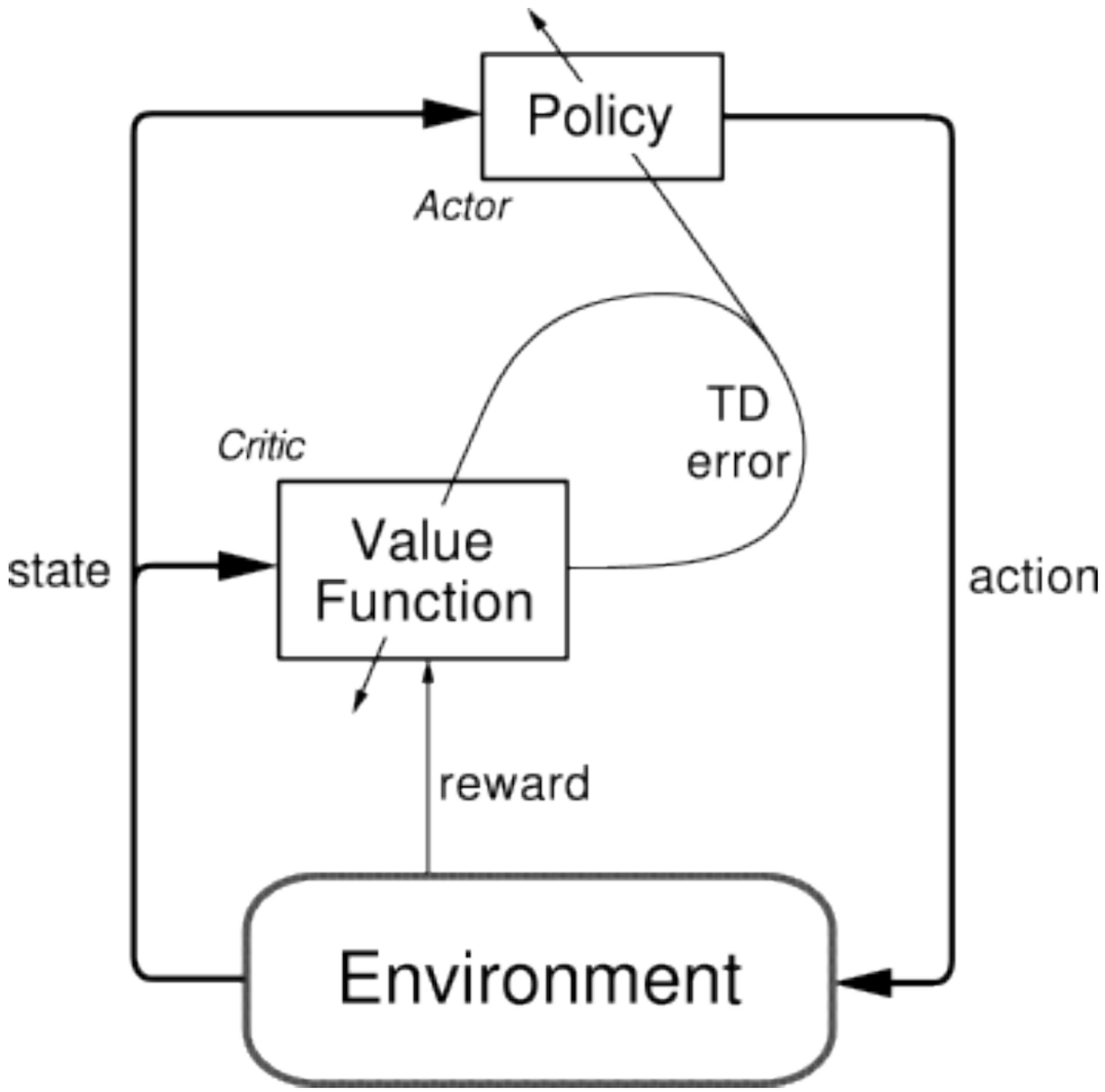
Hyperdirect pathway (stop pathway)
CTX → STN → GPi/SNr → THL → CTX



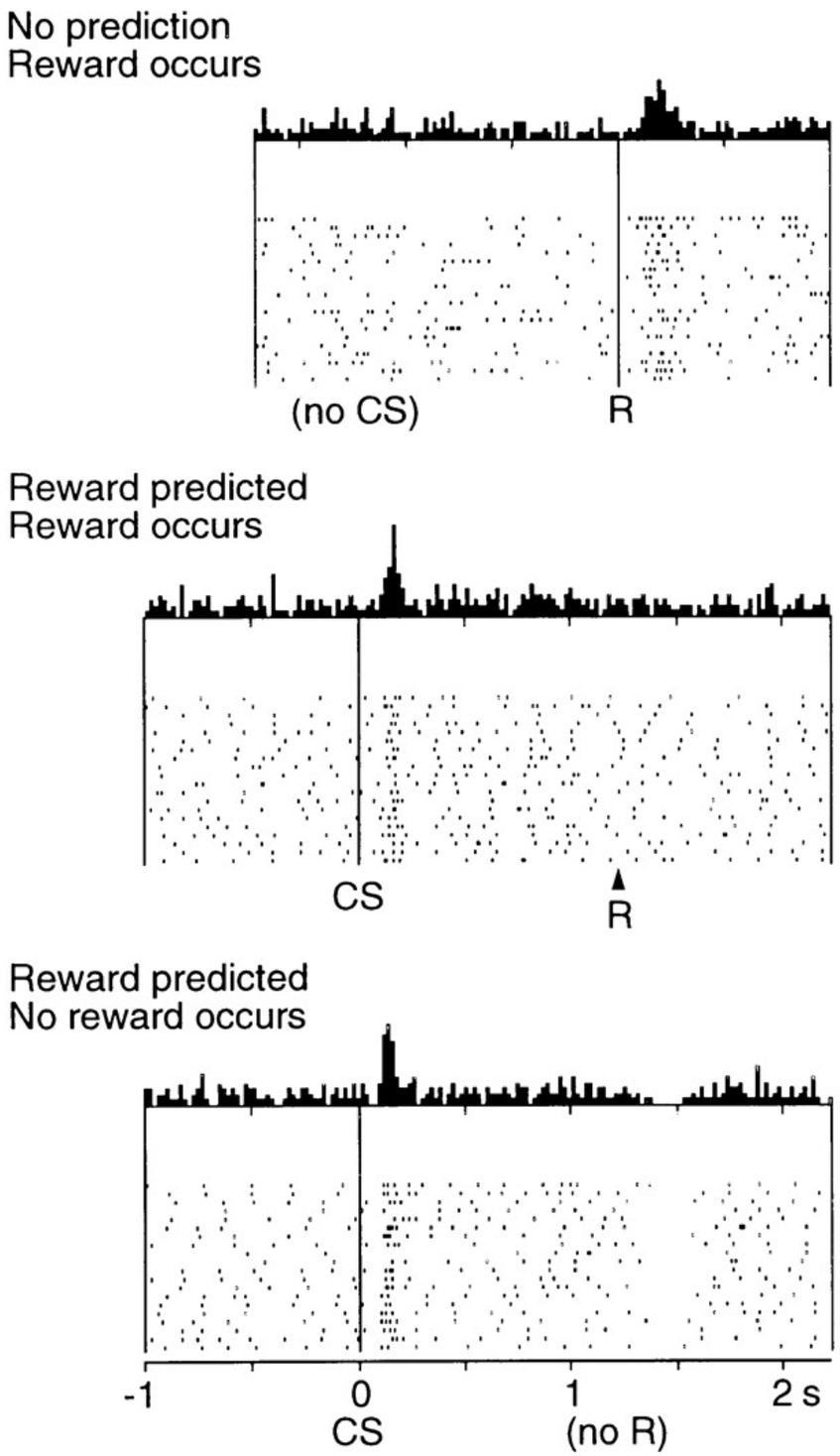
Segregated loops



Dopamine as RPE



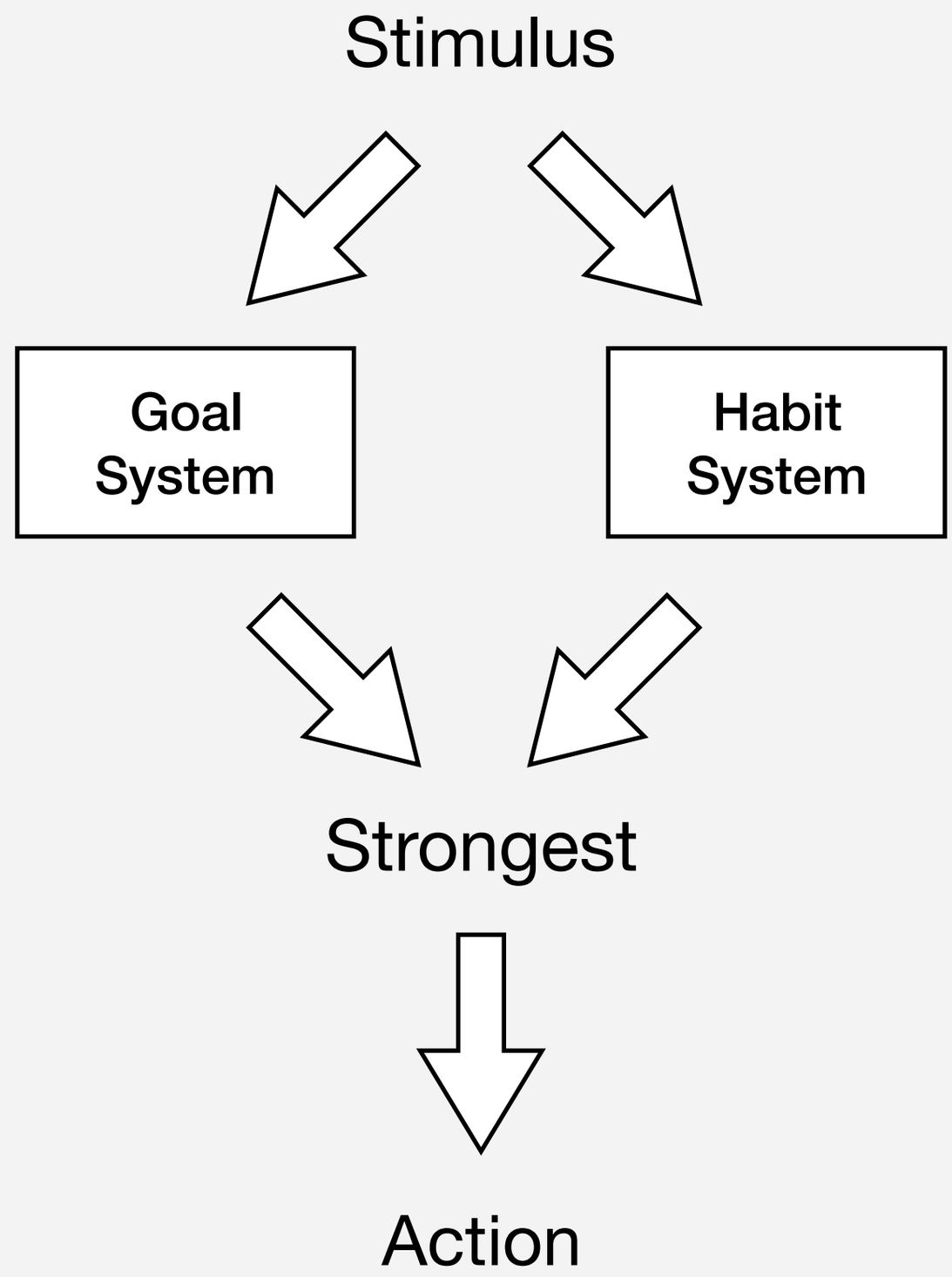
Sutton & Barto, 1998



Schulz et al., 1997

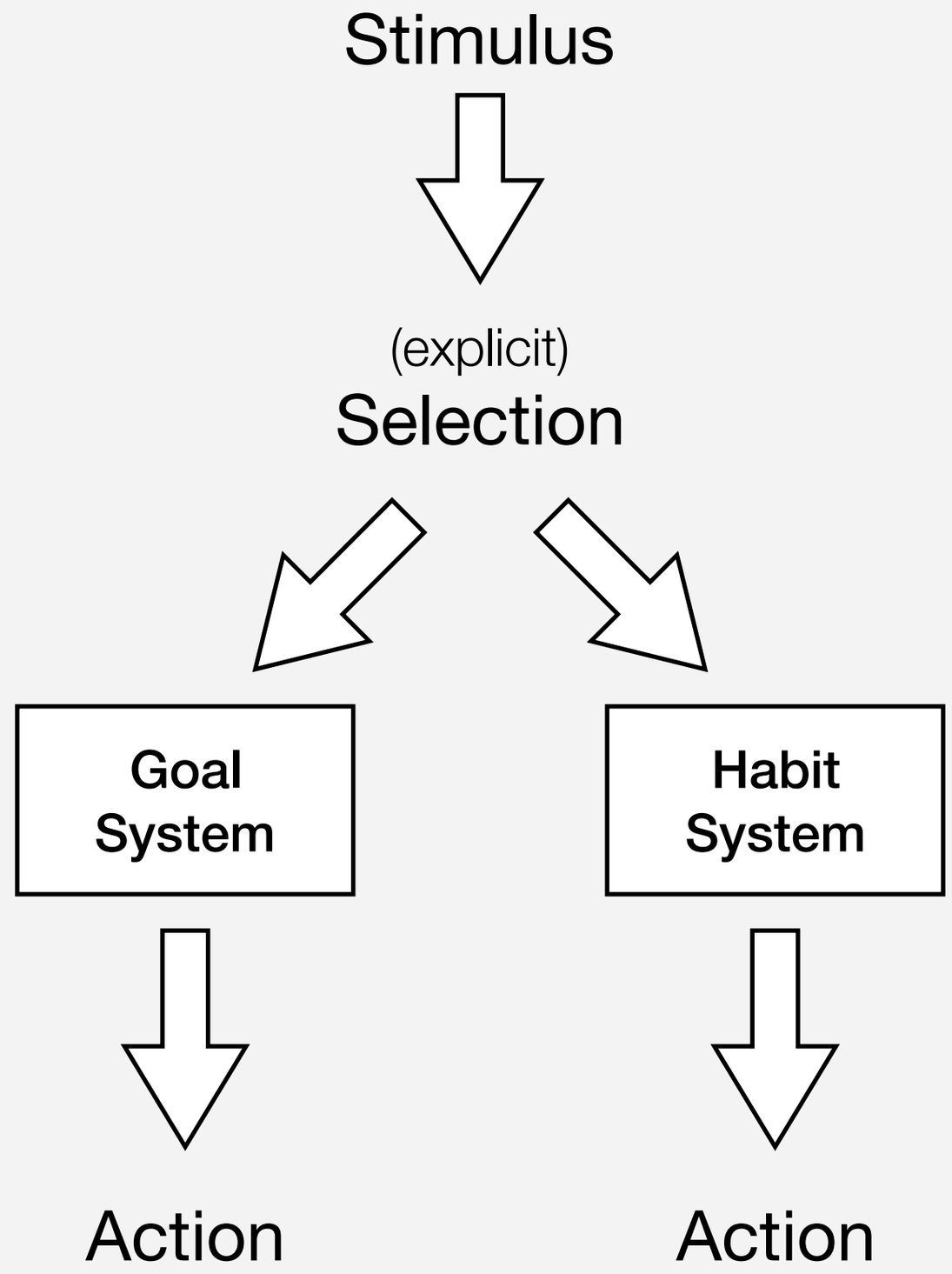
Daw et al. (2005)

"A broad range of neural and behavioral data suggests that the brain contains multiple systems for behavioral choice, including one associated with prefrontal cortex and another with dorsolateral striatum. However, such a surfeit of control raises an additional choice problem: how to arbitrate between the systems when they disagree. Here, we consider dual-action choice systems from a normative perspective, using the computational theory of reinforcement learning. We identify a key trade-off pitting computational simplicity against the flexible and statistically efficient use of experience. The trade-off is realized in **a competition between the dorsolateral striatal and prefrontal systems**…"



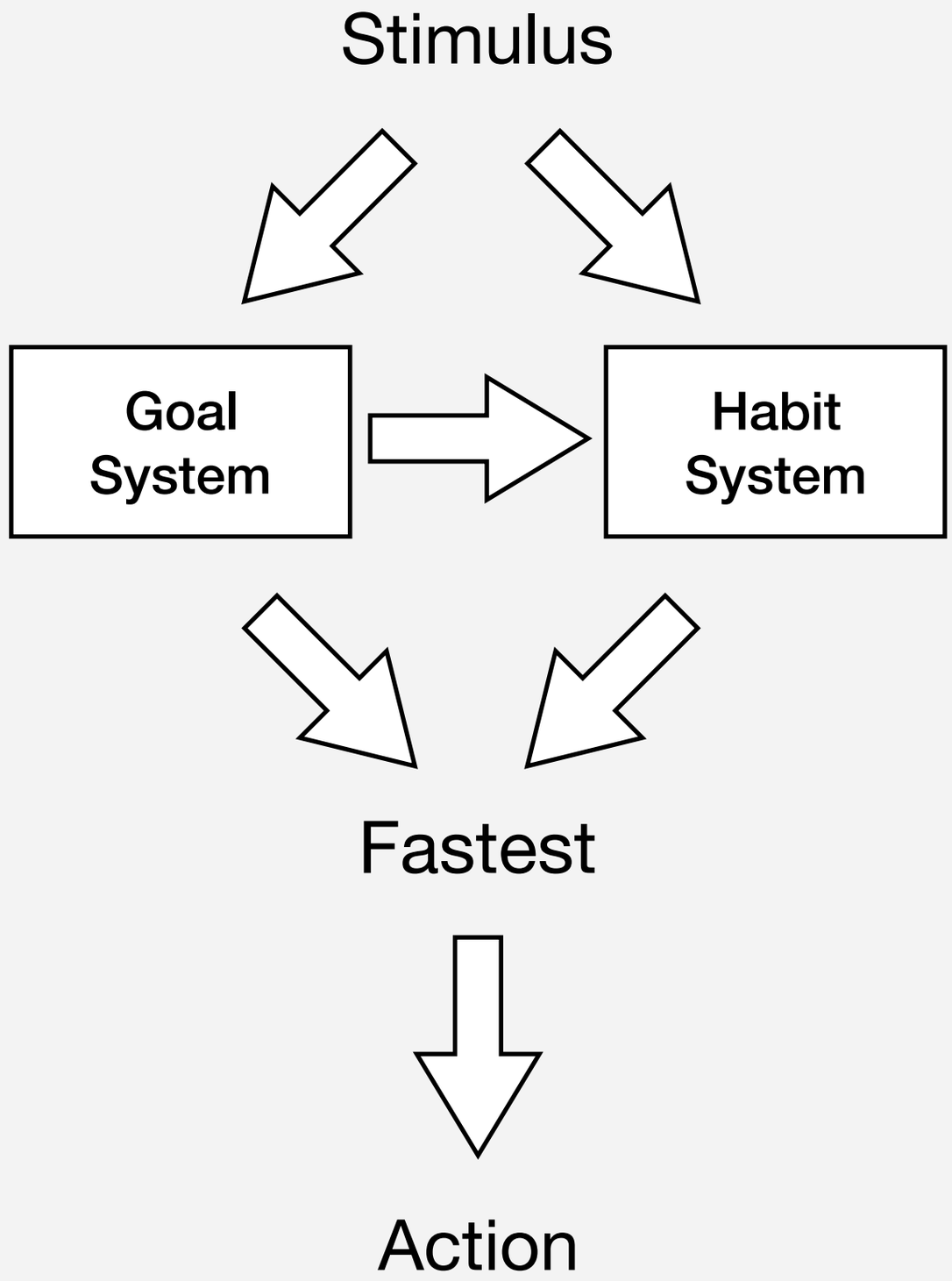
Dezfouli et al. (2013)

"… Model-based reinforcement learning (RL) has been argued to underlie the goal-directed process; however, the way in which it interacts with habits and the structure of the habitual process has remained unclear. According to a flat architecture, the habitual process corresponds to model-free RL, and its interaction with the goal-directed process is coordinated by an external arbitration mechanism. Alternatively, **the interaction between these systems has recently been argued to be hierarchical**, such that the formation of action sequences underlies habit learning and a goal-directed process selects between goal-directed actions and habitual sequences of actions to reach the goal…"



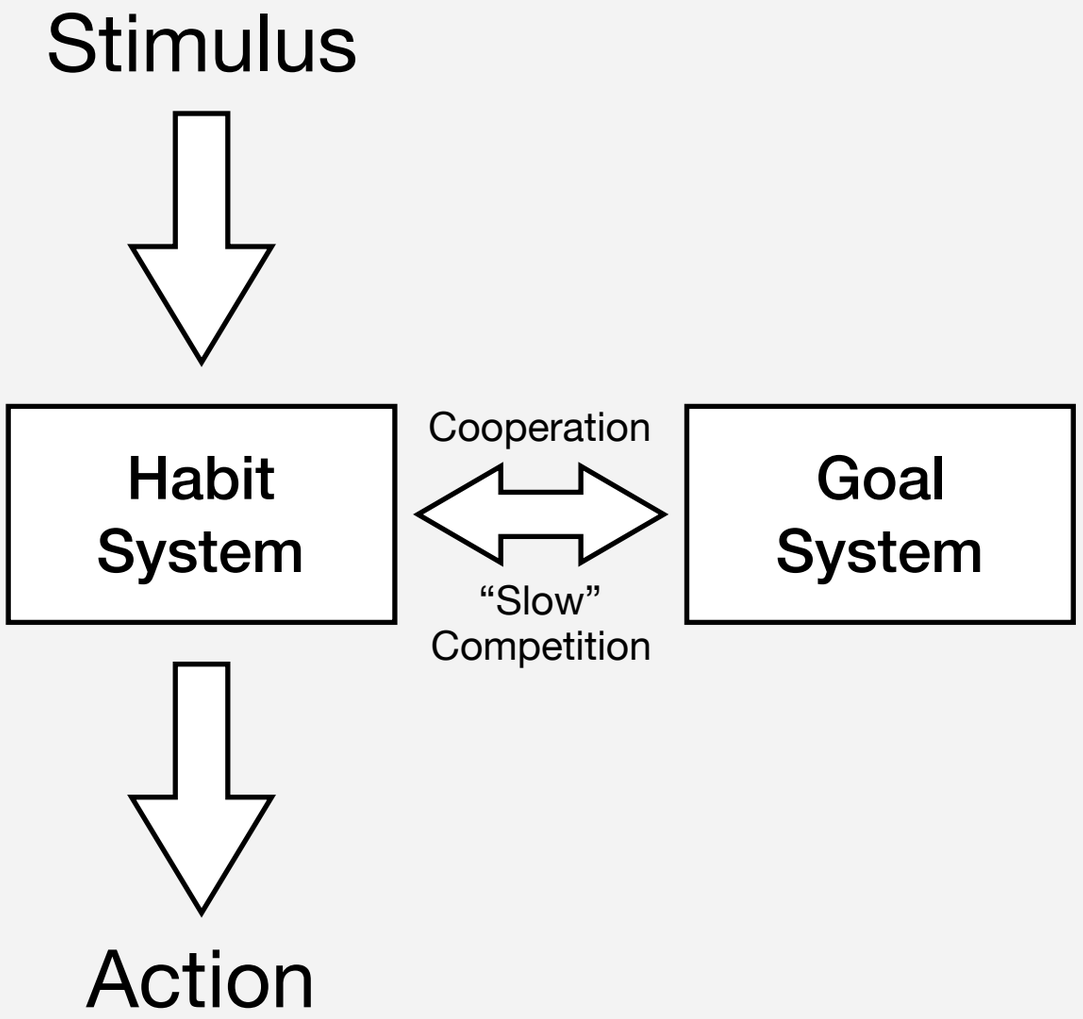
Ashby et al. (2007)

"… The model assumes 2 neural pathways from sensory association cortex to the premotor area that mediates response selection. **A longer and slower path projects to the premotor area via the striatum, globus pallidus, and thalamus. A faster, purely cortical path projects directly to the premotor area.** The model assumes that the subcortical path has greater neural plasticity because of a dopamine-mediated learning signal from the substantia nigra. In contrast, the cortical-cortical path learns more slowly via (dopamine independent) Hebbian learning…"



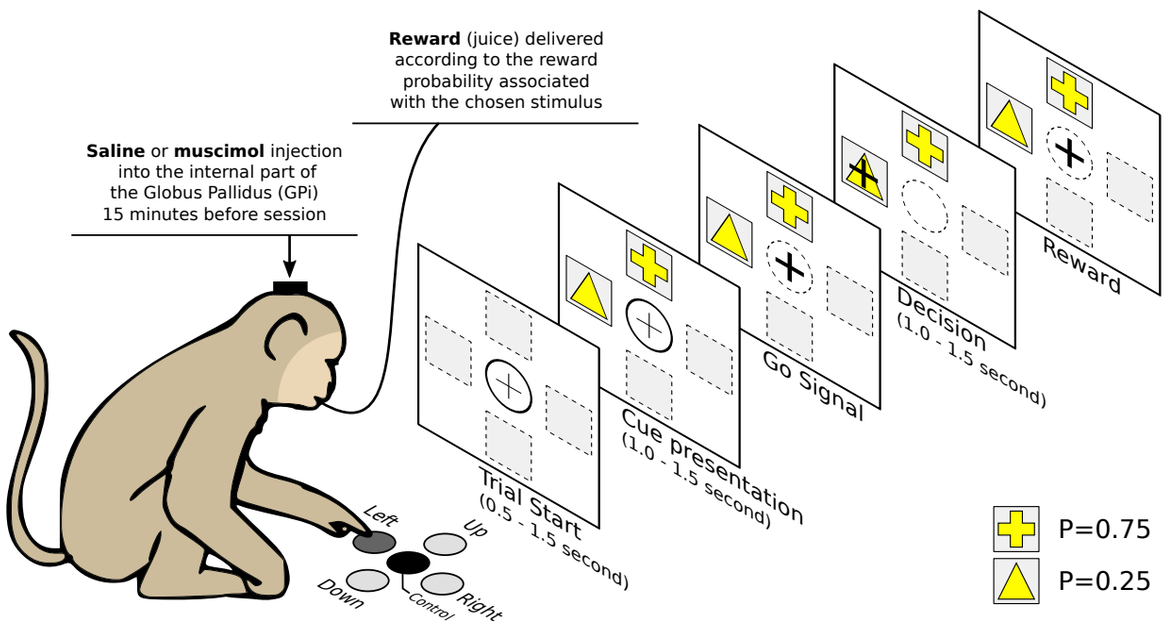
Piron et al. (2016)

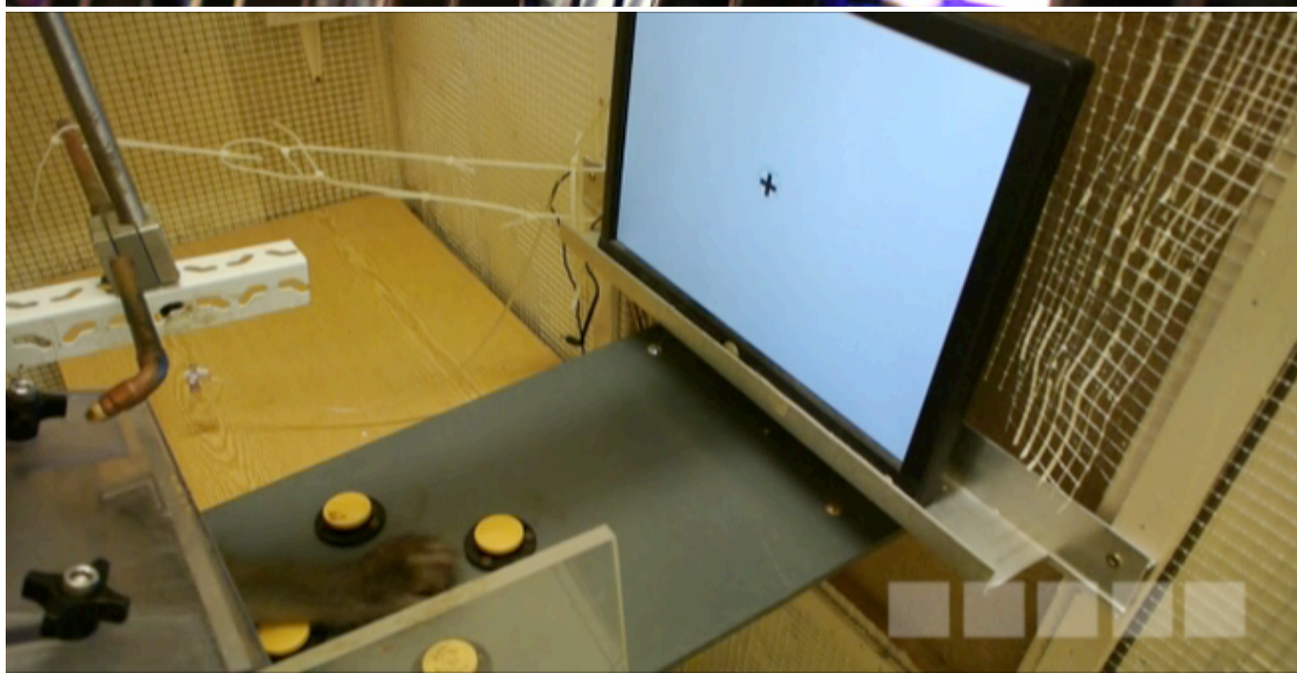
"… Said differently, we managed to explicitly dissociate reinforcement learning from Hebbian learning and demonstrated covert learning inside the basal ganglia. These results suggest that a behavioral decision results from **both the cooperation (acquisition) and competition (expression) of two distinct but entangled memory systems**, the goal-directed system and the habit system that may represent the two ends of the same graded phenomenon.



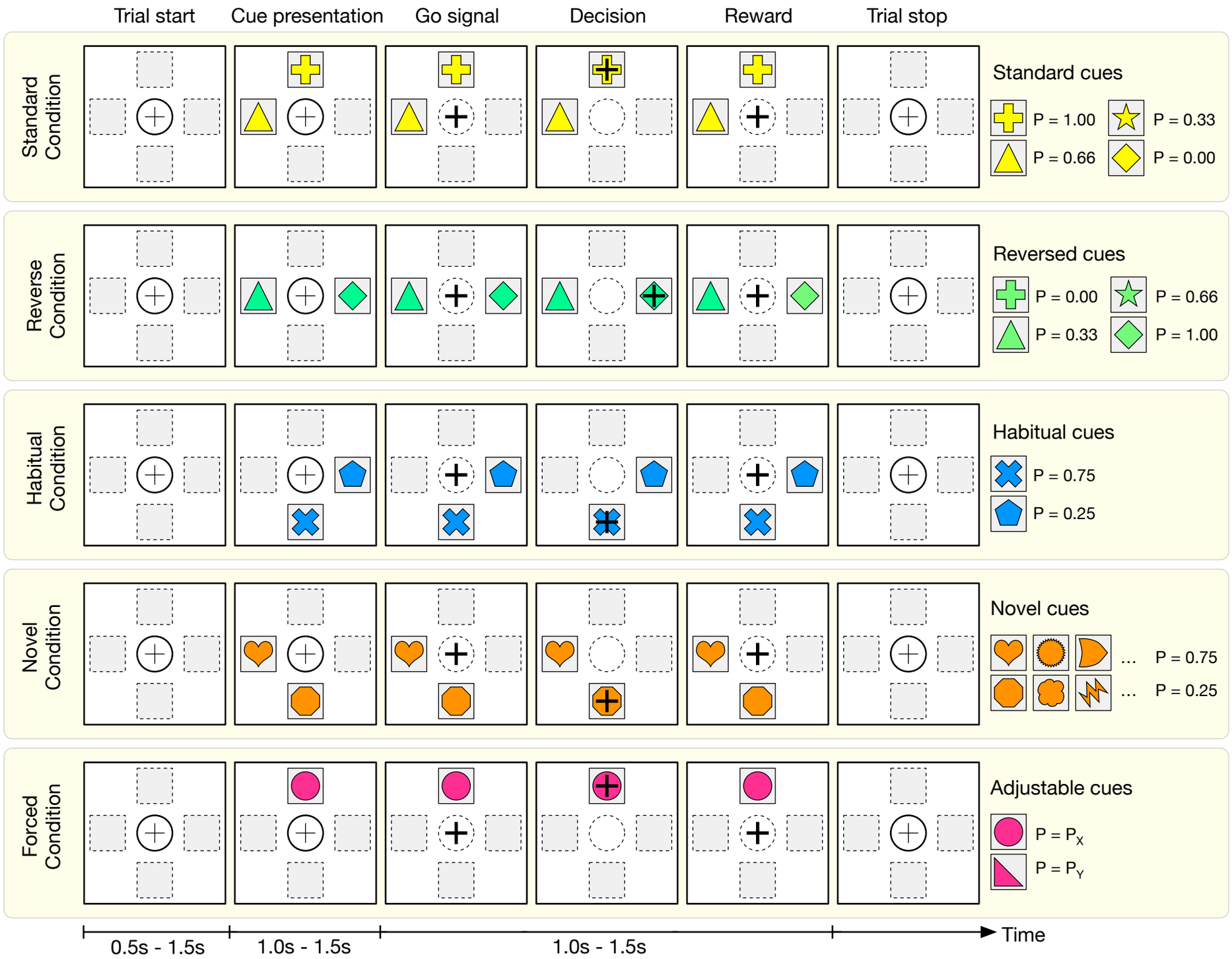
Two-armed bandit

→ Humans
→ Monkeys
→ Rodents
→ Birds (Krebs et al. 1978)
→ Fish (Thomas et al. 1985)
→ Bees (Keasar et al. 2002)
→ Slime mould (Reid et al. 2016)
→ Photon (Naruse et al. 2015)





Two-armed bandit



A long series…

LEBLOIS ET AL. (2006) → MISSING IN ACTION (FEW LINES OF C)

→ GUTHRIE ET AL. (2013) → DEAD (6000 LINES OF DELPHI)

→ TOPALIDOU ET AL. (2015) (200 LINES OF PYTHON) → PIRON ET AL. (2016)

PIRON ET AL. (2016) → TOPALIDOU ET AL. (IN PREP) → …

TOPALIDOU ET AL. (2015) → We redo Science !

We redo Science ! → ESCOBAR ET AL., 2016
NALLAPU ET AL., 2016
CARREIRE ET AL., 2015

RESCIENCE.GITHUB.IO

REGISTER AS REVIEWER
AND YOU CAN GAIN
FAME, FORTUNE, SUCCESS
OR ...
A NICE STICKER !
(YES ! A STICKER !)

BENOÎT G.
RESCIENCE CHANGED MY LIFE

MEHDI K.
IT REALLY WORKS !

XAVIER H.
TOTALLY WORTH IT !

OPEN DATA
OPEN SOURCE
OPEN PEER-REVIEW
OPEN (GREEN) ACCESS
NO "BUZZ" BARRIER
COMMUNITY SUPPORTED
0€ BUDGET

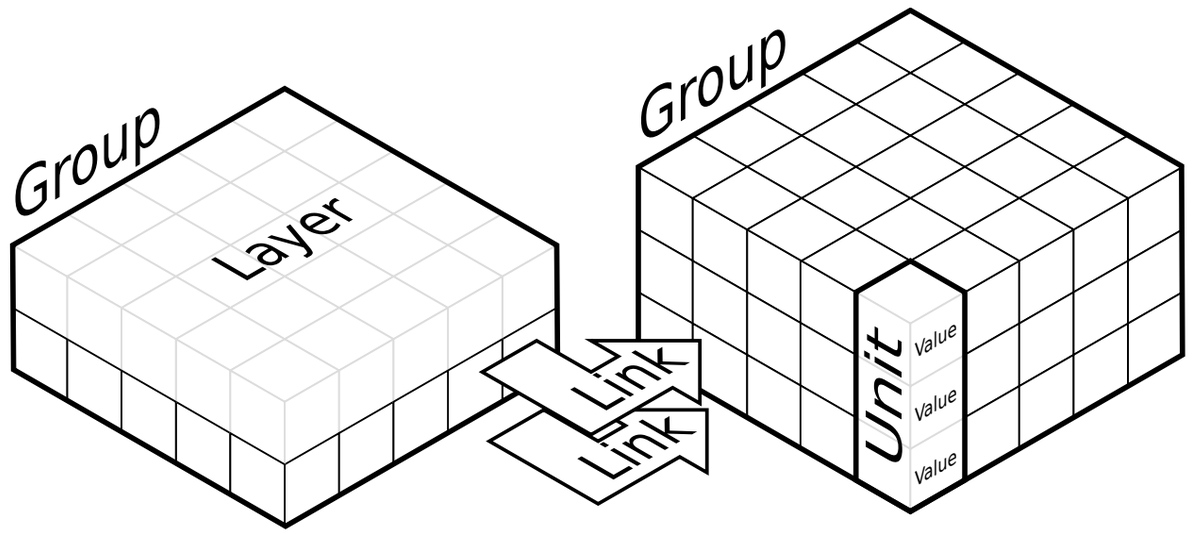
ReScience

Reproducible science is good. Replicated science is better.

DANA framework

A unit is a set of arbitrary values that can vary along time under the influence of other units and learning.

- Distributed
  → no supervisor
- Asynchronous
  → no central clock
- Numerical
  → no symbol
- Adaptative
  → to learn something



We want to make sure that emerging properties are those of the model and not those of the software running the model.

Computational model

Cortex
- Posterior
- Motor / Premotor
- Prefrontal

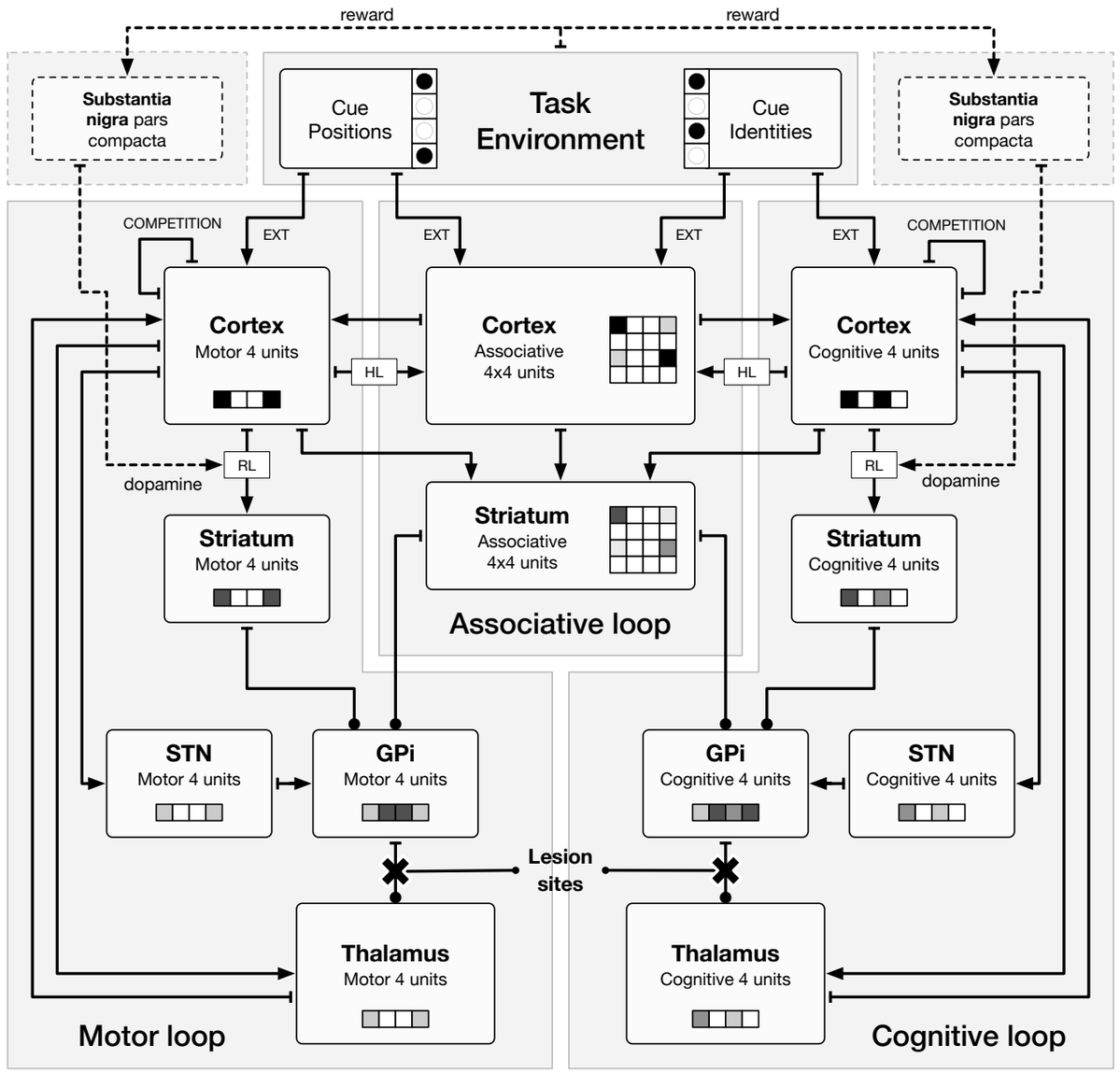
Thalamus

Striatum (STR)
- Caudate
- Putamen
- Nucleus Accumbens

Subthalamic Nucleus (STN)

Globus Pallidus
- Internal (GPi)
- External (GPe)

Subtantia Nigra
- pars Compacta (SNc)
- pars Reticulata (SNr)



Computational model

Dopamine + Reinforcement learning

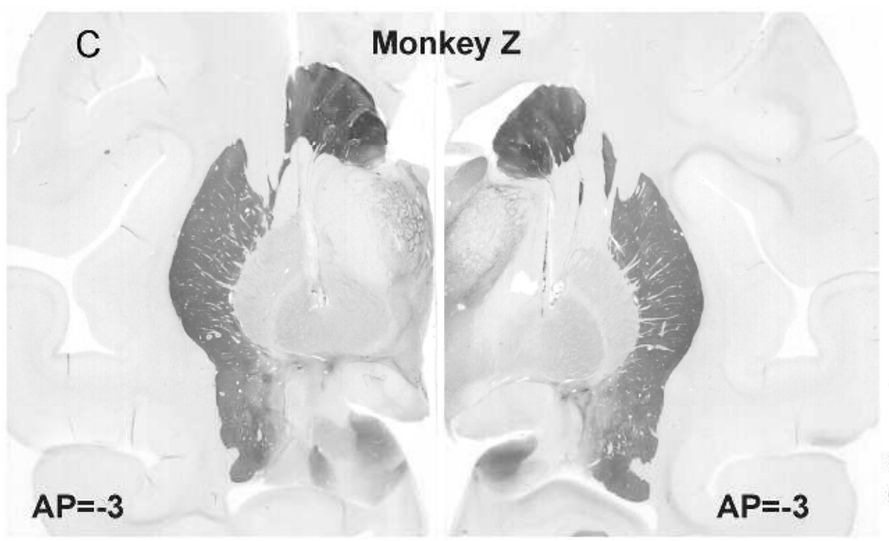
- Cognitive cortex to cognitive striatum
- Motor cortex to motor striatum

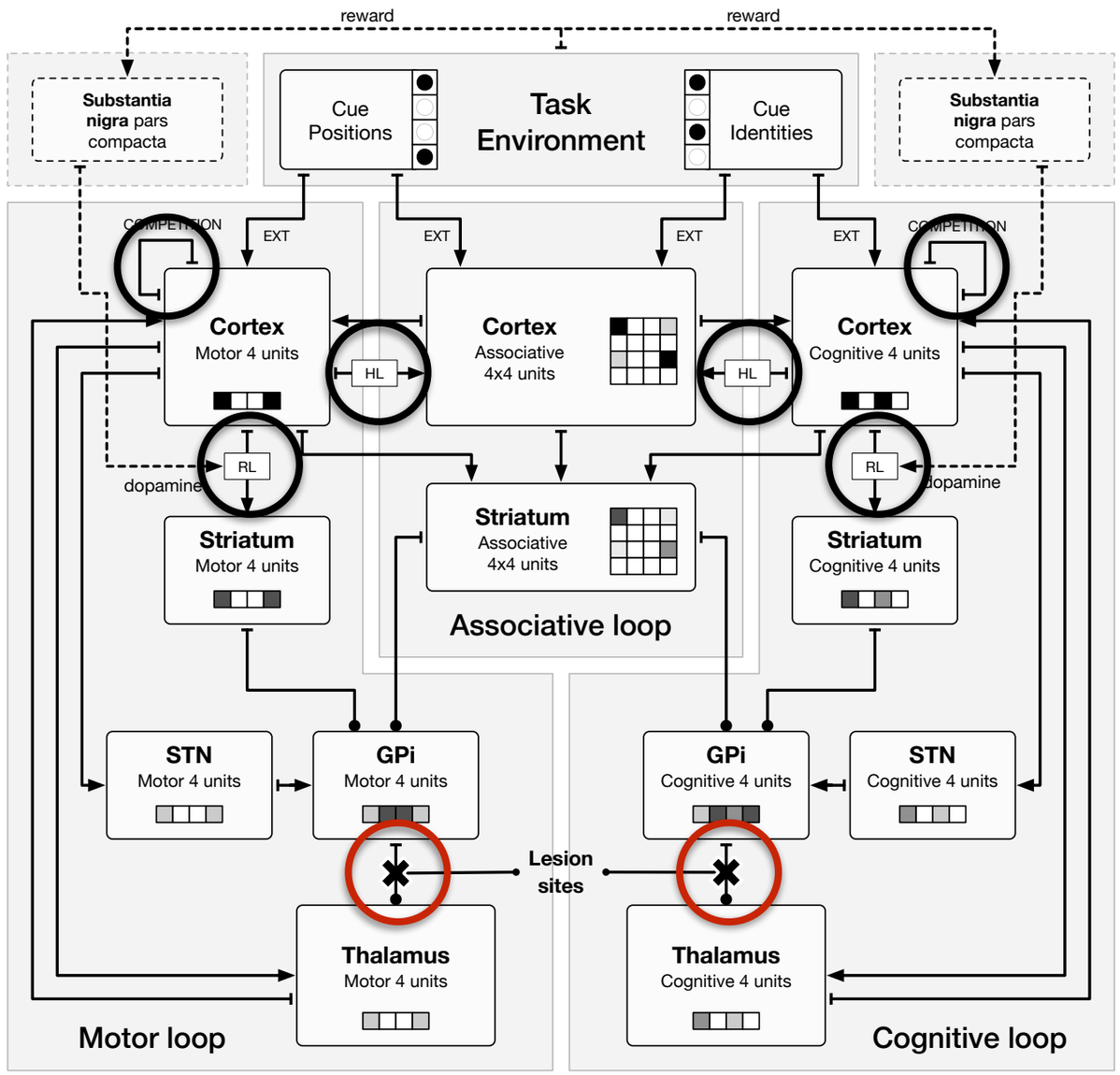
Lateral Competition + Hebbian learning

- Cognitive cortex to associative cortex
- Motor cortex to associative cortex

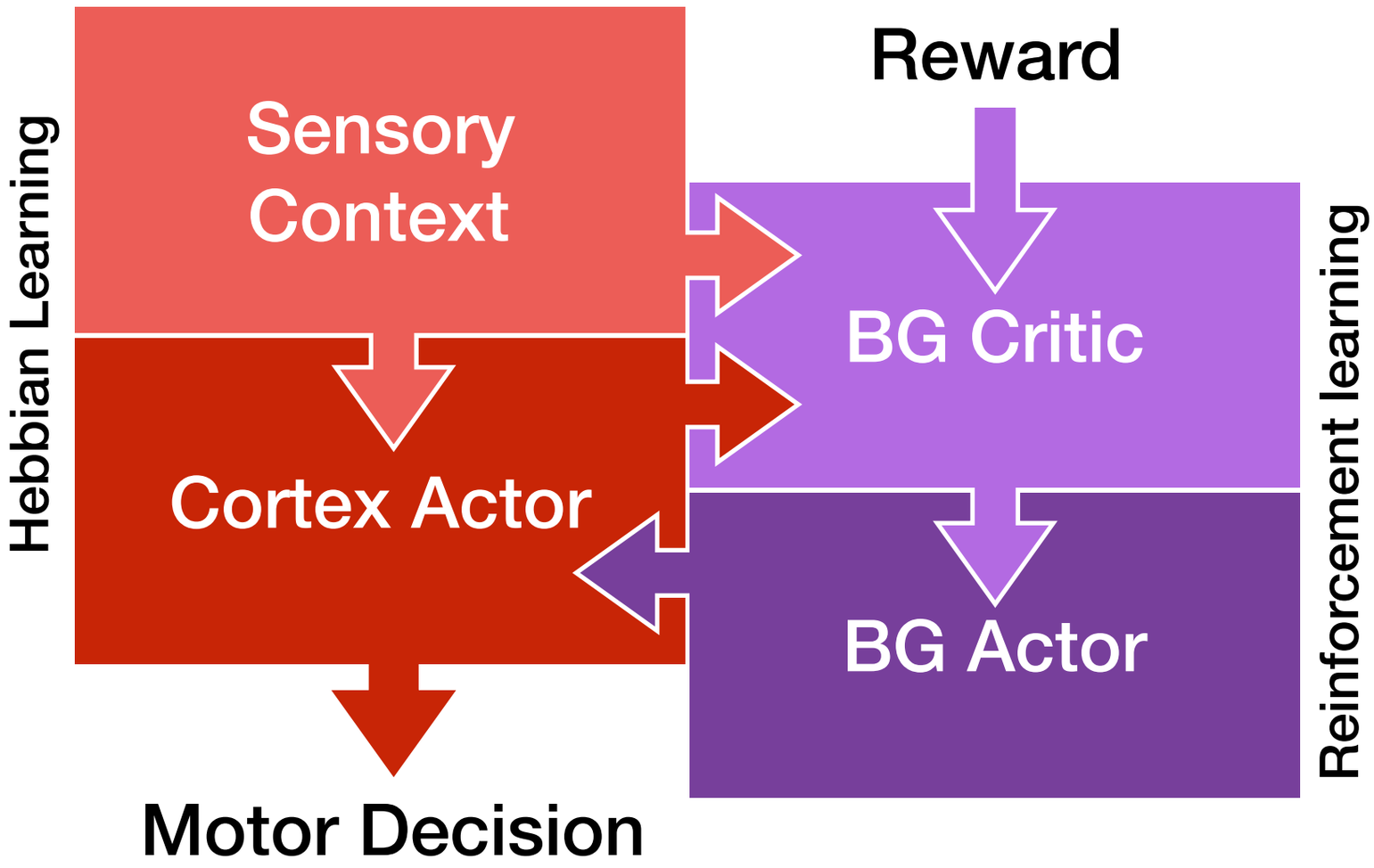
Lesion

- Motor GPi to motor thalamus
- Cognitive GPi to cognitive thalamus





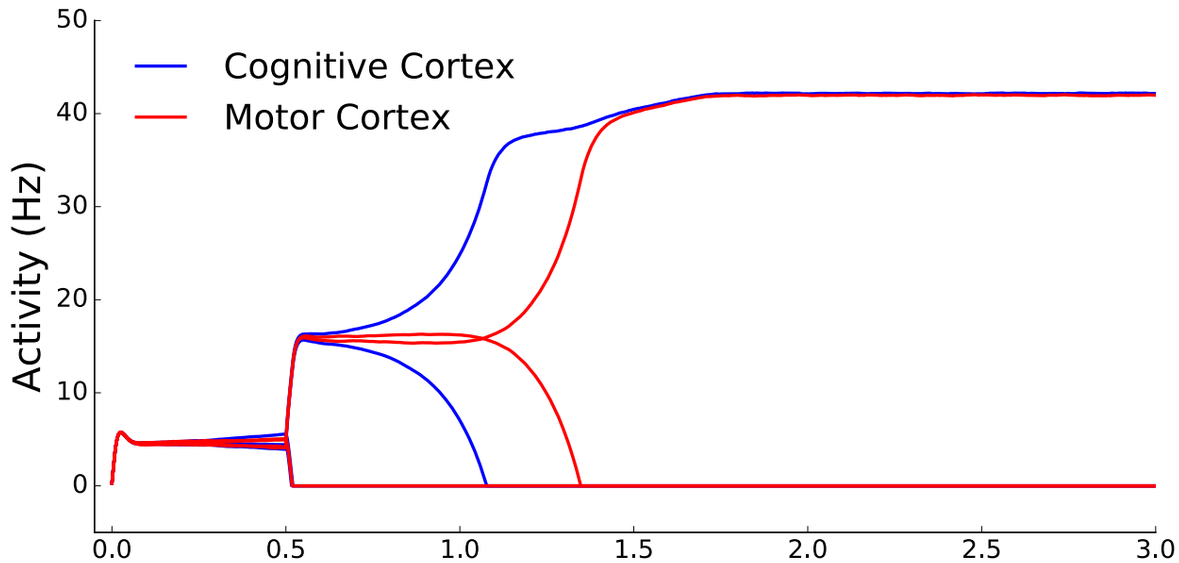
Habit learning



Habit learning

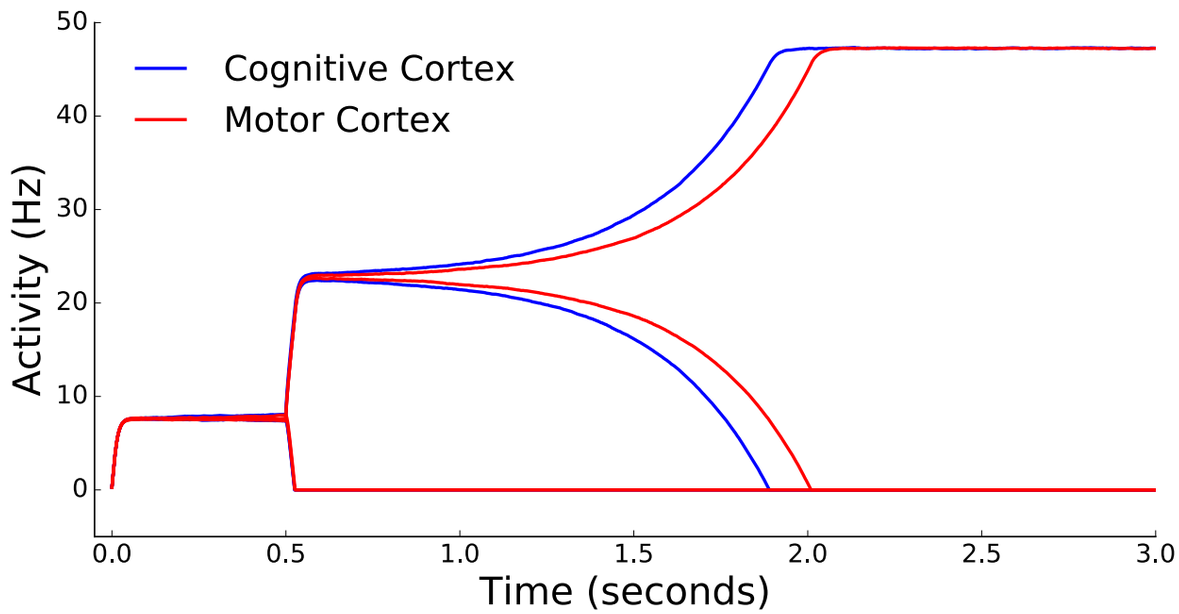
Intact model (GPi On)

Faster decision (before learning)

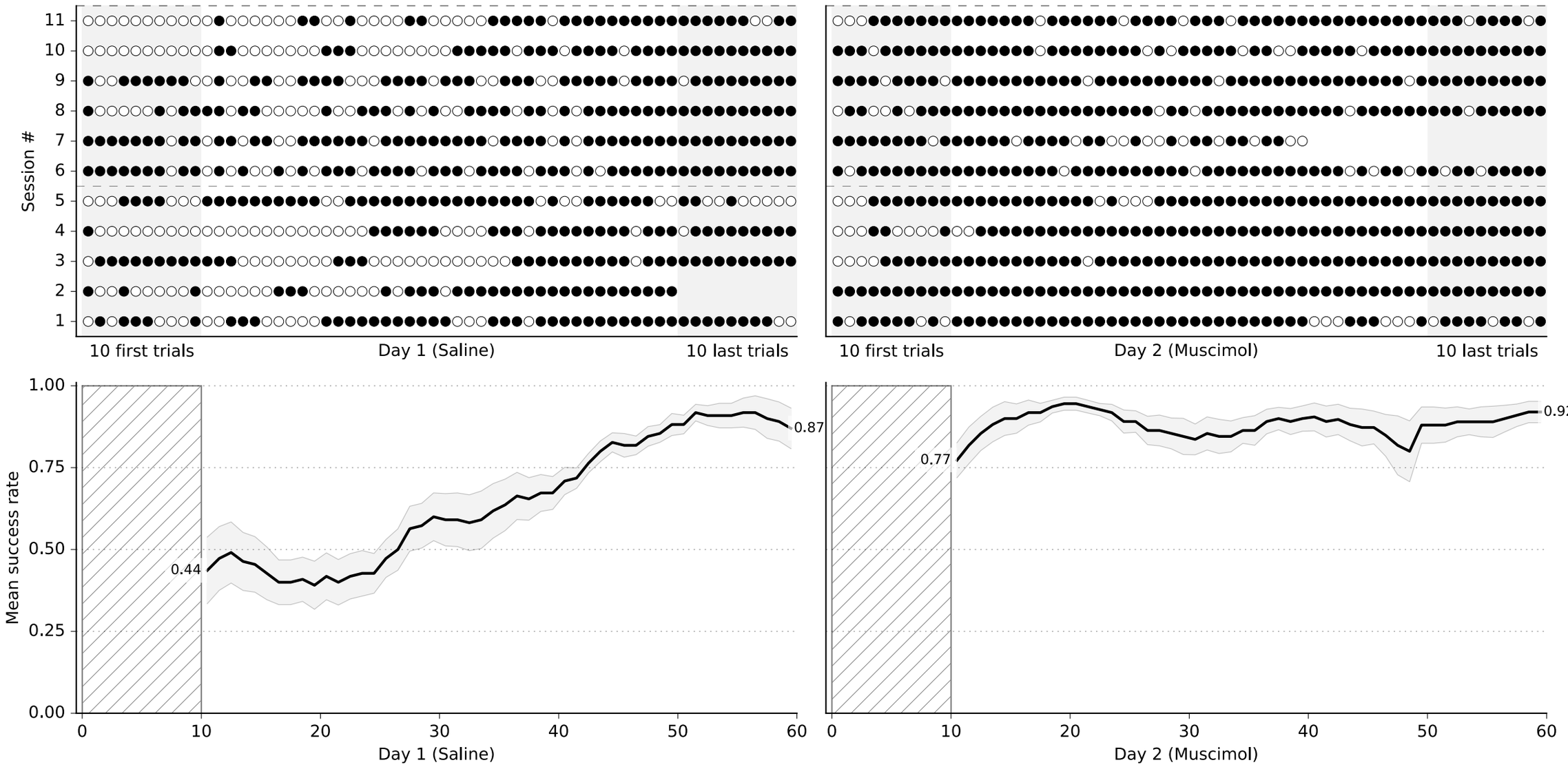


Lesioned model (GPi Off)

Slower decision (before learning)



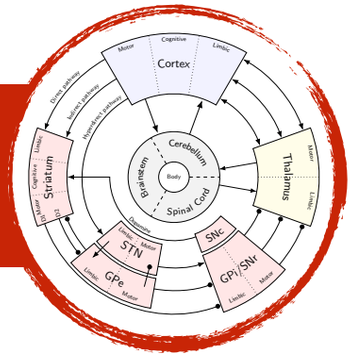
Habit learning

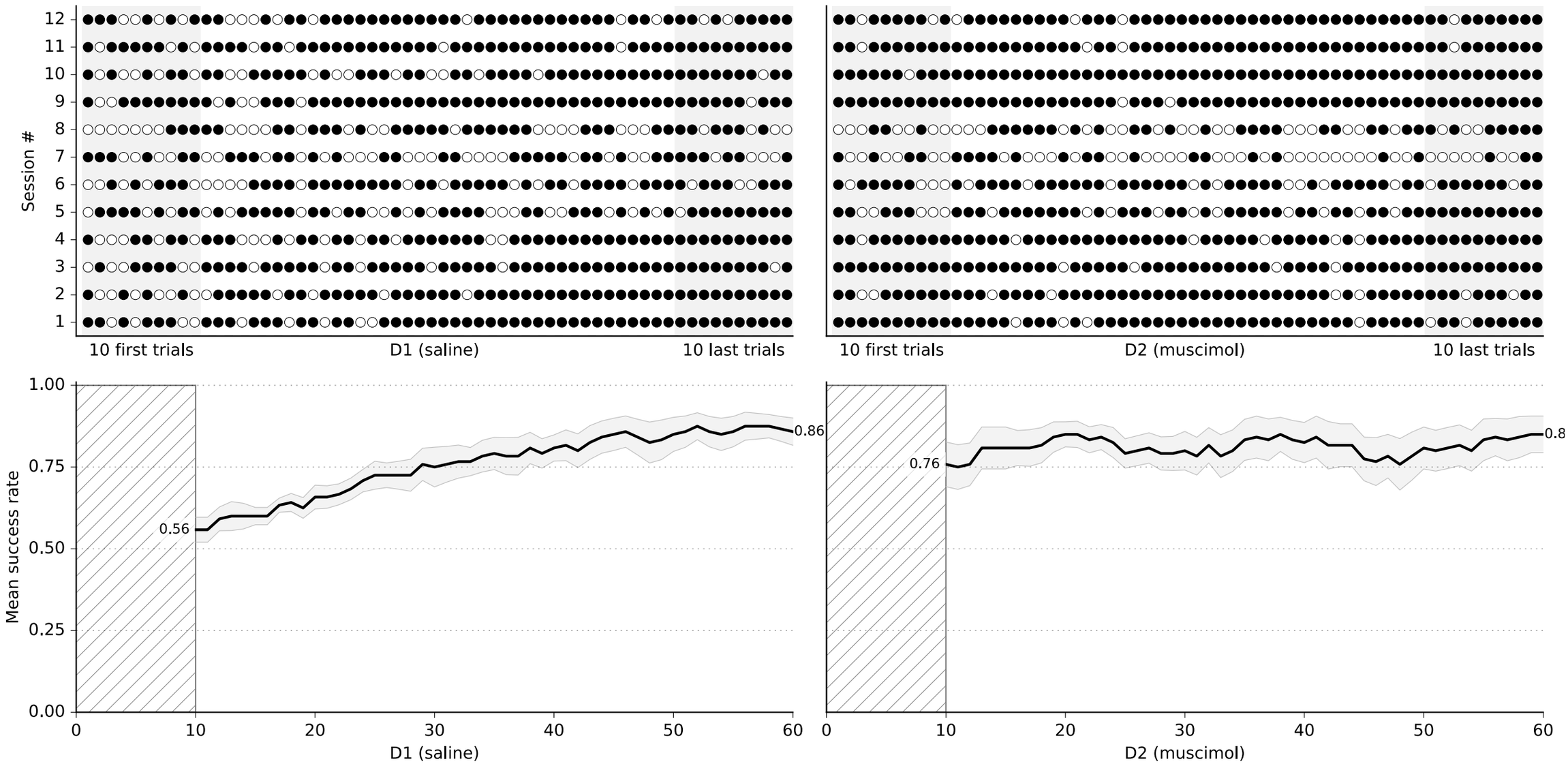


Saline / GPi ON

Muscimol / GPi OFF

Habit learning

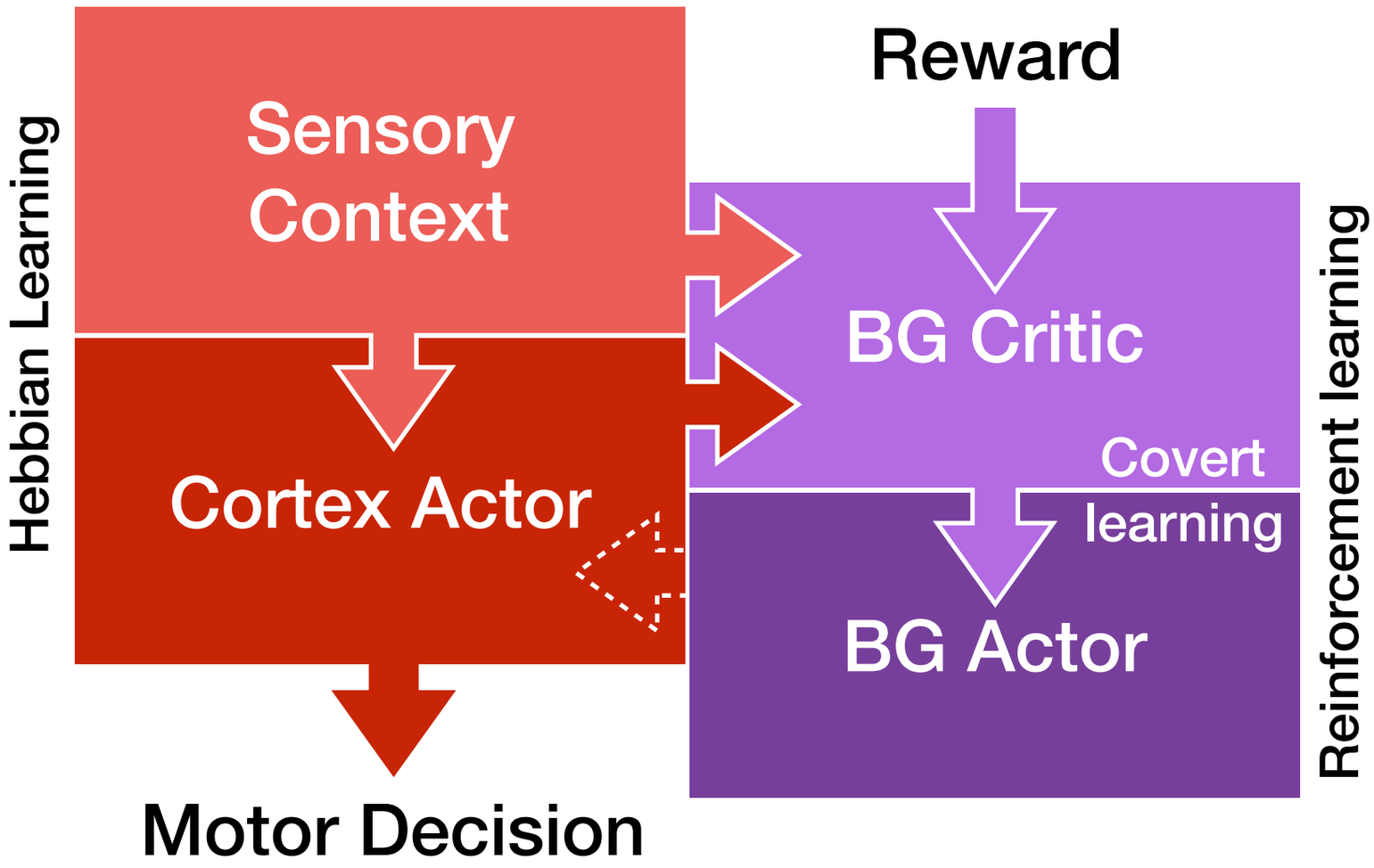




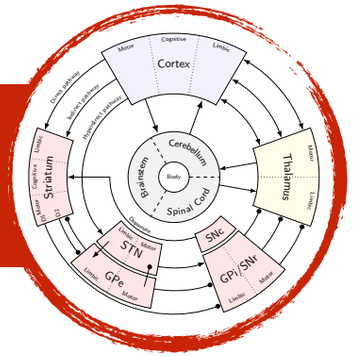
Saline / GPi ON

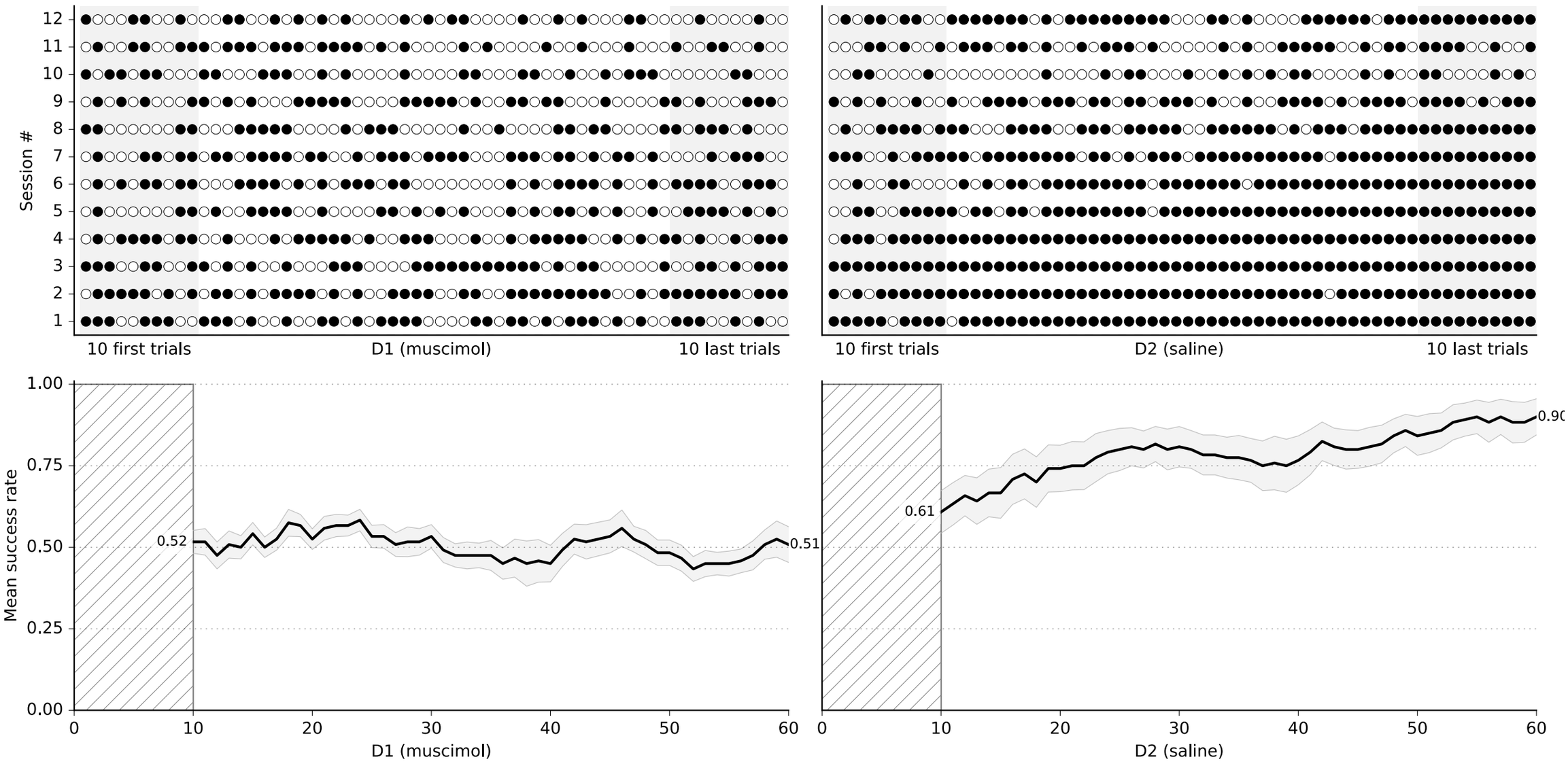
Muscimol / GPi OFF

Covert learning



Covert learning

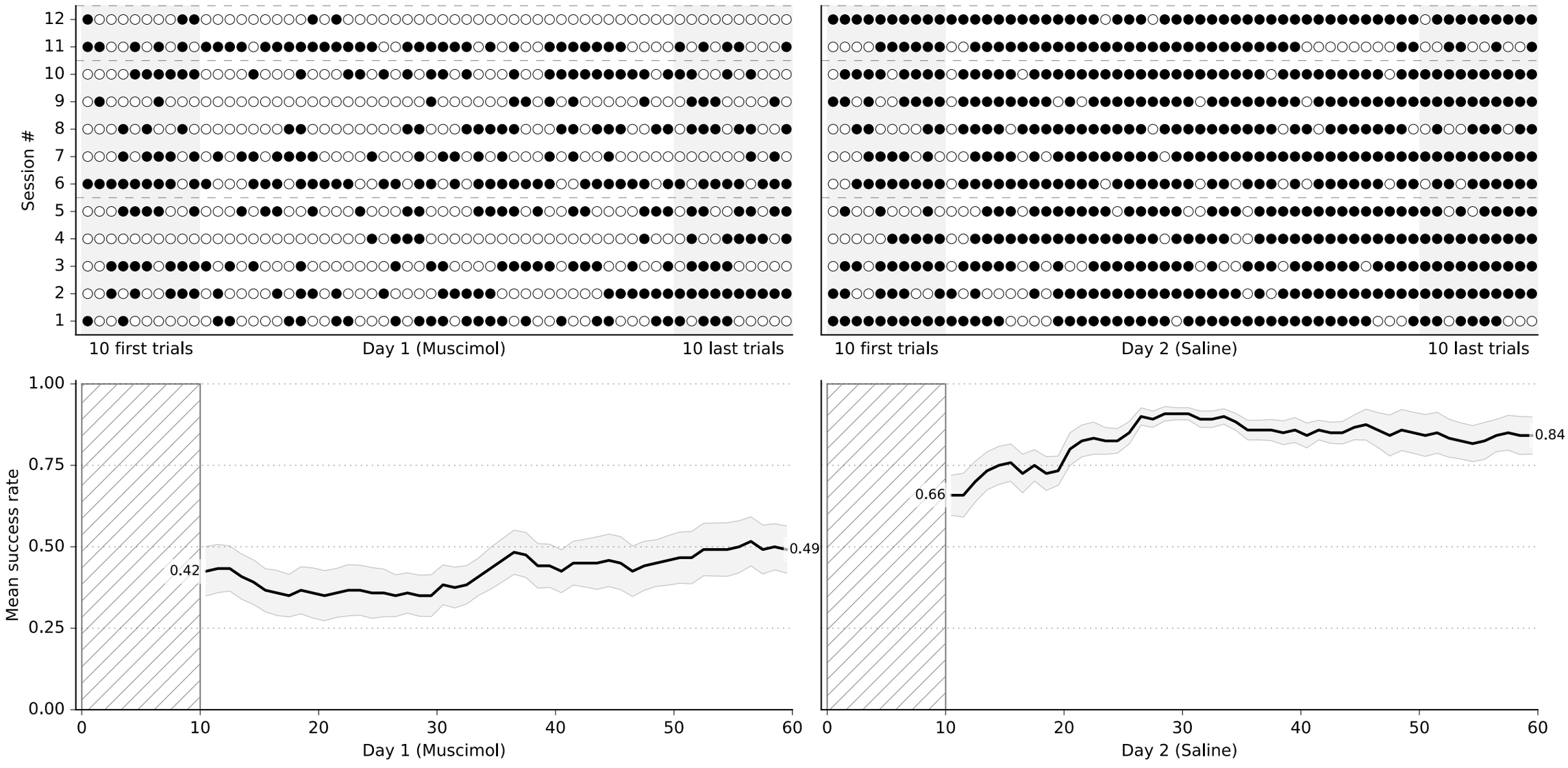




Muscimol / GPi OFF

Saline / GPI ON

Covert learning



Muscimol / GPi OFF

Saline / GPI ON

Conclusion

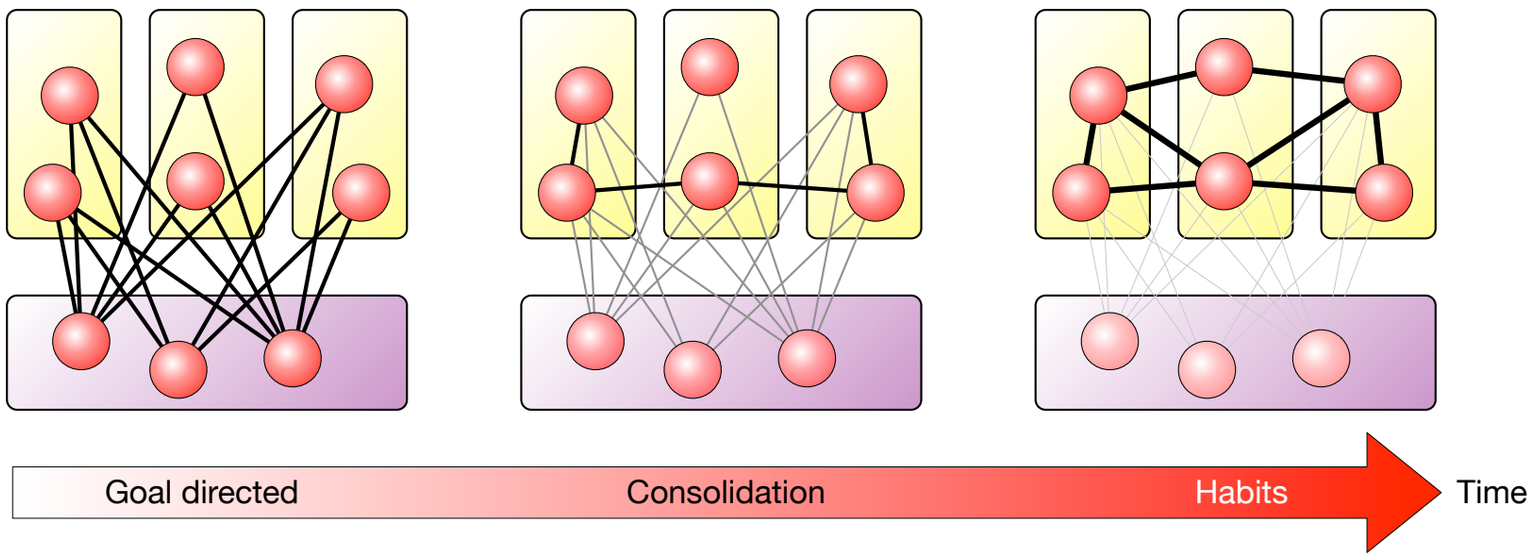
Habit acquisition and habit expression

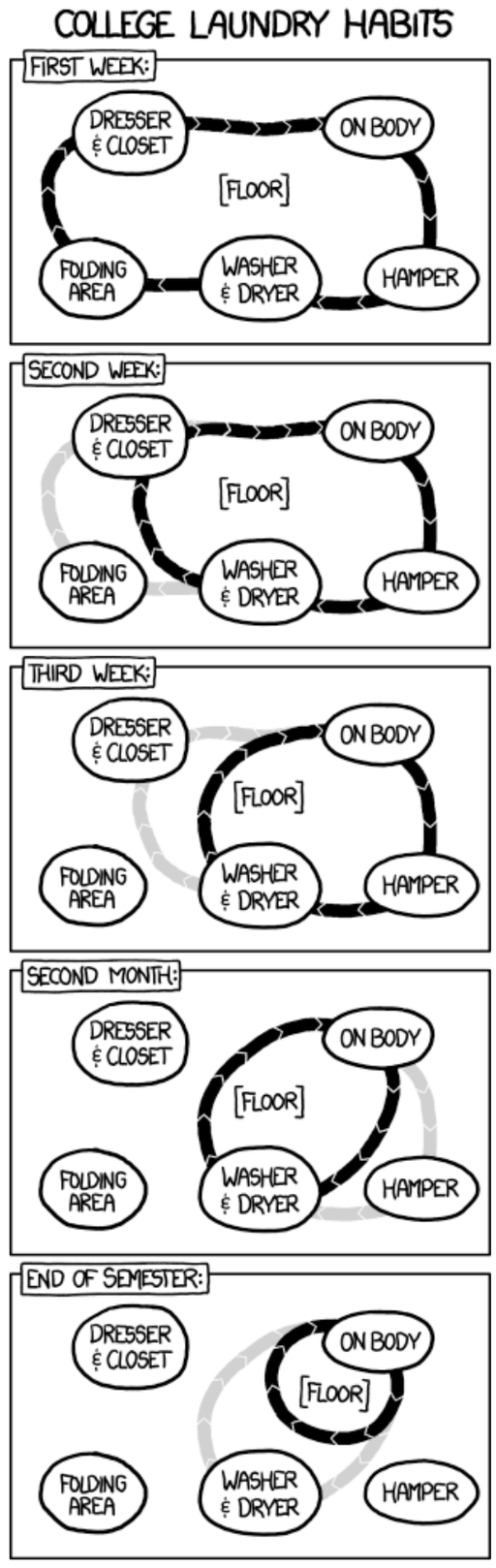
→ These are two different processes even though they're entangled
→ Basal ganglia serves as an implicit supervisor
→ Habit can be expressed outside BG (at least in the primate)

The critic role of the BG

→ Basal ganglia serves as a generic critic, for any "actor"
→ No experimental evidence yet for the role of the cortex
→ Ongoing experiments to measure RL vs HL influence on behavior

Habits are a graded phenomenon





XKCD #1066

Questions ?

NICOLAS ROUGIER, INRIA
ROBOTICS & NEUROSCIENCES, BORDEAUX, 2016
HTTP://WWW.LABRI.FR/PERSO/NROUGIER/
PLAYING NOVEMBER 17TH, 2016


M. TOPALIDOU


T. BORAUD


C. PIRON


D. KASE

NICOLAS ROUGIER, INRIA
ROBOTICS & NEUROSCIENCES, BORDEAUX, 2016
HTTP://WWW.LABRI.FR/PERSO/NROUGIER/
PLAYING NOVEMBER 17TH, 2016

M. TOPALIDOU

T. BORAUD

C. PIRON

D. KASE

Questions ?

NICOLAS ROUGIER, INRIA
ROBOTICS & NEUROSCIENCES, BORDEAUX, 2016
HTTP://WWW.LABRI.FR/PERSO/NROUGIER/
PLAYING NOVEMBER 17TH, 2016


M. TOPALIDOU


T. BORAUD


C. PIRON


D. KASE

Questions ?

NICOLAS ROUGIER, INRIA
ROBOTICS & NEUROSCIENCES, BORDEAUX, 2016
HTTP://WWW.LABRI.FR/PERSO/NROUGIER/
PLAYING NOVEMBER 17TH, 2016


M. TOPALIDOU


T. BORAUD


C. PIRON


D. KASE