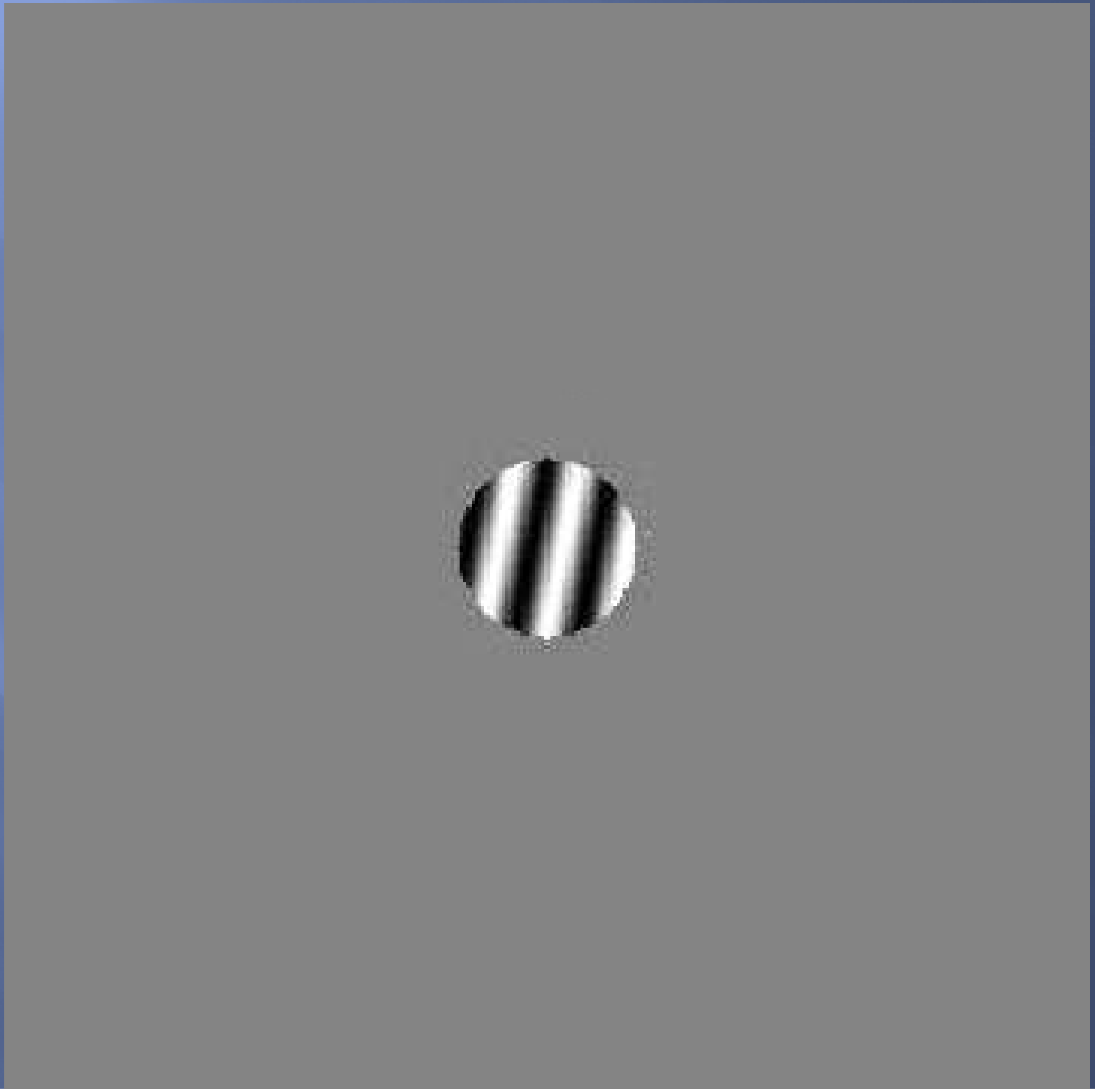
The Neurobiology of Perceptual Categorization: From Learning to Automaticity

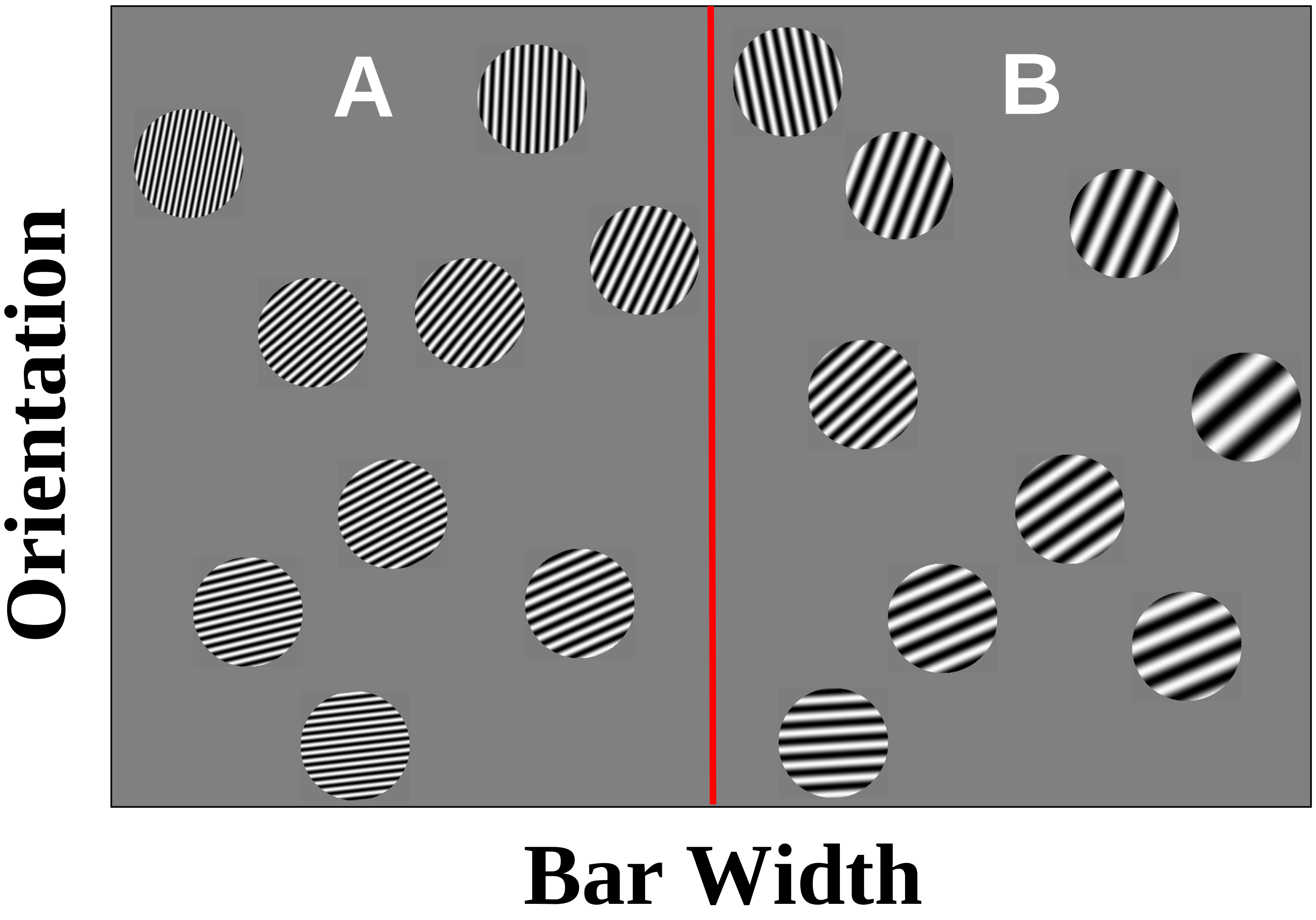
F. Gregory Ashby

Laboratory for Computational Cognitive Neuroscience
University of California, Santa Barbara

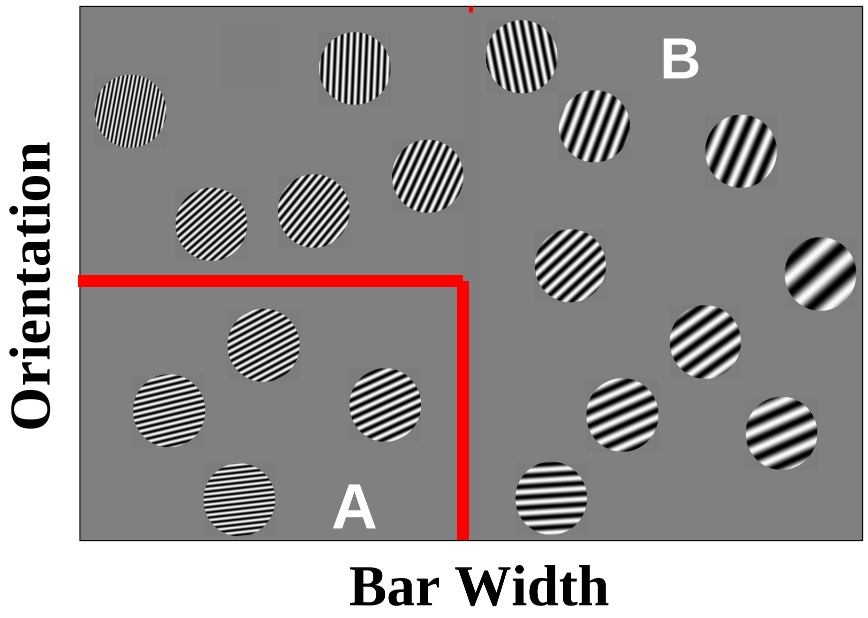
STIMULUS ON A SINGLE CATEGORY-LEARNING TRIAL



RULE-BASED CATEGORY LEARNING



Rule-Based Category Learning



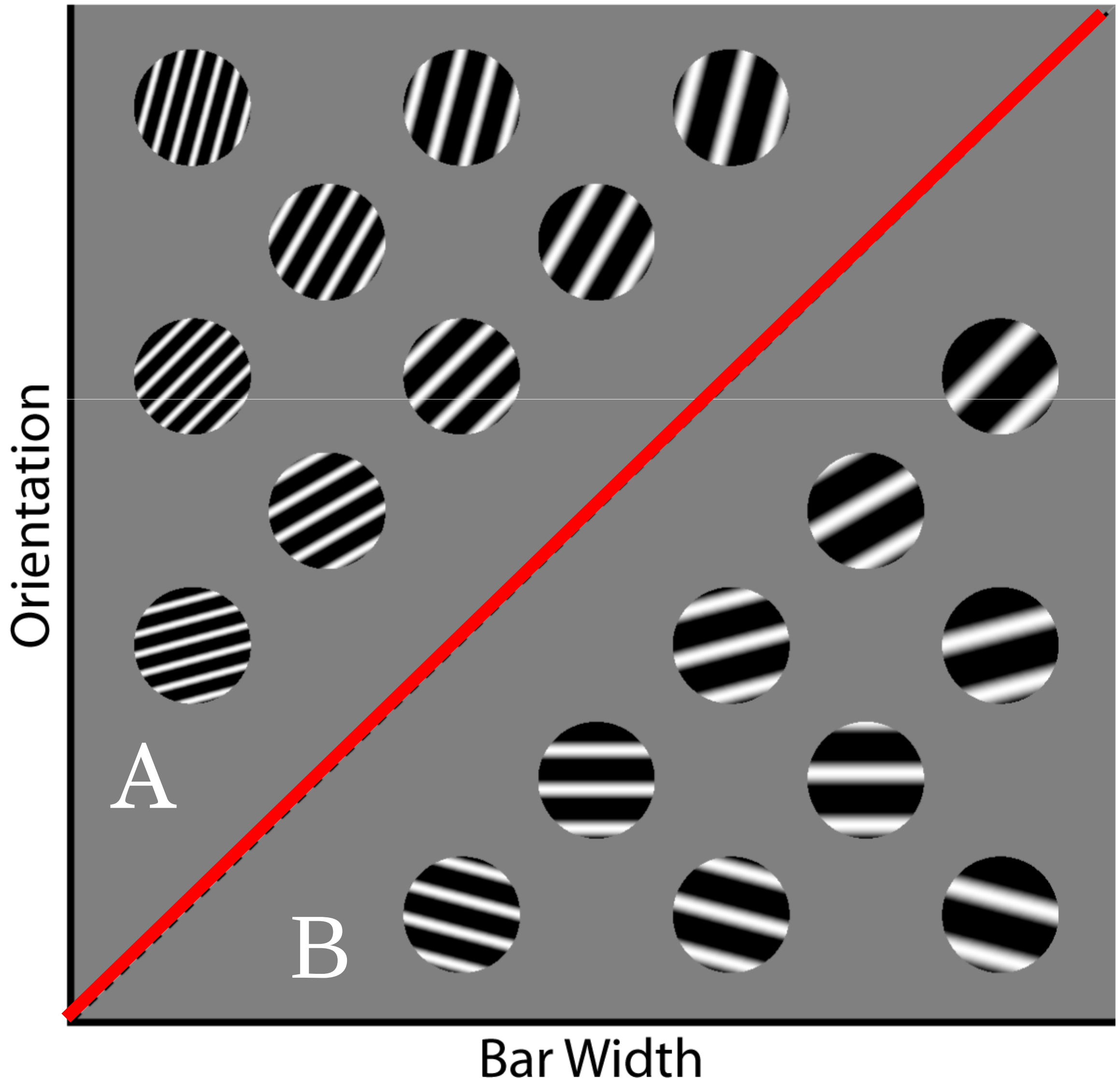
Categorization rule is easy to describe

Effective learning requires:

- no distractions
- active and effortful processing of feedback

But the nature and timing of feedback is not critical

Information-Integration Category Learning



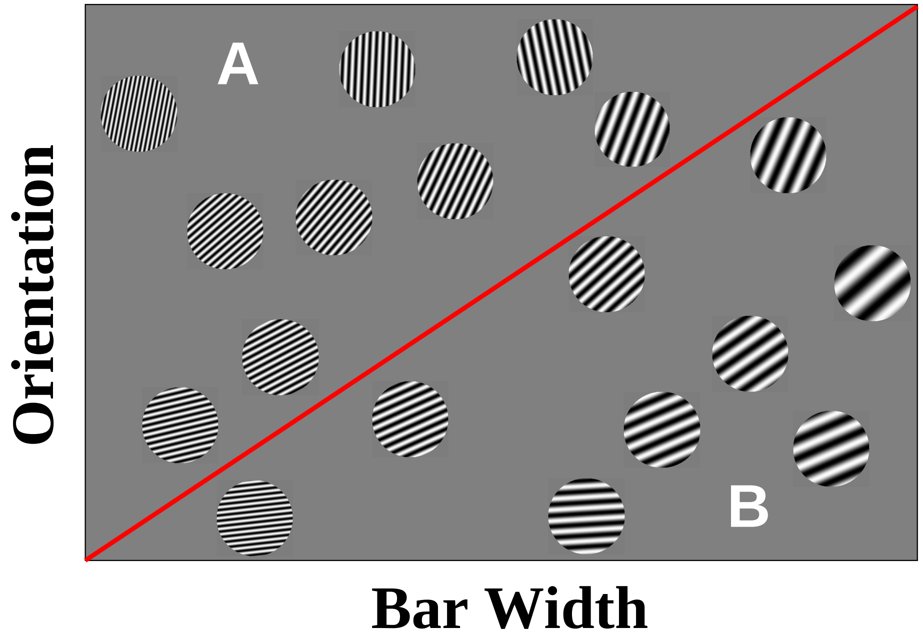
A REAL-LIFE II TASK?

Does this mammogram show a tumor?

i.e., is it in the category "tumor" or the category "nontumor"?

A REAL-LIFE II TASK?

Tumor!



Information-Integration Category Learning

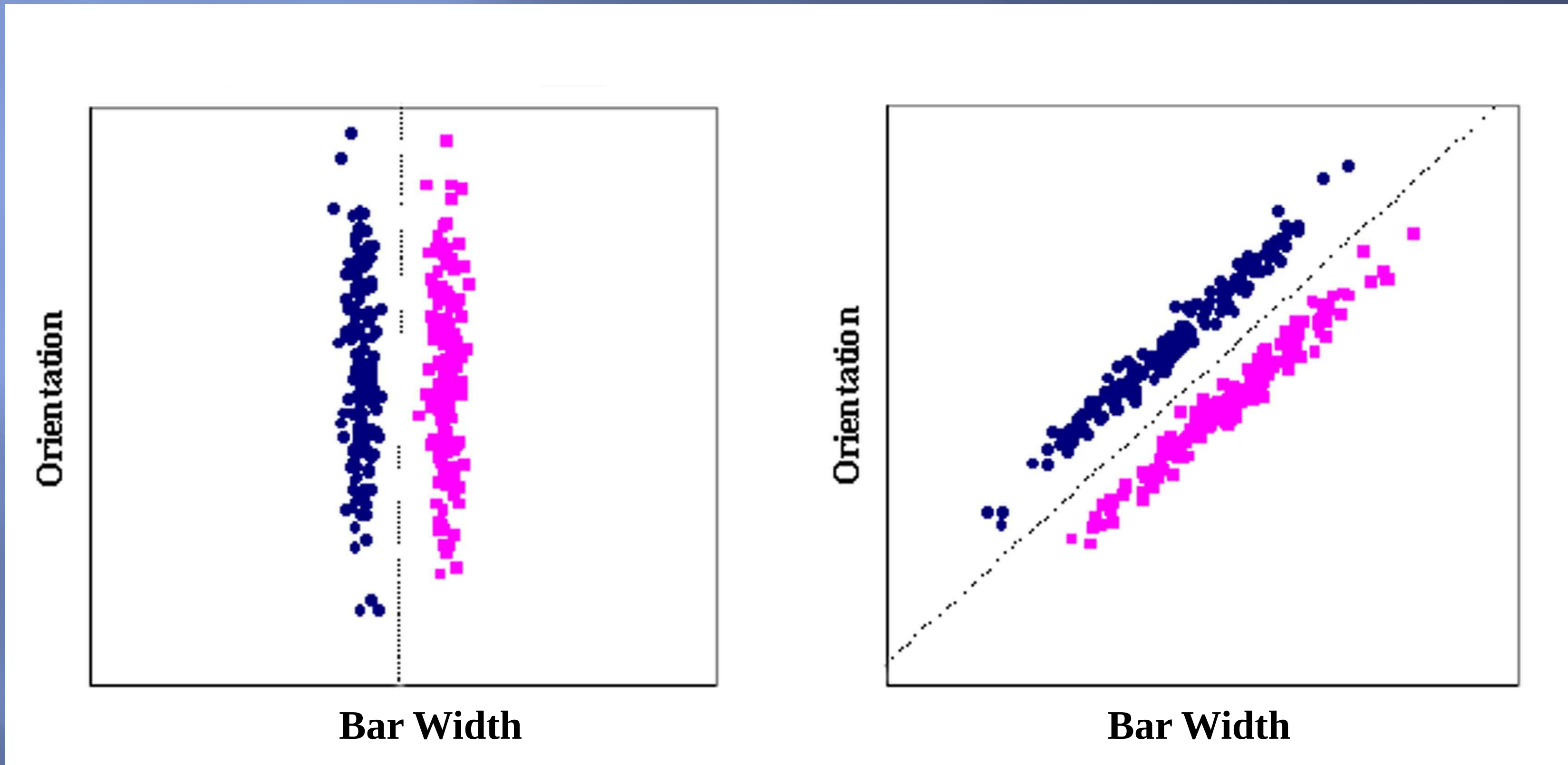
Categorization rule is difficult to describe

Effective learning requires:

- consistent feedback immediately after response
- consistent mapping from category to response location
- no active feedback processing

Categories

Rule Based | Information Integration



Is the information-integration task inherently more difficult?

THE TWO CATEGORY LEARNING SYSTEMS OF COVIS

(Ashby, Alfonso-Reese, Turken, & Waldron, *Psychological Review*, 1998)

- explicit, logical-reasoning system
  -- quickly learns explicit rules

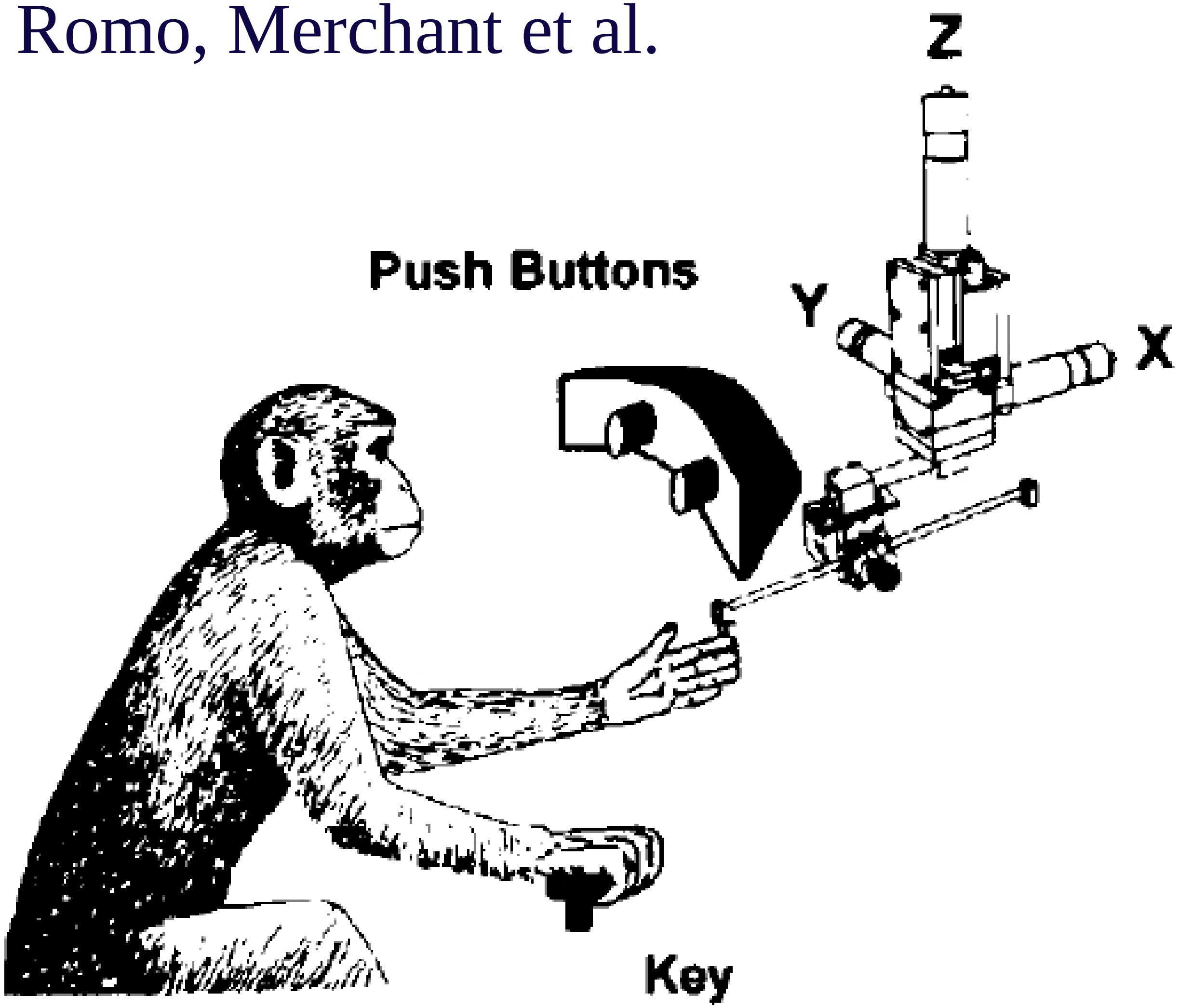
- procedural- or habit-learning system
  -- slowly learns similarity-based rules

- simultaneously active in all tasks (at least initially)

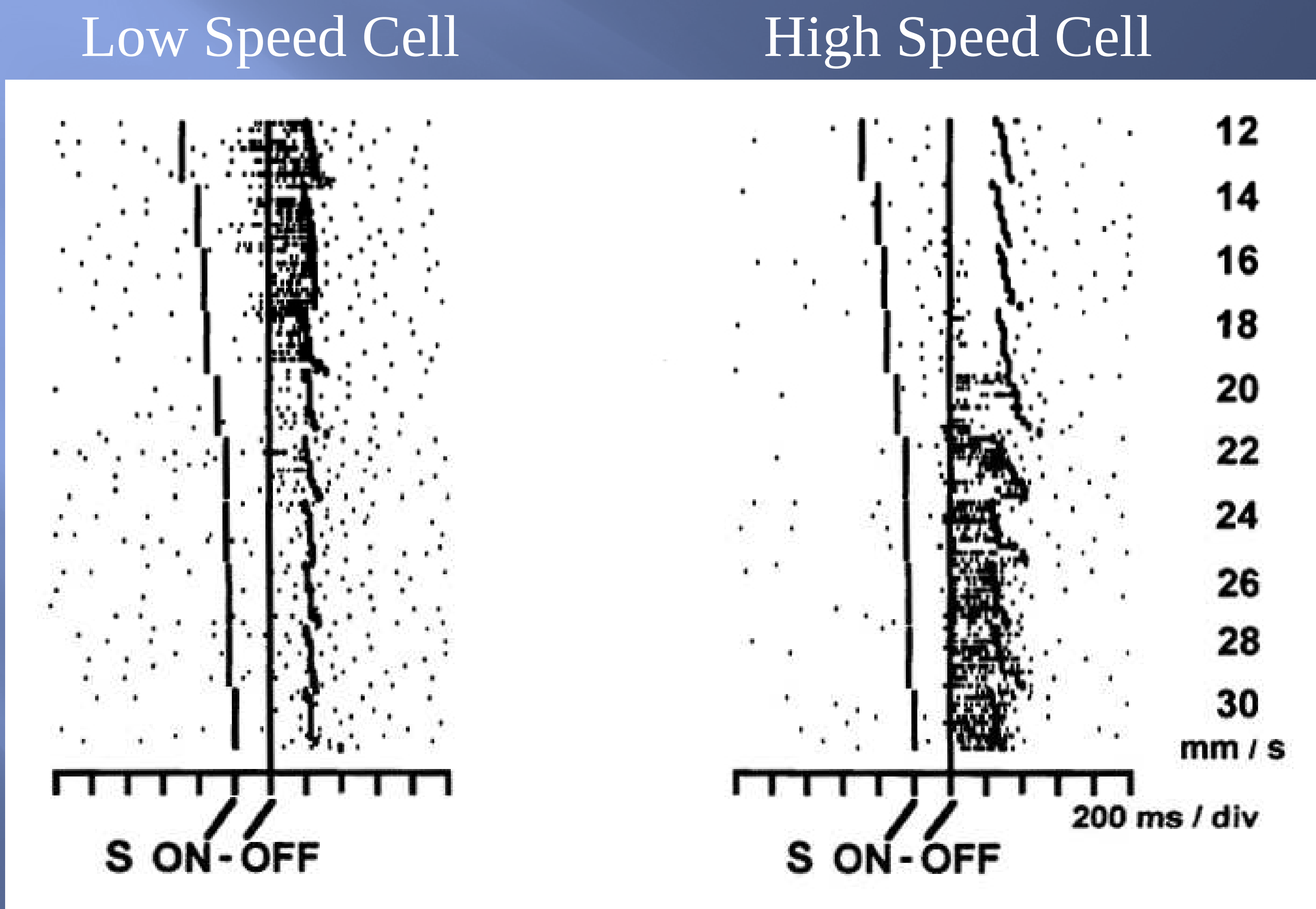
The Caudate Nucleus

Tactile Category Learning

Romo, Merchant et al.



Single Cell Responses – Putamen



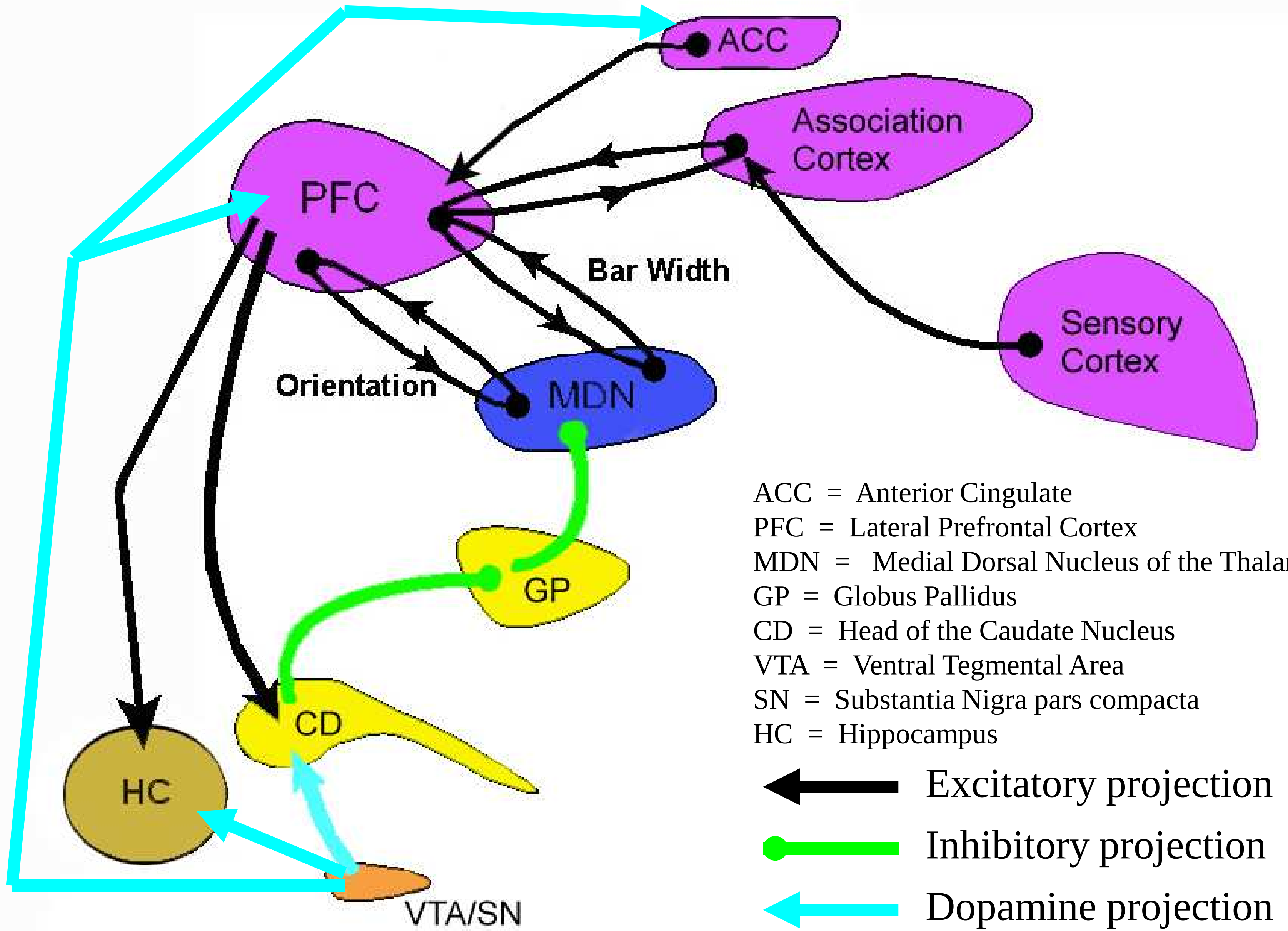
Merchant et al. (1997, J. of Neurophysiology)

THE COVIS EXPLICIT SYSTEM

- logical reasoning system
- uses working memory and executive attention
- prefrontal cortex, anterior cingulate, head of the caudate nucleus, thalamo-cortical loops, medial temporal lobe structures

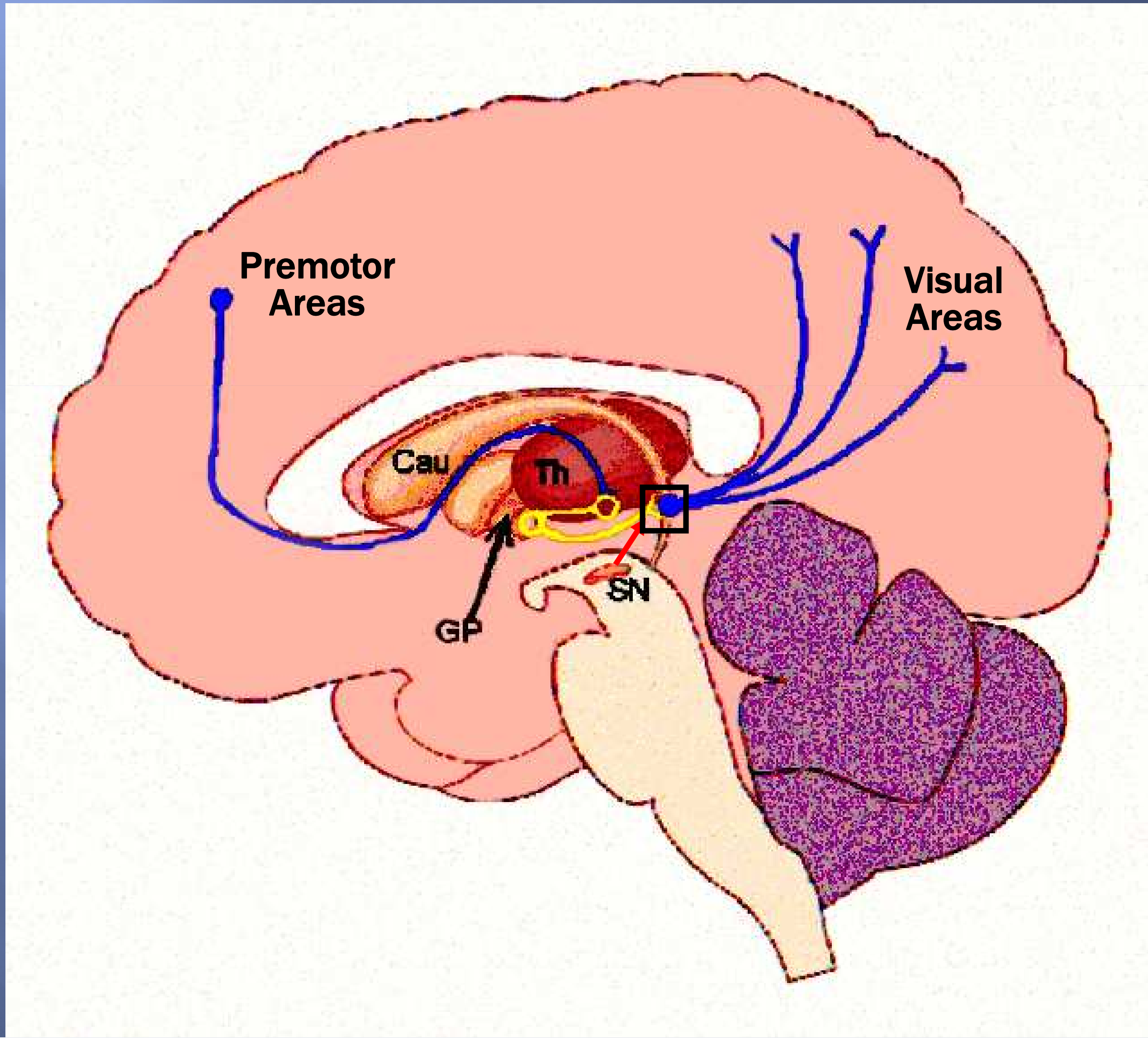
- Working memory & attentional switching component – FROST (Ashby, Ell, Valentin, & Casale, 2005, *J. of Cognitive Neuroscience*)

The COVIS Explicit System

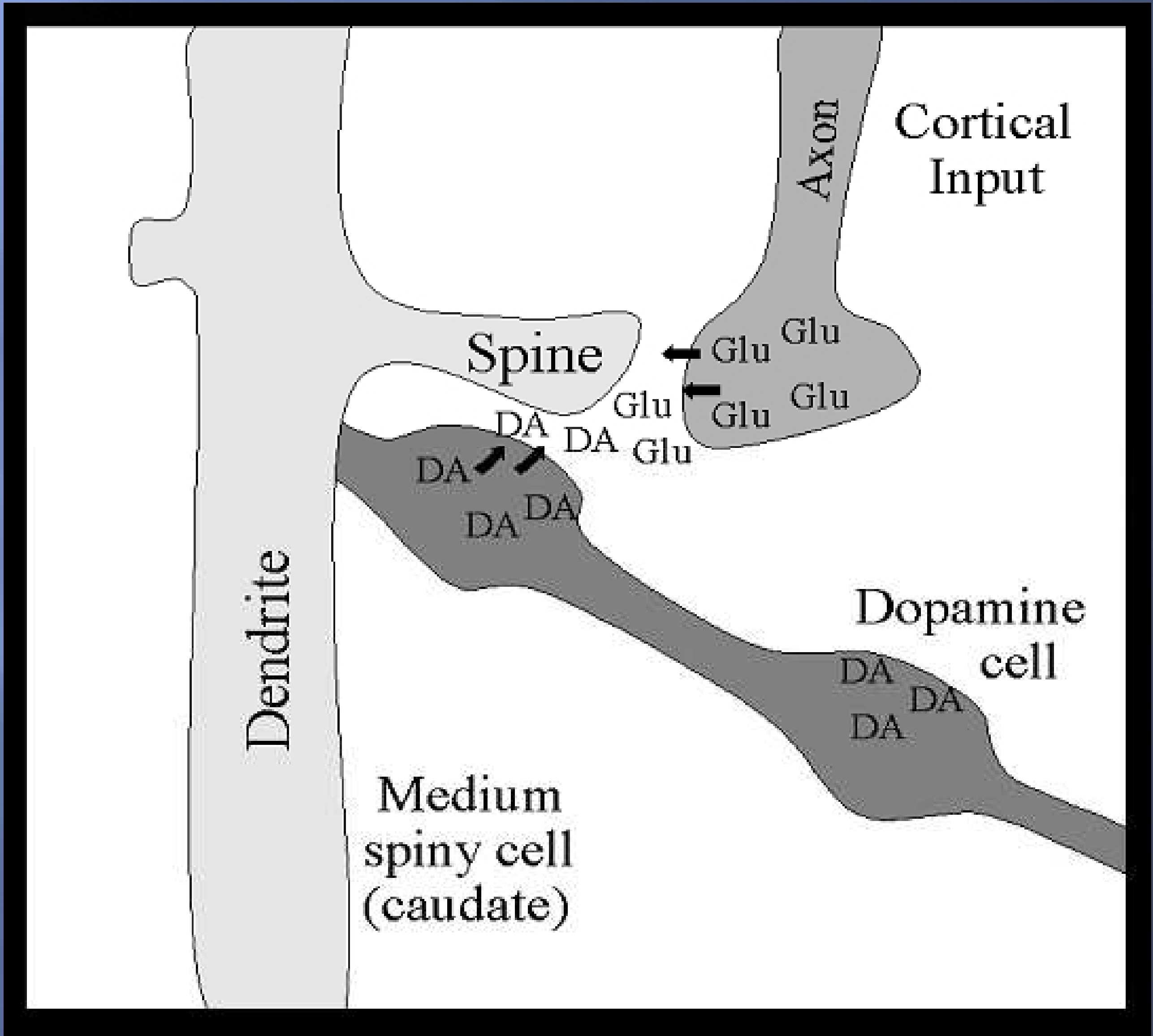


The COVIS Procedural-Learning System

The Striatal Pattern Classifier (Ashby & Waldron, 1999)



A Cortical-Striatal Synapse

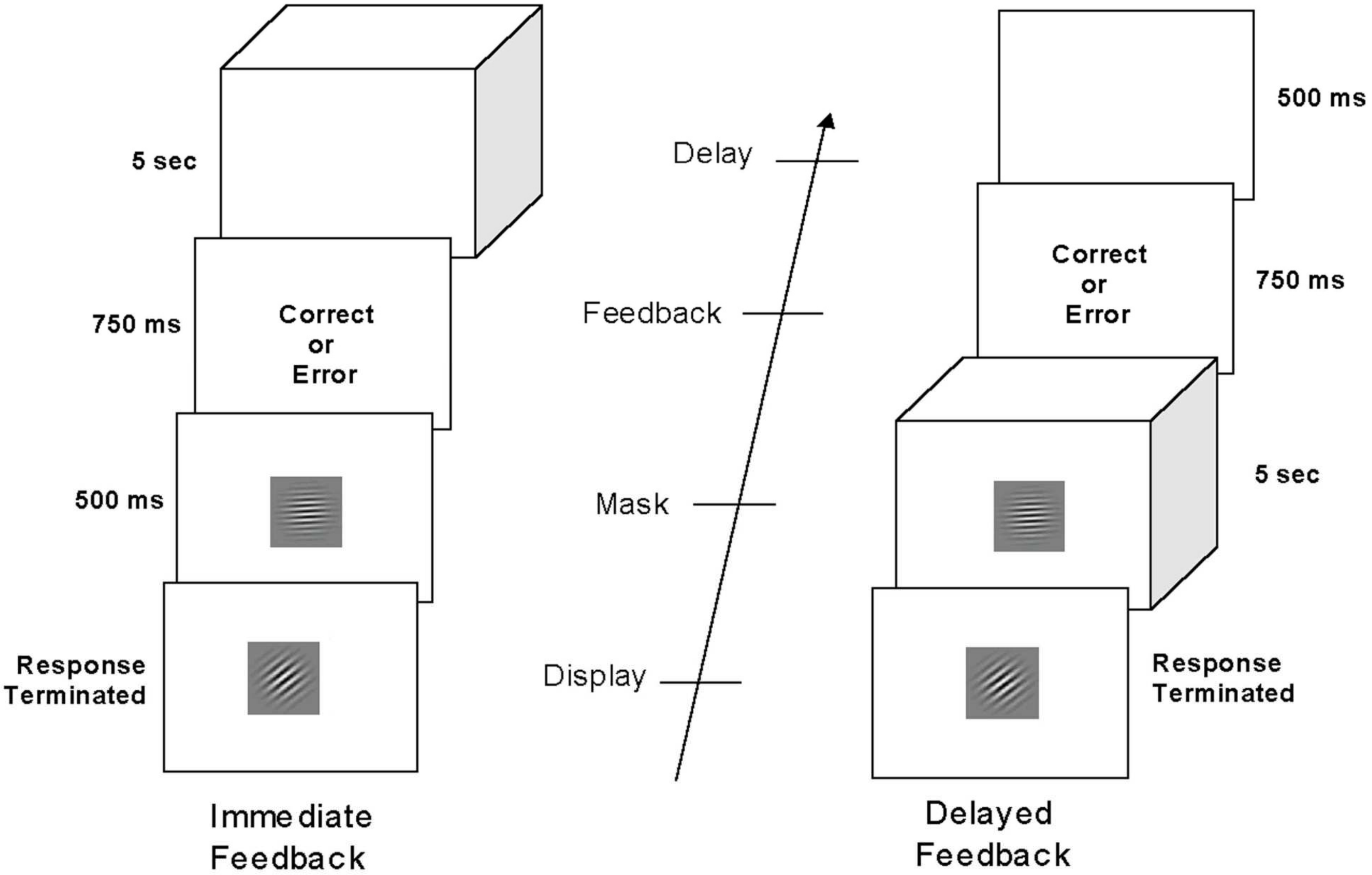


FEEDBACK PREDICTION

- Information-integration category learning should be sensitive to feedback delay

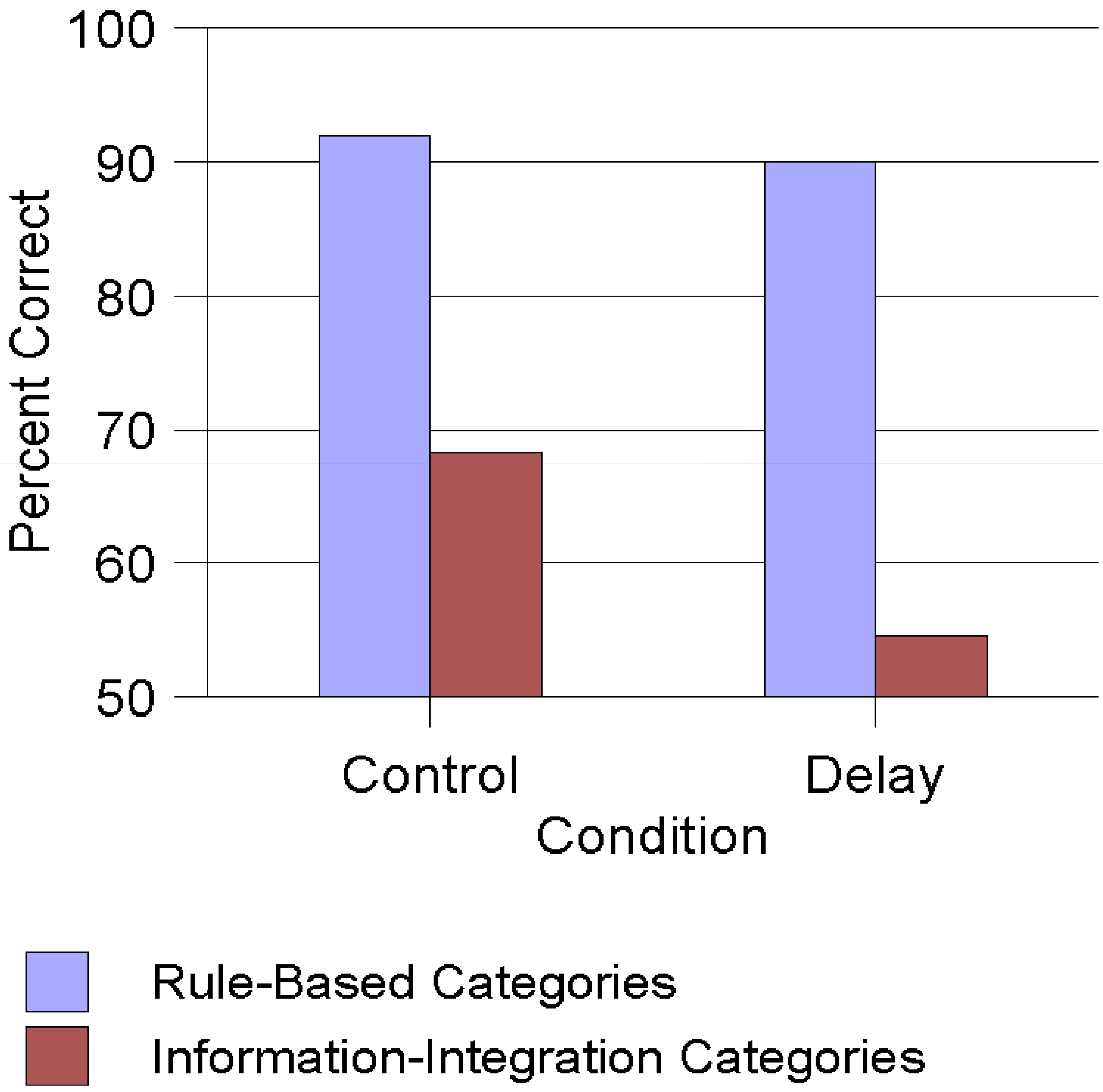
- Rule-based category learning should not be sensitive to feedback delay

Design of Feedback-Delay Experiment



Maddox, Ashby, & Bohil (2003, *JEP:LM&C*)

Effects of Feedback Delay



Maddox, Ashby, & Bohil (2003, *JEP:LM&C*)

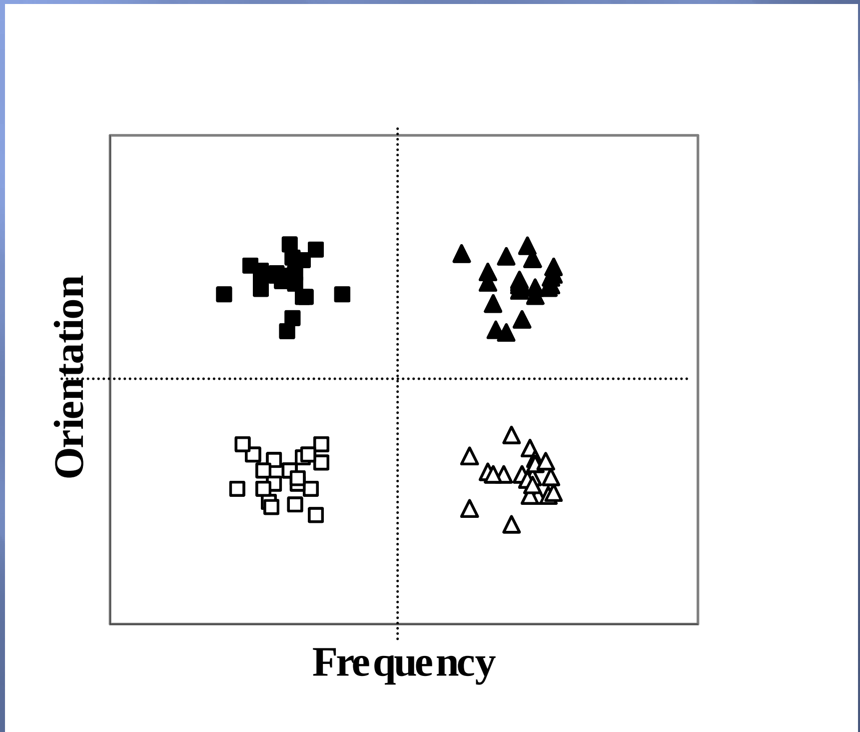
FOLLOW-UP EXPERIMENTS

- Results identical with 2.5 and 10 sec delays
- RB results replicated at 4 increased levels of difficulty
- Replication with a rule-based task that uses a conjunction rule?

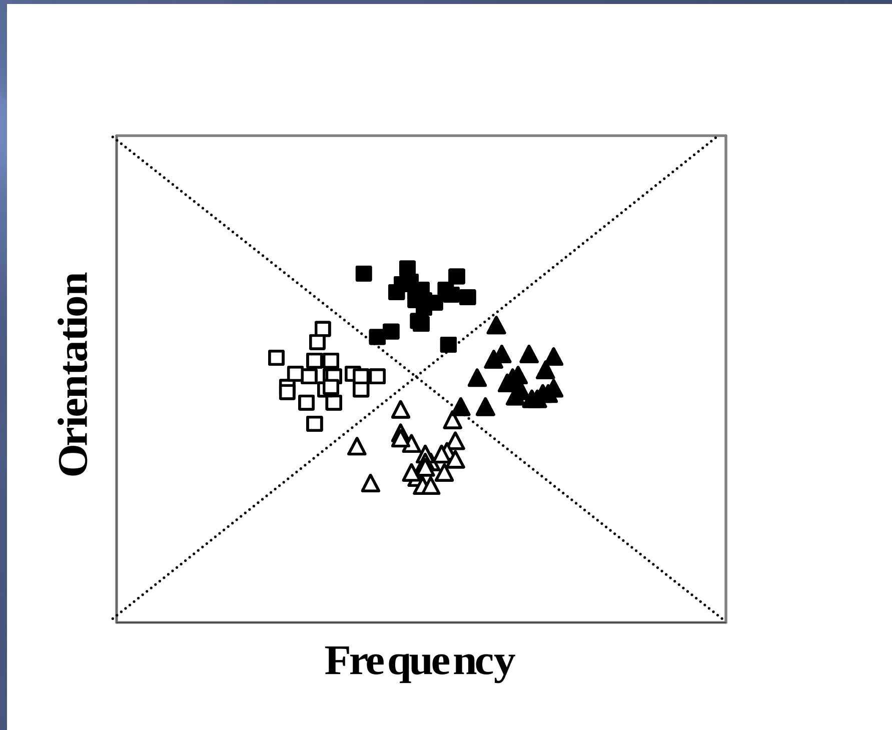
Category Structures

(Note: Rule-based discriminability higher)

Rule-Based

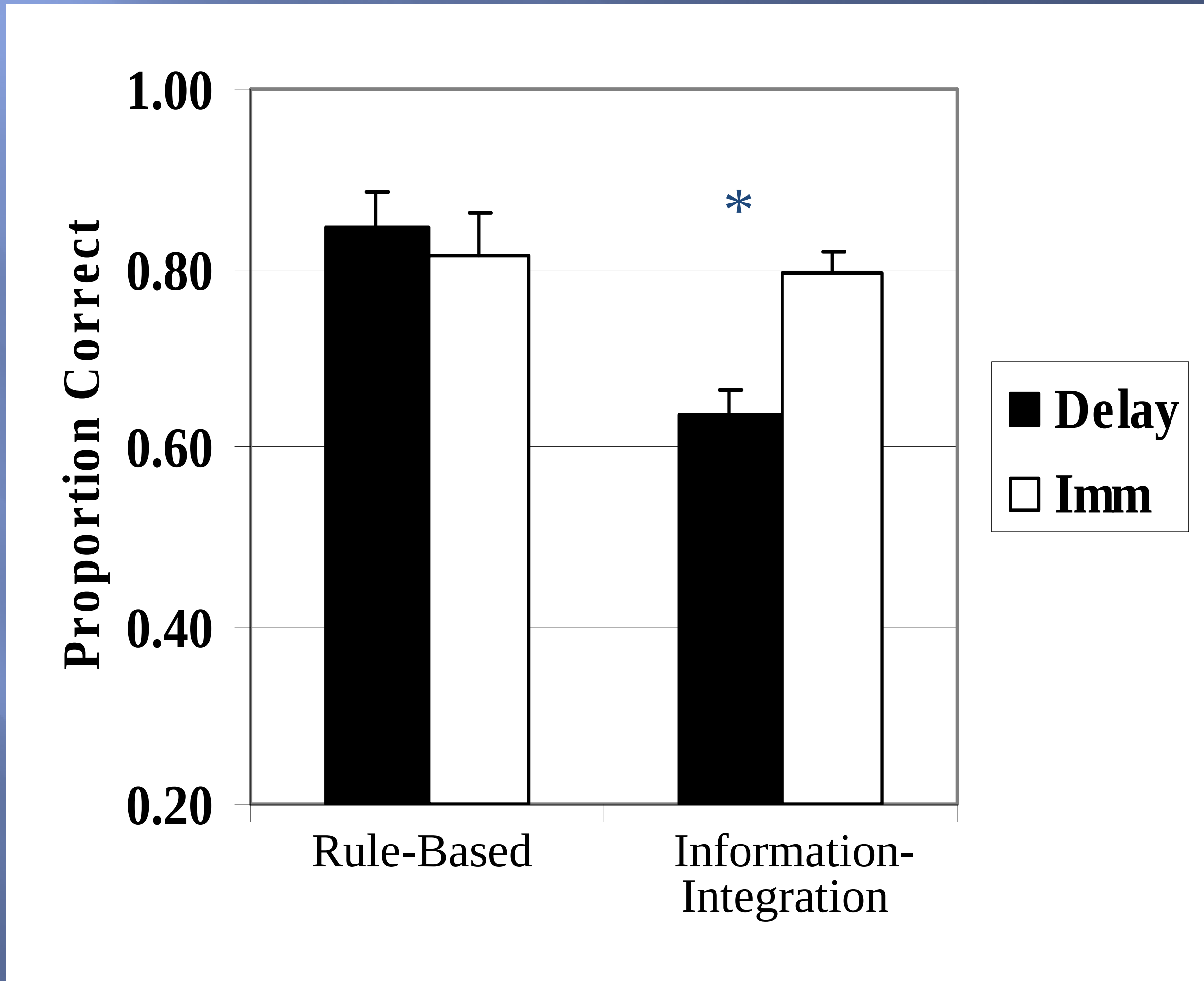


Information-Integration



Maddox & Ing (2005, JEP:LM&C)

Final Block Accuracy



Maddox & Ing (2005, JEP:LM&C)

CONCLUSIONS

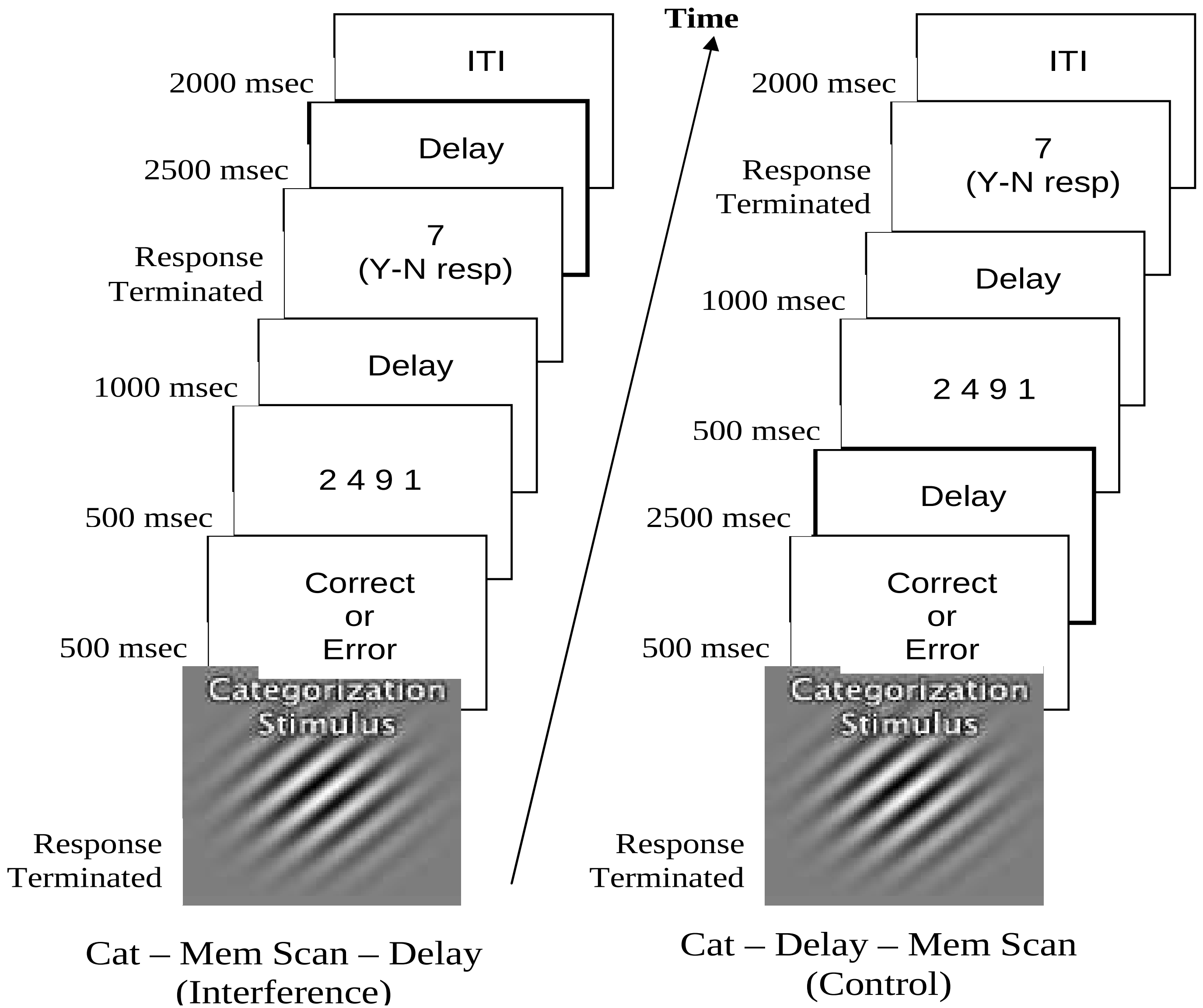
Feedback delay interferes with information-integration category learning, but not with rule-based category learning.

FEEDBACK PREDICTION

- Rule-based category learning requires active processing of feedback signal

- Feedback processing is automatic in information-integration category learning

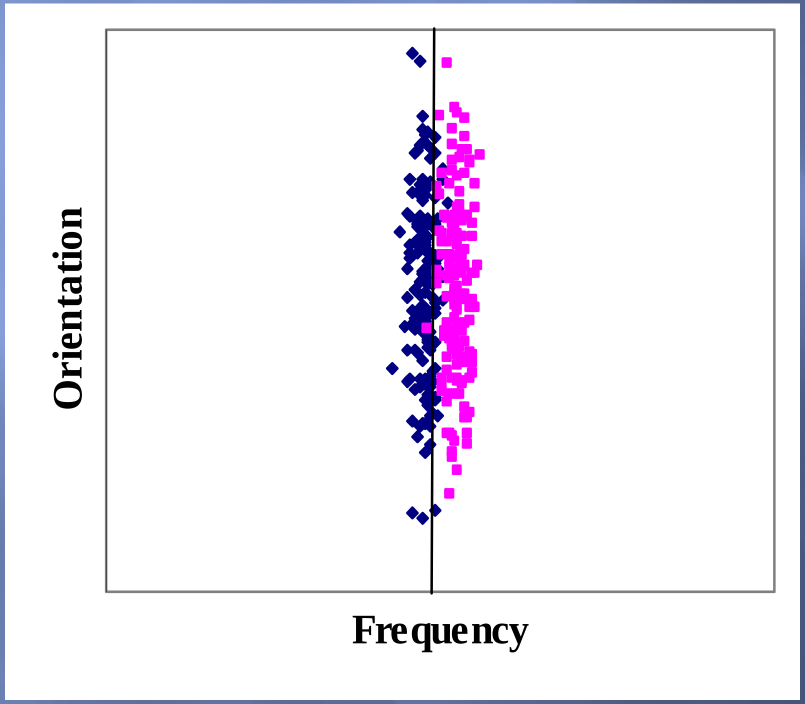
Feedback Interference Design



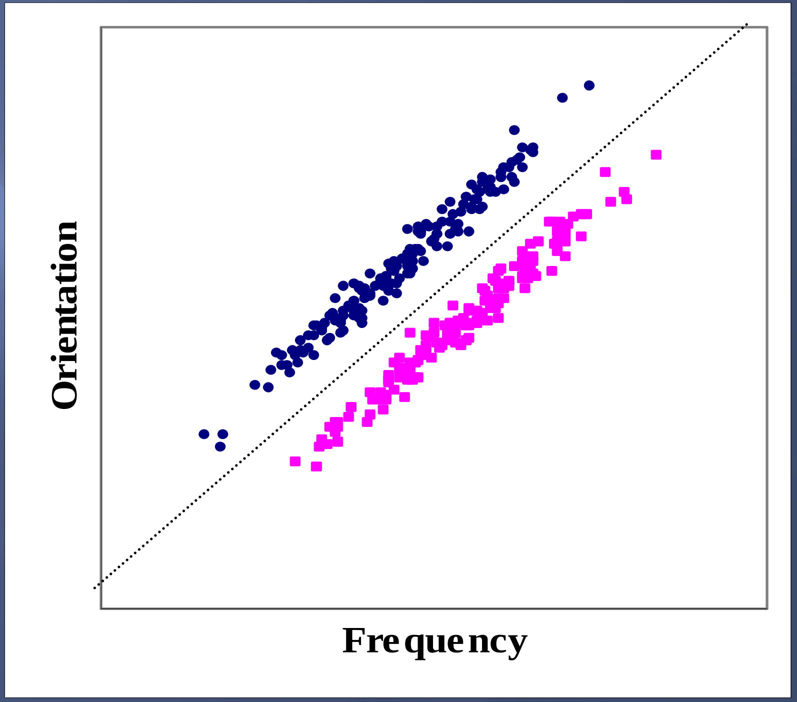
Maddox, Ashby, Ing, & Pickering (2004, *Memory & Cognition*)

Category Structures

Rule-Based

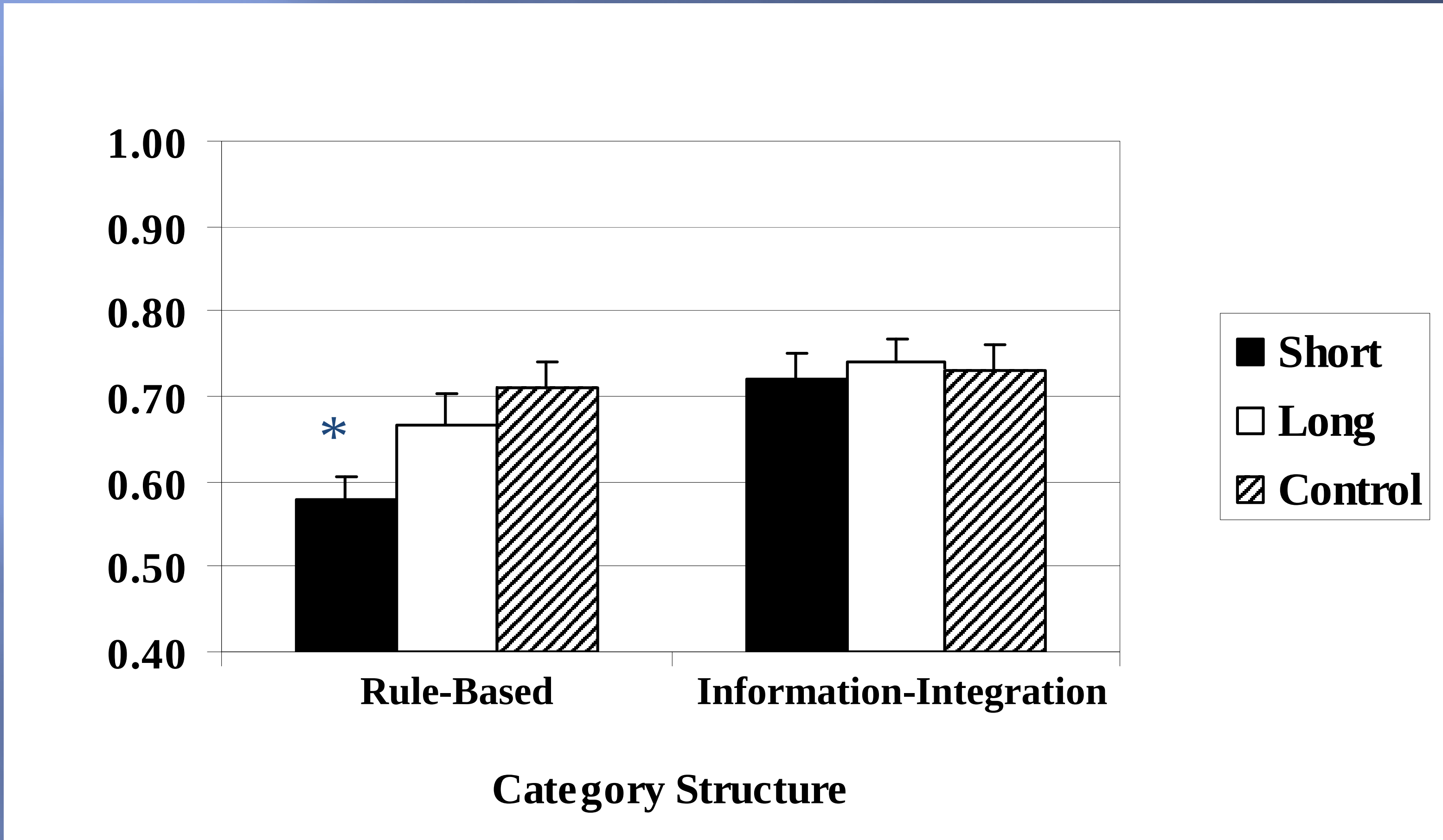


Information-Integration



Maddox, Ashby, Ing, & Pickering (2004, *Memory & Cognition*)

Final Block Proportion Correct



Maddox, Ashby, Ing, & Pickering (2004, *Memory & Cognition*)

EVIDENCE SUPPORTING COVIS

Single-cell recording studies

Asaad, Rainer, & Miller, 2000; Hoshi, Shima, & Tanji, 1998; Merchant, Zainos, Hernadez, Salinas, & Romo, 1997; Romo, Merchant, Ruiz, Crespo, & Zainos, 1995; White & Wise, 1999

Animal lesion experiments

Eacott & Gaffan, 1991; Gaffan & Eacott, 1995; Gaffan & Harrison, 1987; McDonald & White, 1993, 1994; Packard, Hirsch, & White, 1989; Packard & McGaugh, 1992; Roberts & Wallis, 2000

Neuropsychological patient studies

Ashby, Noble, Filoteo, Waldron, & Ell, 2003; Brown & Marsden, 1988; Cools et al., 1984; Downes et al., 1989; Filoteo, Maddox, & Davis, 2001a, 2001b; Filoteo, Maddox, Ing, Zizak, & Song, in press; Filoteo, Maddox, Salmon, & Song, 2005; Janowsky, Shimamura, Kritchevsky, & Squire, 1989; Knowlton, Mangels, & Squire, 1996; Leng & Parkin, 1988; Snowden et al., 2001

EVIDENCE SUPPORTING COVIS

Neuroimaging experiments

Konishi et al., 1999; Lombardi et al., 1999; Nomura et al., in press; Poldrack, et al., 2001; Rao et al., 1997; Rogers, Andrews, Grasby, Brooks, & Robbins, 2000; Seger & Cincotta, 2002; Volz et al., 1997

Traditional cognitive behavioral experiments

Ashby & Ell, 2002; Ashby, Ell, & Waldron, 2003; Ashby, Maddox, & Bohil, 2002; Ashby, Queller, & Berretty, 1999; Ashby, Waldron, Lee, & Berkman, 2001; Maddox, Ashby, & Bohil, 2003; Maddox, Ashby, Ing, & Pickering, 2004; Maddox, Bohil, & Ing, in press; Waldron & Ashby, 2001; Zeithamova & Maddox, in press

AUTOMATICITY IN II-TYPE TASKS

EARLY NOTIONS OF AUTOMATICITY

"As I write, my mind is not preoccupied with how my fingers form the letters; my attention is fixed simply on the thought the words express. But there was a time when the formation of the letters, as each one was written, would have occupied my whole attention."

Sir Charles Sherrington (1906)

EARLY NOTIONS OF AUTOMATICITY

"It has been widely held that although memory traces are at first formed in the cerebral cortex, they are finally reduced or transferred by long practice to subcortical levels" (p. 466)

Karl Lashley (1950) In search of the engram.

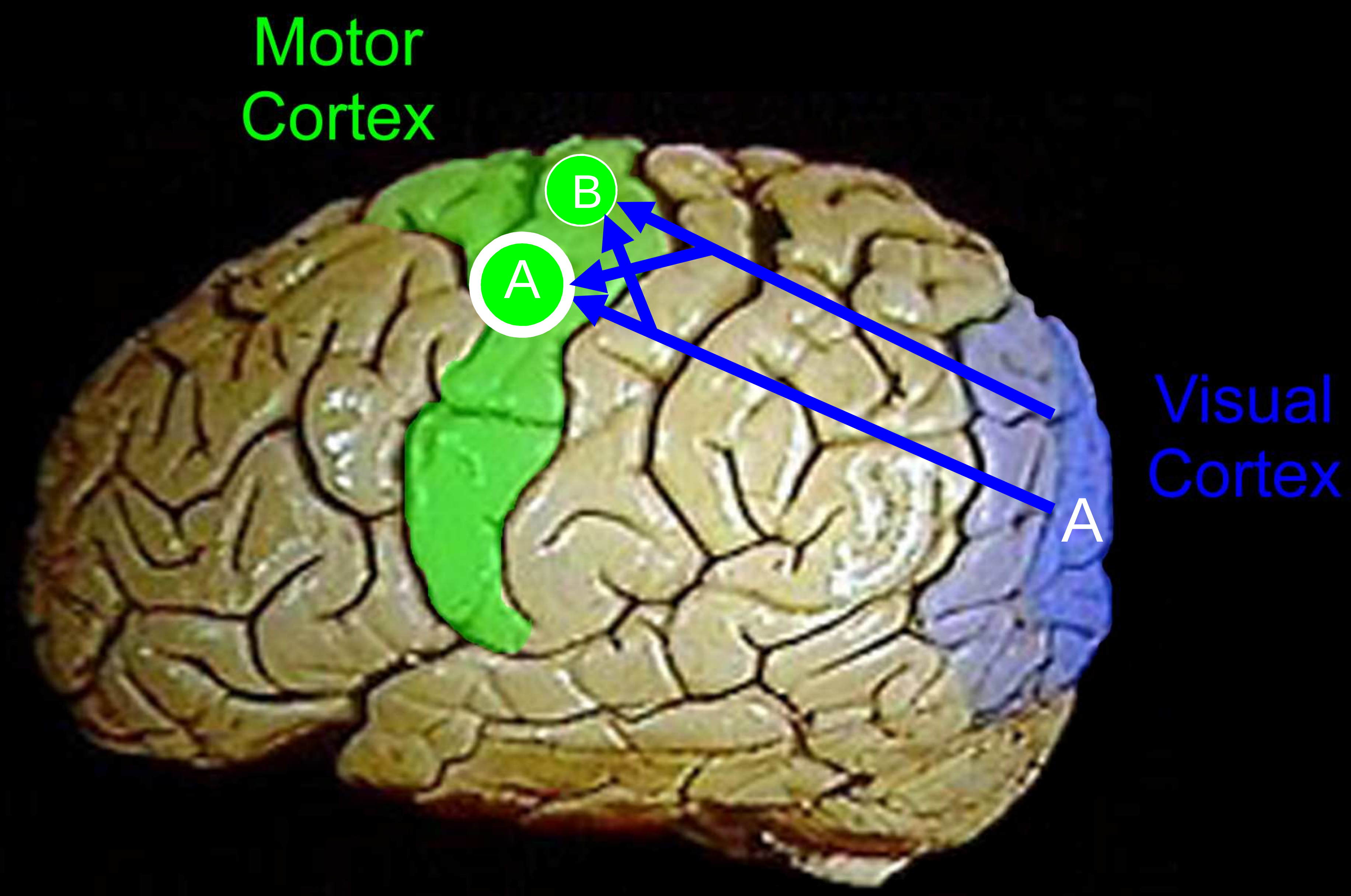
"Routine, automatic, or overlearned behavioral sequences, however complex, do not engage the PFC and may be entirely organized in subcortical structures" (p. 323)

Joaquin Fuster (2001). The prefrontal cortex – an update.

A DOUBLE DISSOCIATION?

	Category Learning	Categorization Expertise
Patients with Basal Ganglia Dysfunction (Parkinson's disease, Huntington's disease)	Impaired	Unimpaired
Patients with certain visual cortex lesions (category-specific agnosia)	Unimpaired if stimuli are perceived normally?	Impaired

BUILDING A MODEL OF AUTOMATICITY

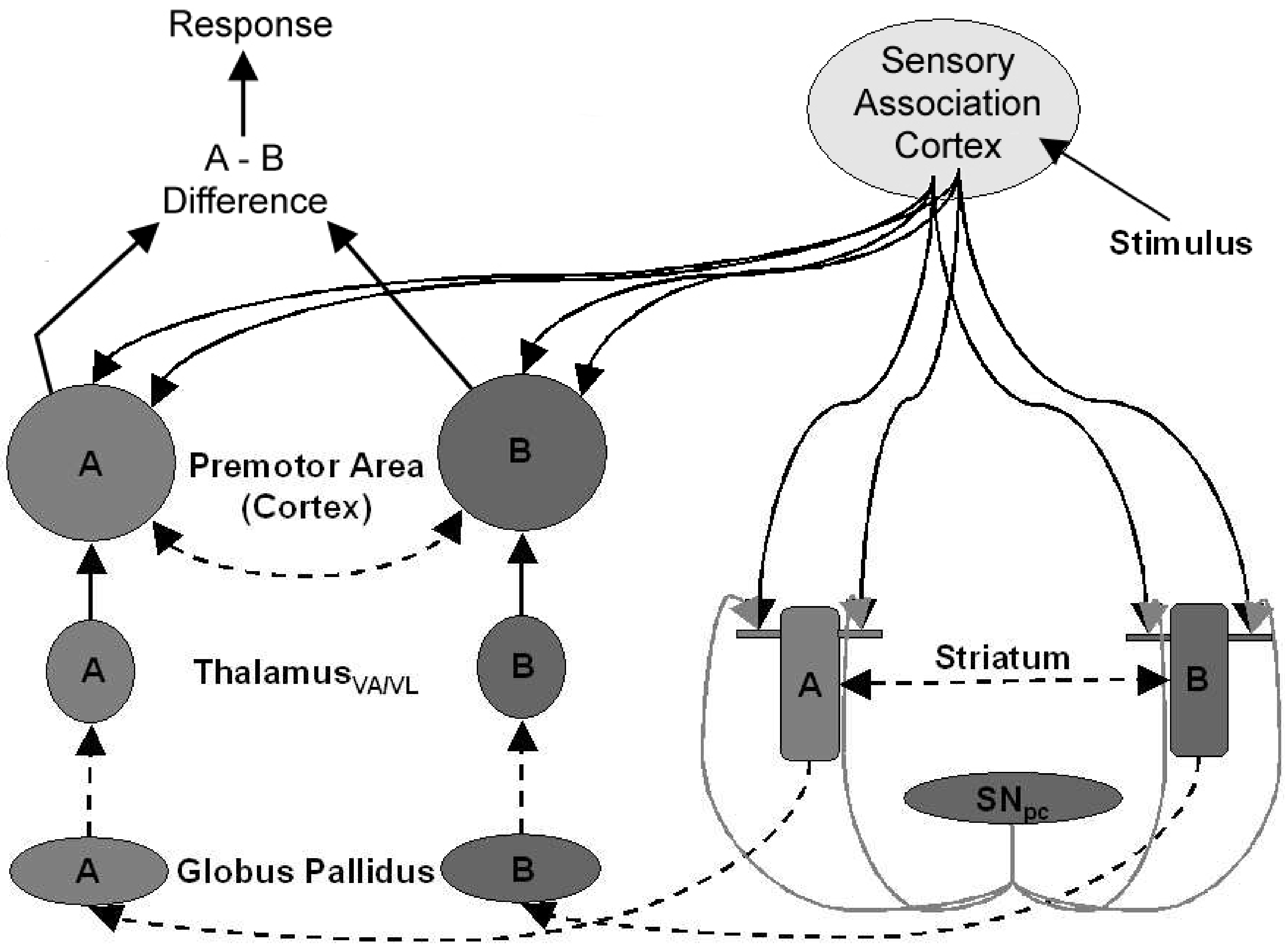


SPEED

Excitatory projection (glutamate)

Inhibitory projection (GABA)

Dopamine projection



Ashby, Ennis, & Spiering (2007, *Psych Review*)

Activation in Striatum (Medium Spiny Cells)

Activation in striatal unit *J* at time *t*, denoted $S_J(t)$ equals

$$\frac{dS_J(t)}{dt} = \left[\sum_K w_{K,J}(n) I_K(t)\right][1 - S_J(t)] - \beta_S S_M(t) - \gamma_S [S_J(t) - S_{base}] + \sigma_S \varepsilon(t) S_J(t)[1 - S_J(t)],$$

where $I_K(t)$ is the input from visual cortical unit *K* at time *t*, and $w_{K,iJ}(n)$ is the strength of the synapse between cortical unit *K* and spine *i* on medium spiny cell *J*, and ϵ(t) is white noise.

Modeling Activation in Other Units

Globus Pallidus

$$\frac{dG_J(t)}{dt} = -\alpha_G S_J(t) G_J(t) - \beta_G \left[G_J(t) - G_{base} \right]$$

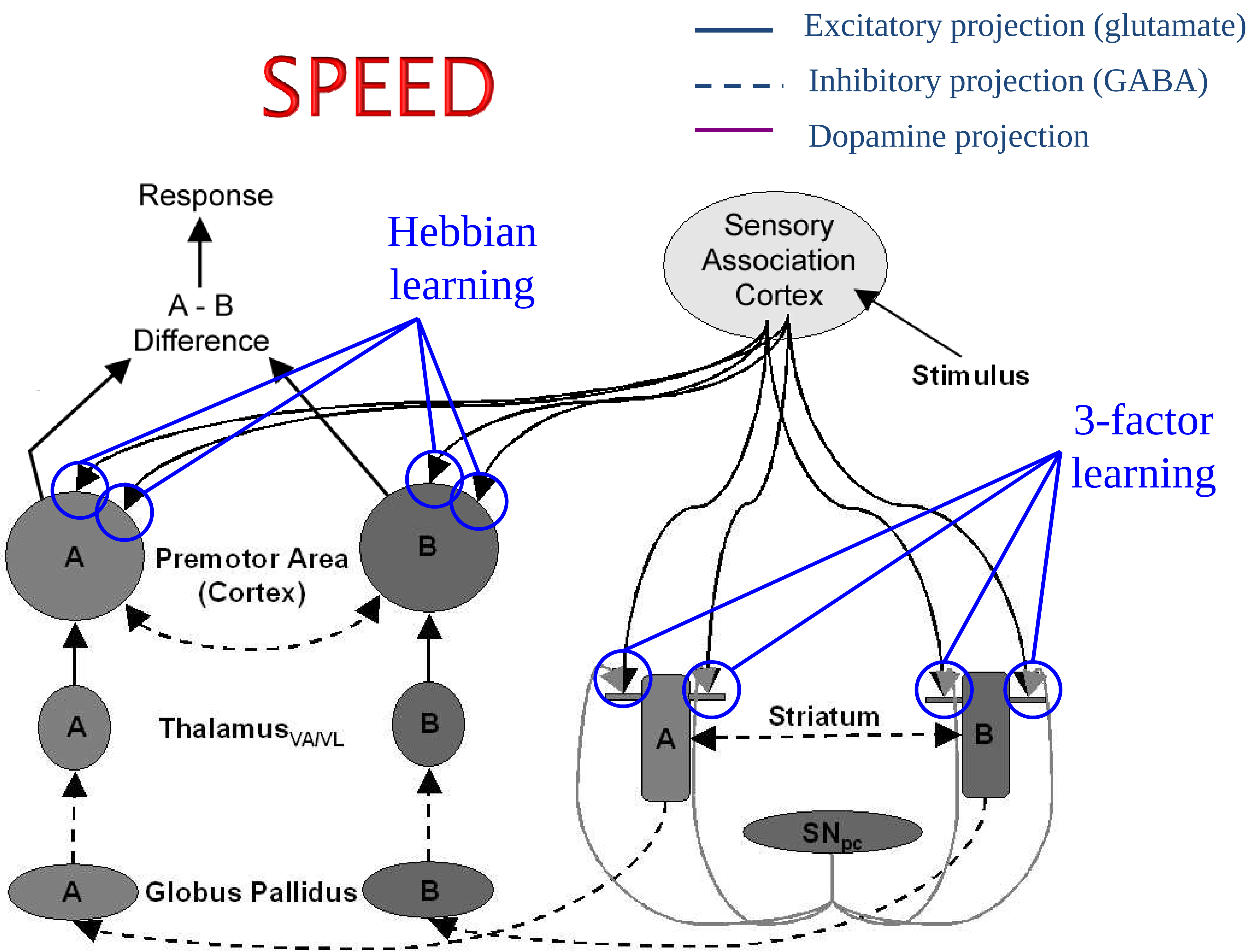
Thalamus:

$$\frac{dT_J(t)}{dt} = -\alpha_T G_J(t) T_J(t) - \beta_T T_J(t),$$

Premotor Area

$$\frac{dE_J(t)}{dt} = \left[\alpha_E T_J(t) + \sum_K v_{K,J}(n) I_K(t) \right] [1 - E_J(t)] - \beta_E E_K(t) - \gamma_E \left[E_J(t) - E_{base} \right] + \sigma_E \varepsilon(t) E_J(t) [1 - E_J(t)],$$

SPEED


Excitatory projection (glutamate)
Inhibitory projection (GABA)
Dopamine projection
Response
A - B
Difference
Hebbian
learning
Sensory
Association
Cortex
Stimulus
3-factor
learning
A
Premotor Area
(Cortex)
B
A
Thalamus$_{VA/VL}$
B
A
Globus Pallidus
B
Striatum
A
B
SN$_{pc}$

Cortical-Cortical Learning (Hebbian)

LTP

presynaptic activation

postsynaptic activation (above NMDA threshold)

$$v_{K,J}(n+1) = v_{K,J}(n) + \alpha_v I_k(t)\left[E_J(t) - \theta_{NMDA}\right]^+ \left[1 - v_{K,J}(n)\right]$$

$$- \beta_v I_k(t)\left[\theta_{NMDA} - E_J(t)\right]^+ v_{K,J}(t)$$

LTD

postsynaptic activation (below NMDA threshold)

Cortical-Striatal Learning (3-factor)

$$w_{K,iJ}(n+1) = w_{K,iJ}(n)$$
$$+\alpha_w S_J(t)\left[r_{K,iJ}(t)-\theta_{NMDA}\right]^+\left[D(n)-D_{base}\right]^+\left[1-w_{K,iJ}(n)\right]$$
$$-\beta_w S_J(t)\left[r_{K,iJ}(t)-\theta_{NMDA}\right]^+\left[D_{base}-D(n)\right]^+ w_{K,iJ}(n)$$
$$-\gamma_w\left[\theta_{NMDA}-r_{K,iJ}(t)\right]^+ w_{K,iJ}(n),$$

LTP

activation above NMDA threshold

dopamine above baseline (Correct Response)

LTD

activation below NMDA threshold

activation above NMDA threshold

dopamine below baseline (error)

Dopamine Release

Increases with:

Obtained Reward – Predicted Reward

where obtained reward on trial $n + 1$ equals

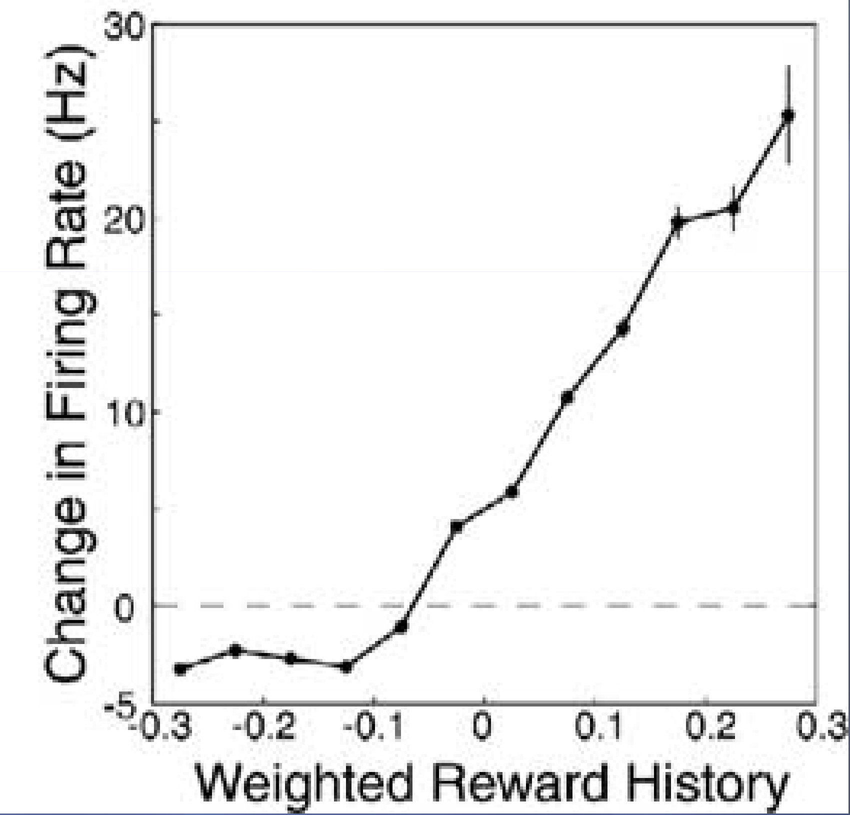
$$R_{n+1} = \begin{cases} 1 & \text{if correct feedback is received} \\ 0 & \text{if no feedback is received} \\ -1 & \text{if error feedback is received} \end{cases}$$

and

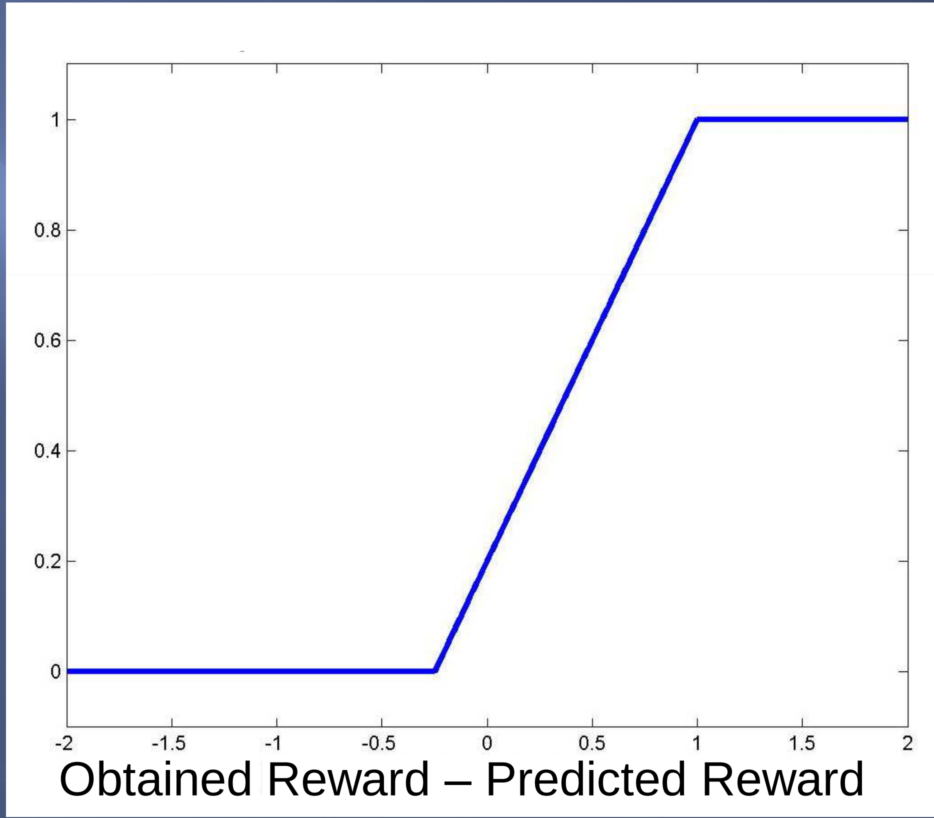
$$\text{Predicted Reward} = C \sum_i e^{\theta i} R_i$$

Dopamine Release

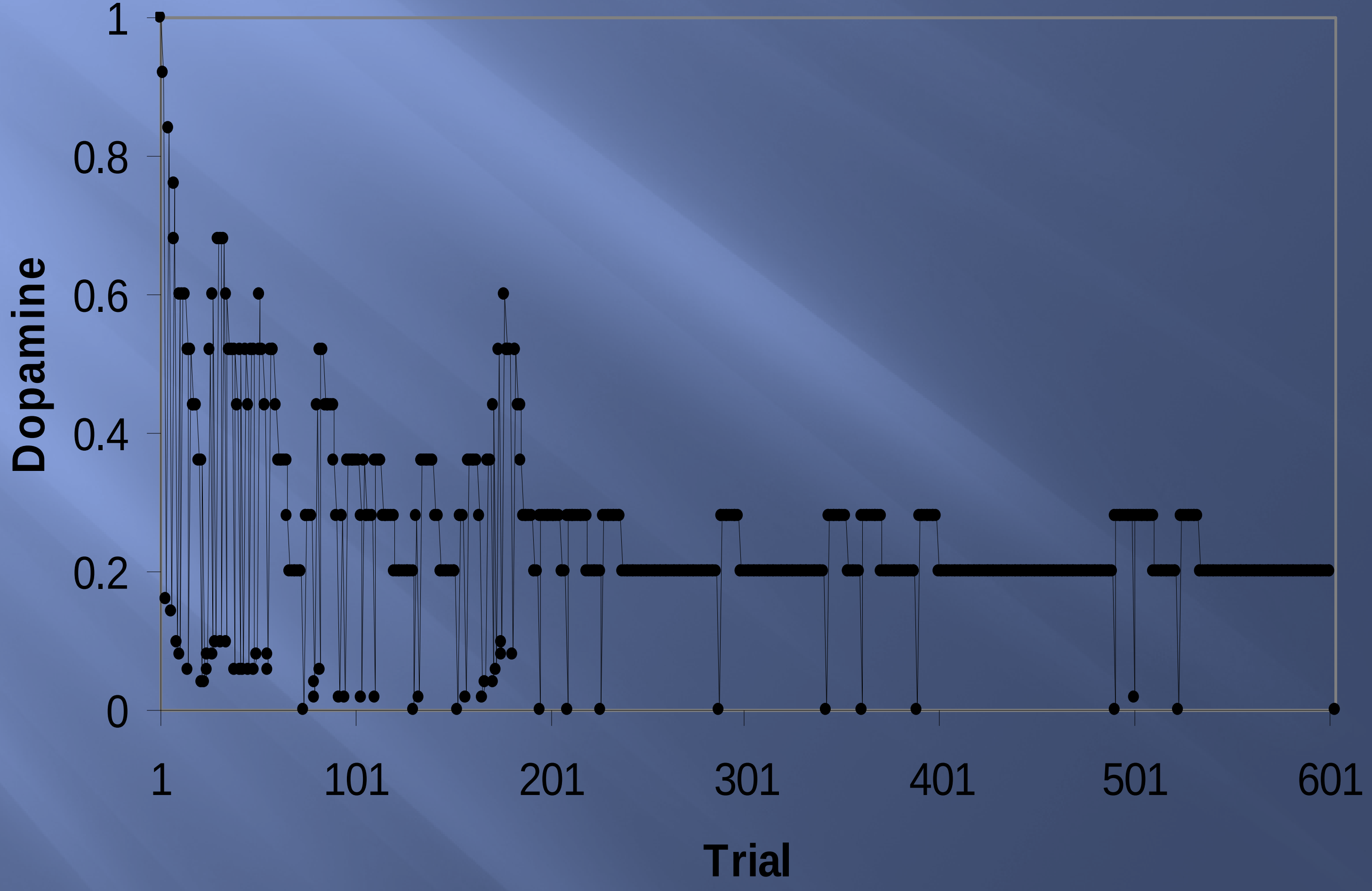
Bayer & Glimcher (2005, *Neuron*)



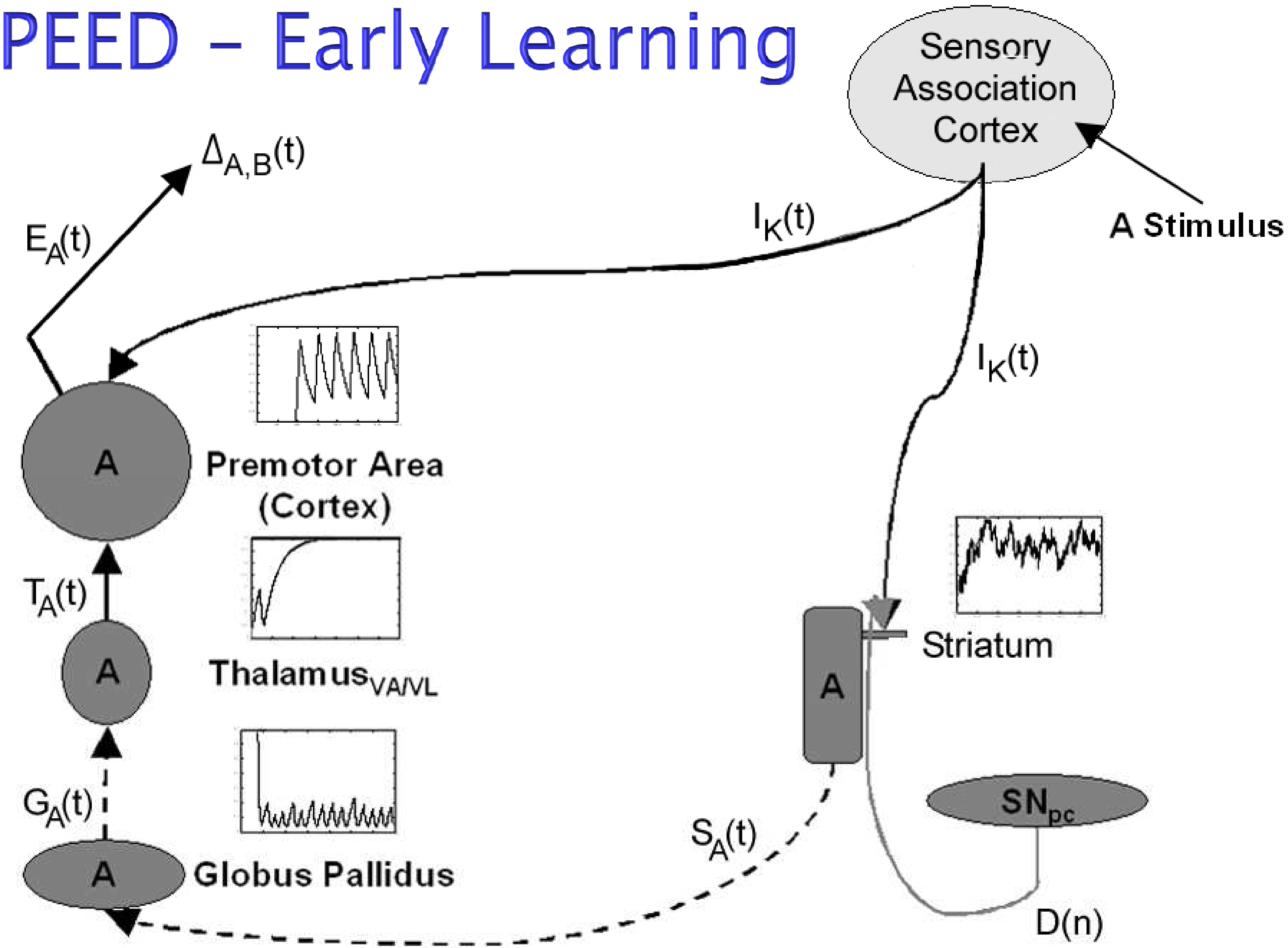
Dopamine Release in SPEED



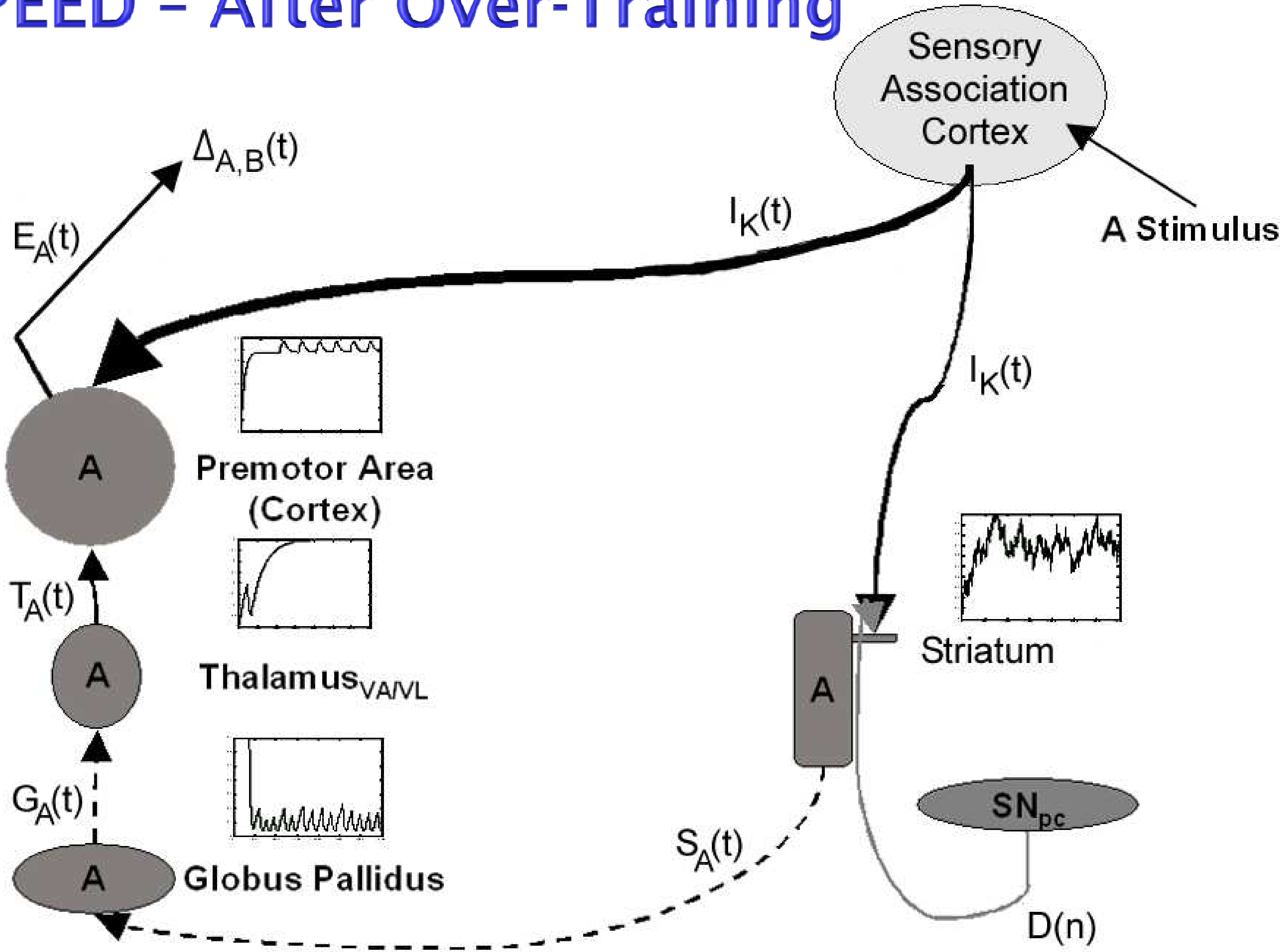
Dopamine Release



SPEED – Early Learning



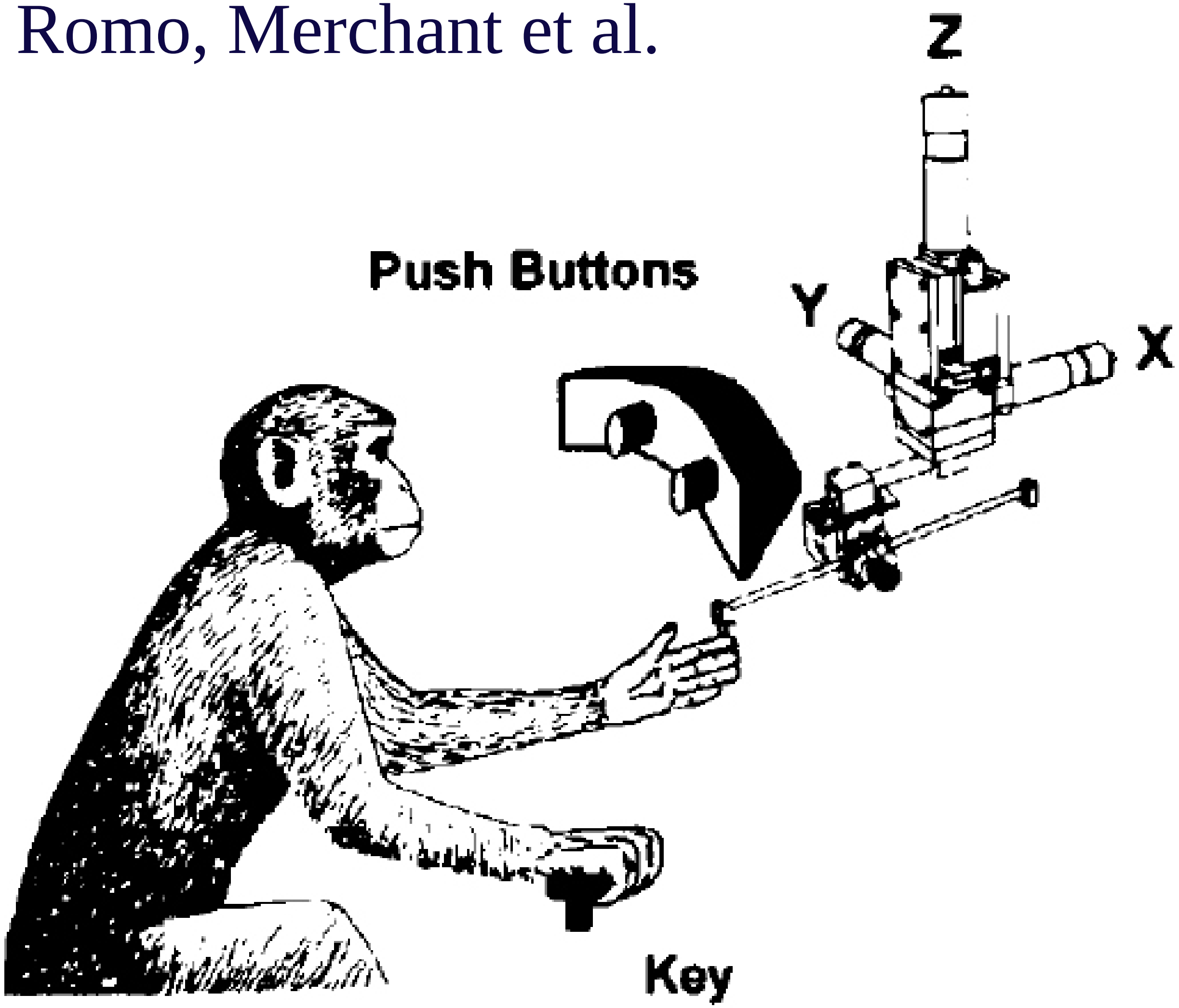
SPEED – After Over-Training



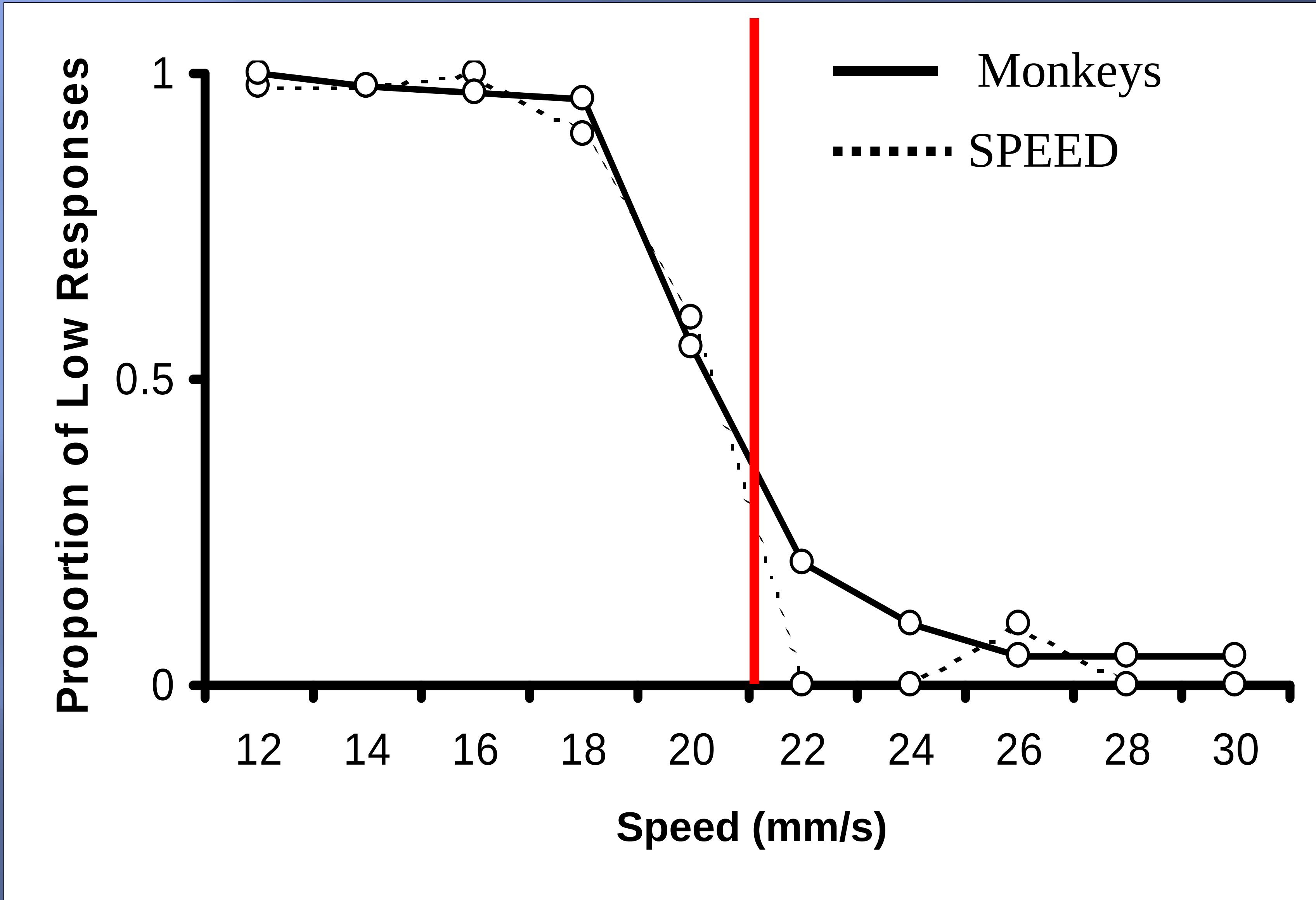
Experimental Tests

Tactile Category Learning

Romo, Merchant et al.



Model Fits

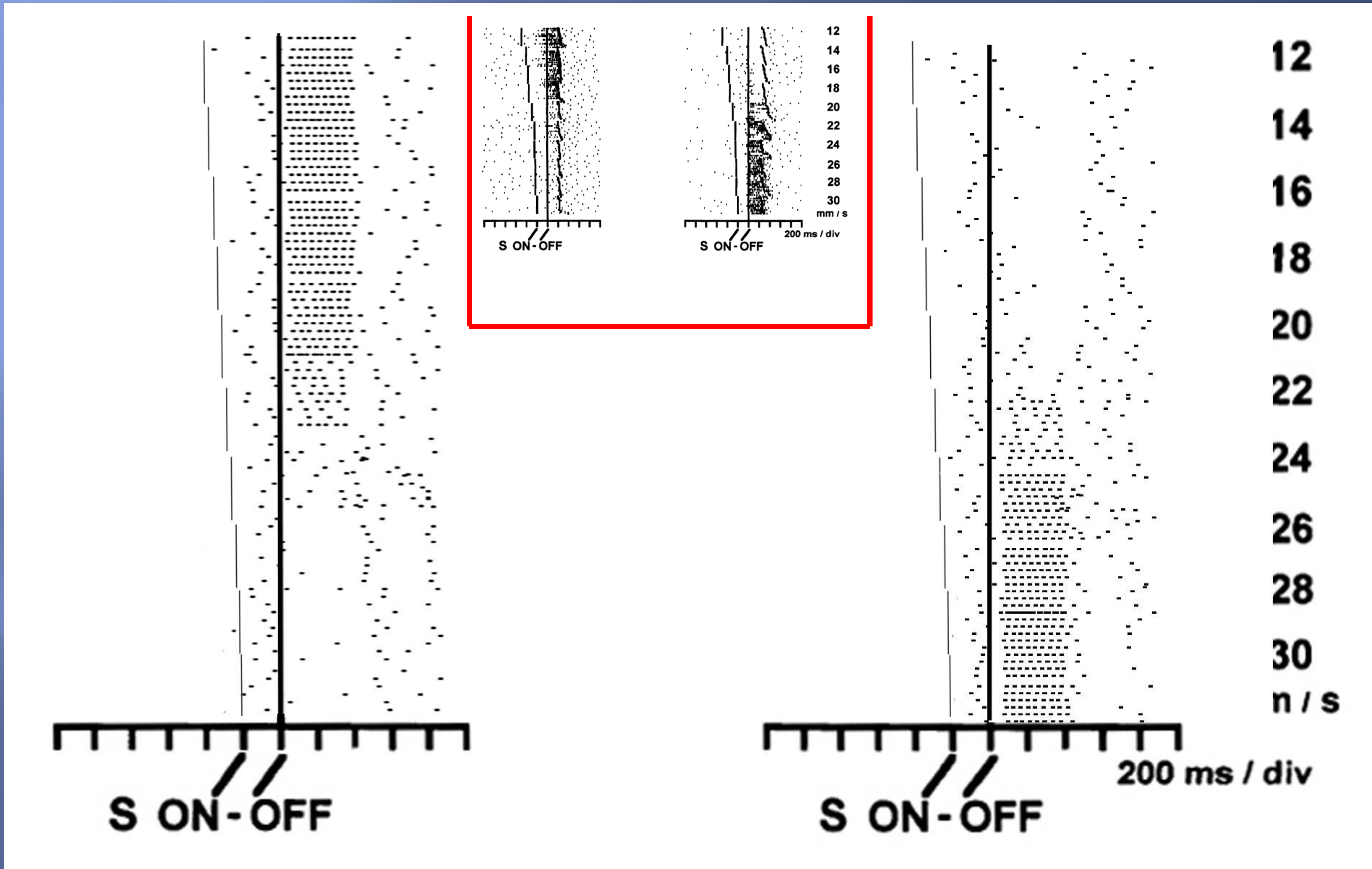


Proportion of Low Responses
1
0.5
0
12
14
16
18
20
22
24
26
28
30
Speed (mm/s)
Monkeys
SPEED

SPEED's Single Cell Responses –Putamen

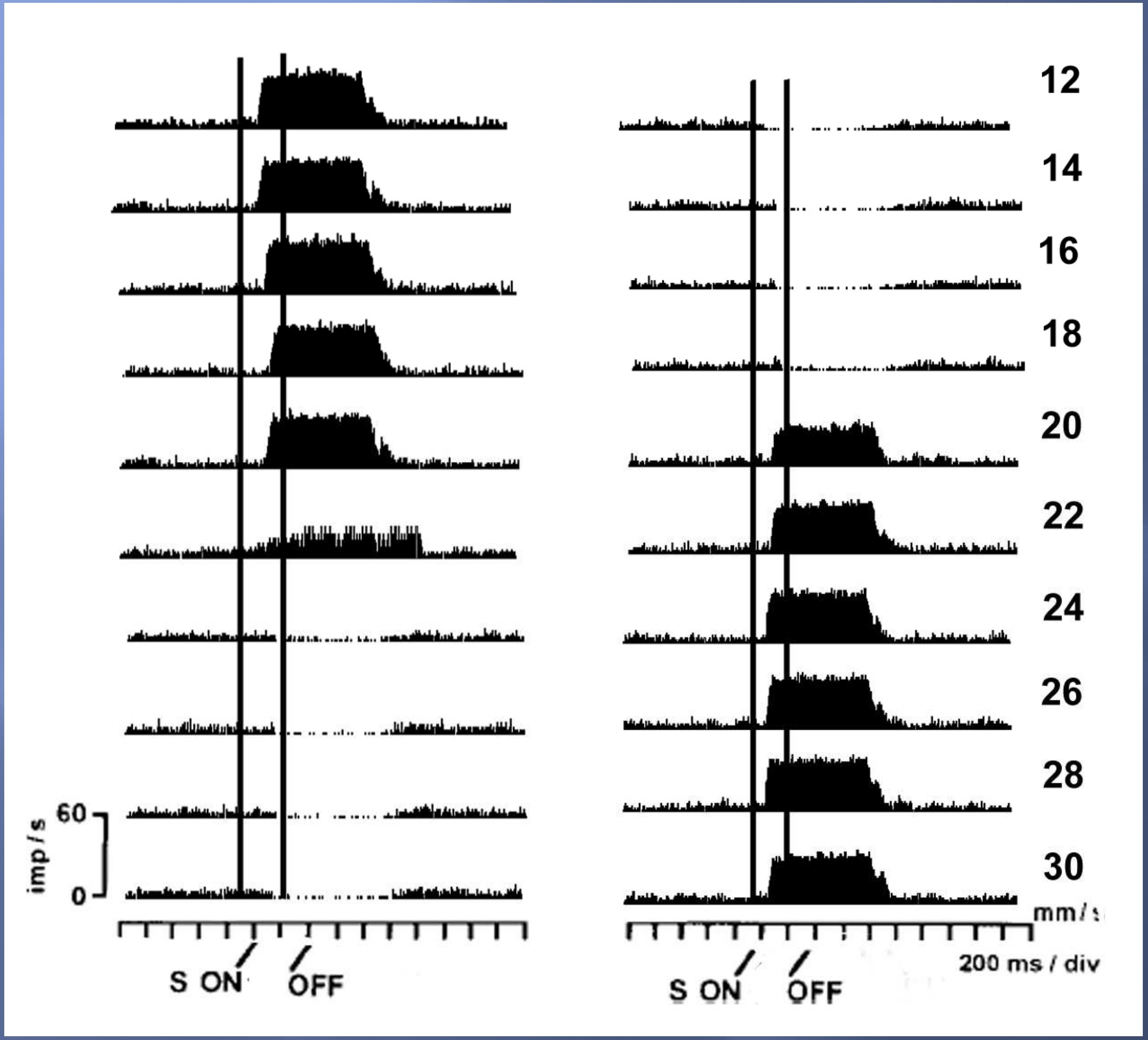
Low Speed Cell

High Speed Cell

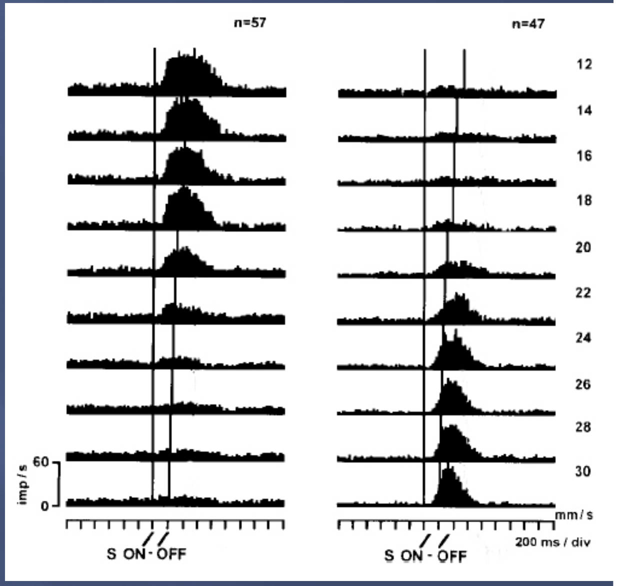


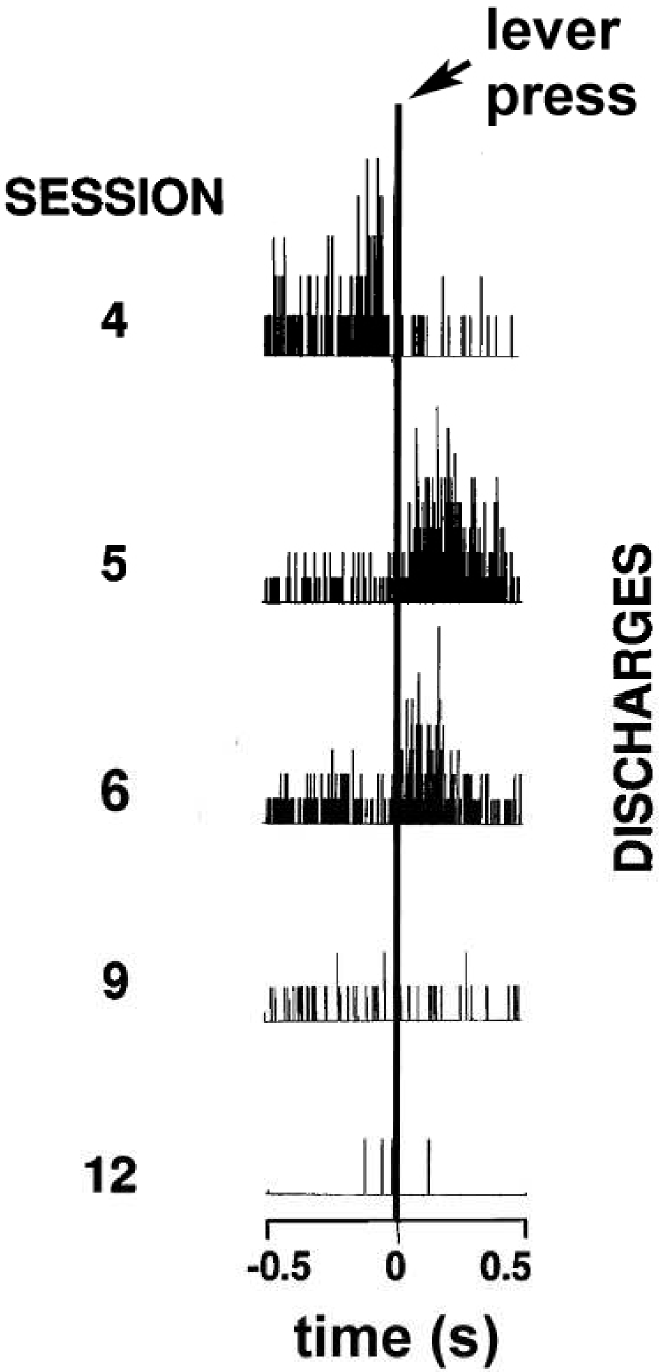
SPEED's Responses – Premotor Cortex

Low Speed Cells   High Speed Cells



Romo et al., 1997





Carelli, Wolske, & West

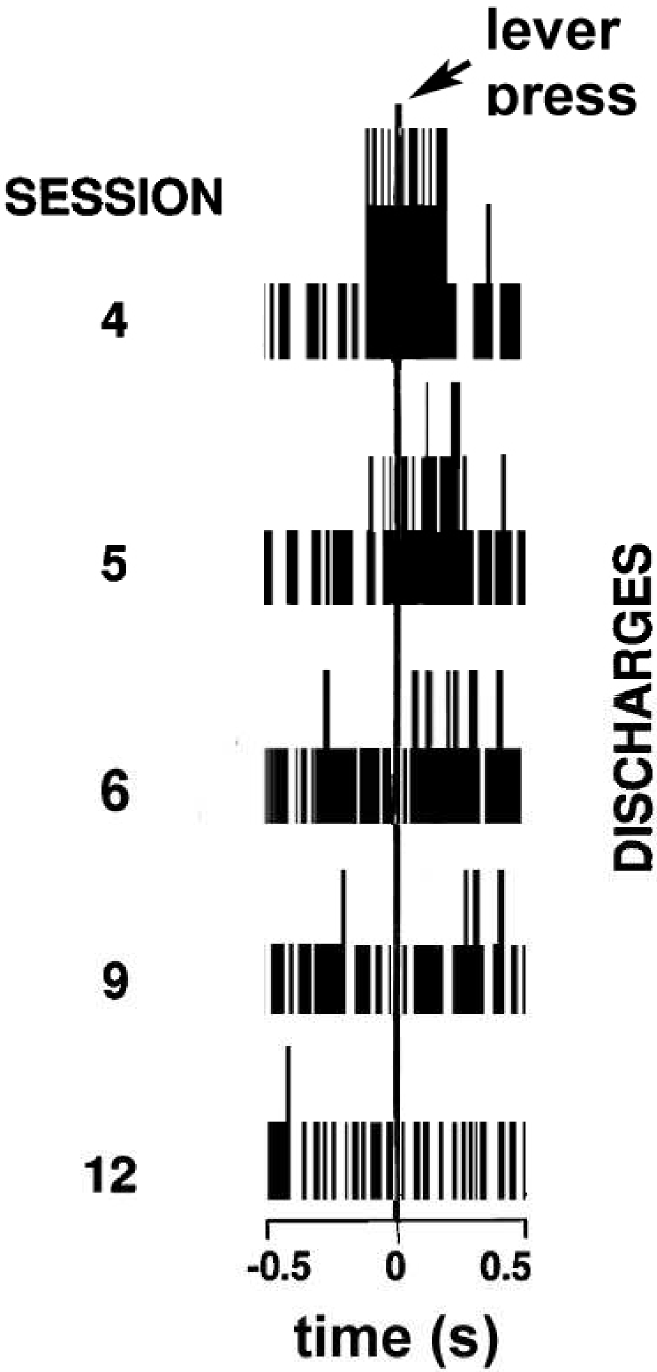
(1997, J. of Neuroscience)

Lever press to tone

70 trials/day

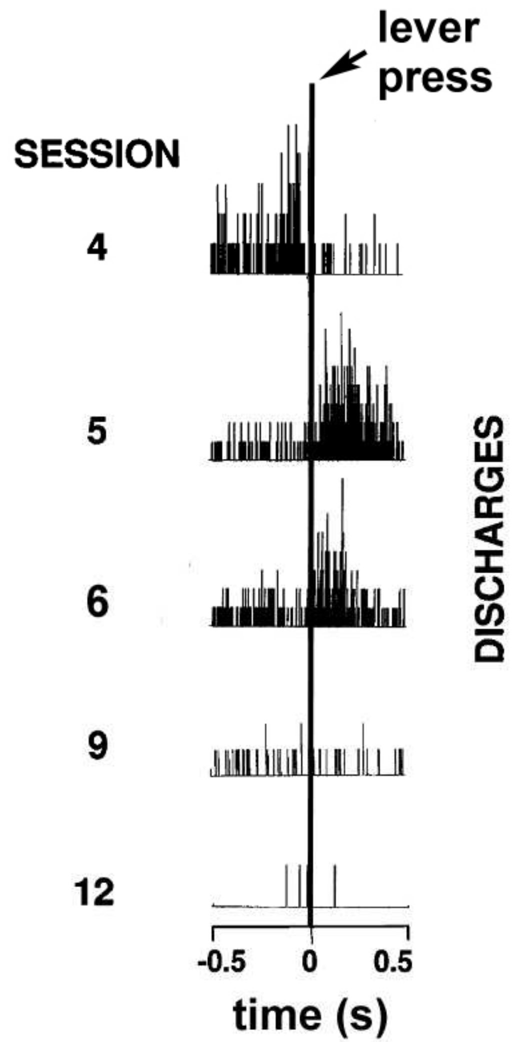
18 days

Striatal Response



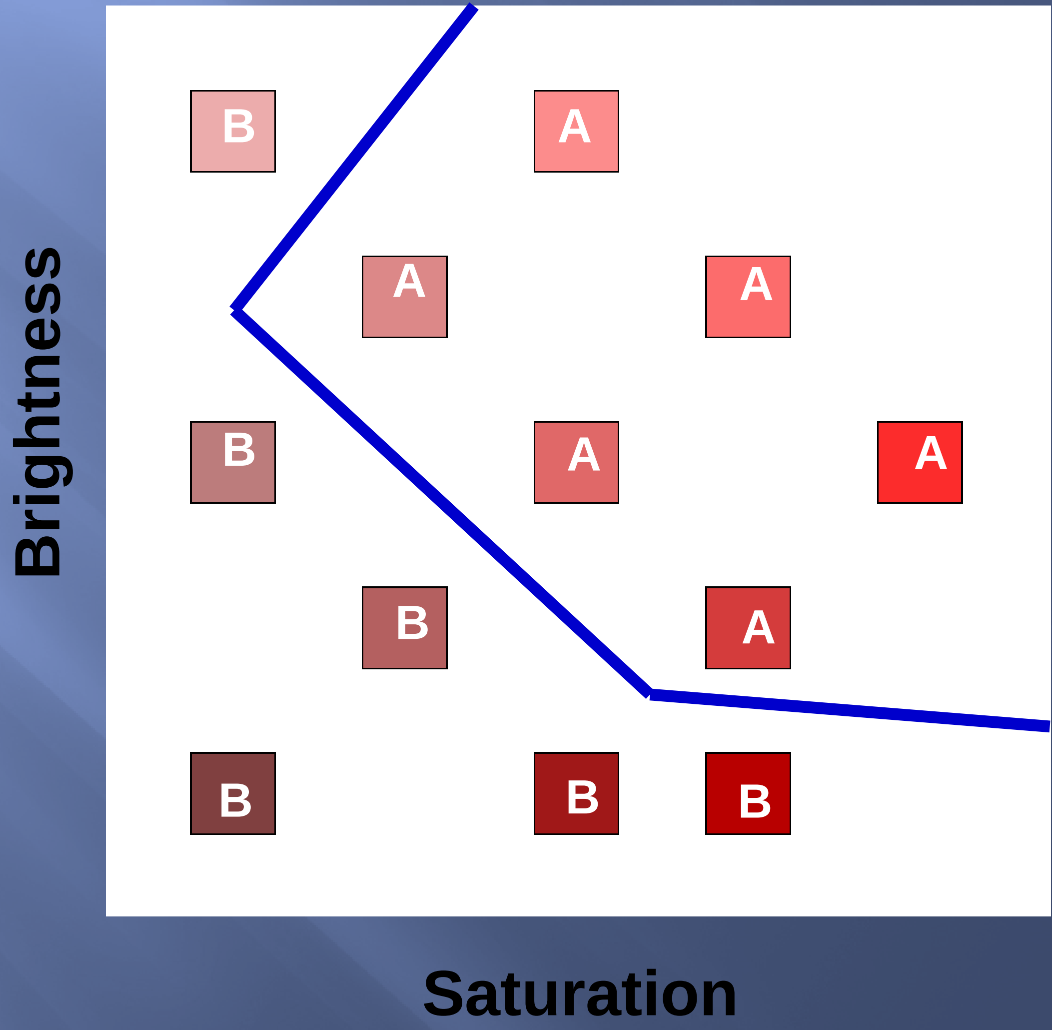
SPEED's Striatal Responses

Carelli et al. (1997, Journal of Neuroscience)

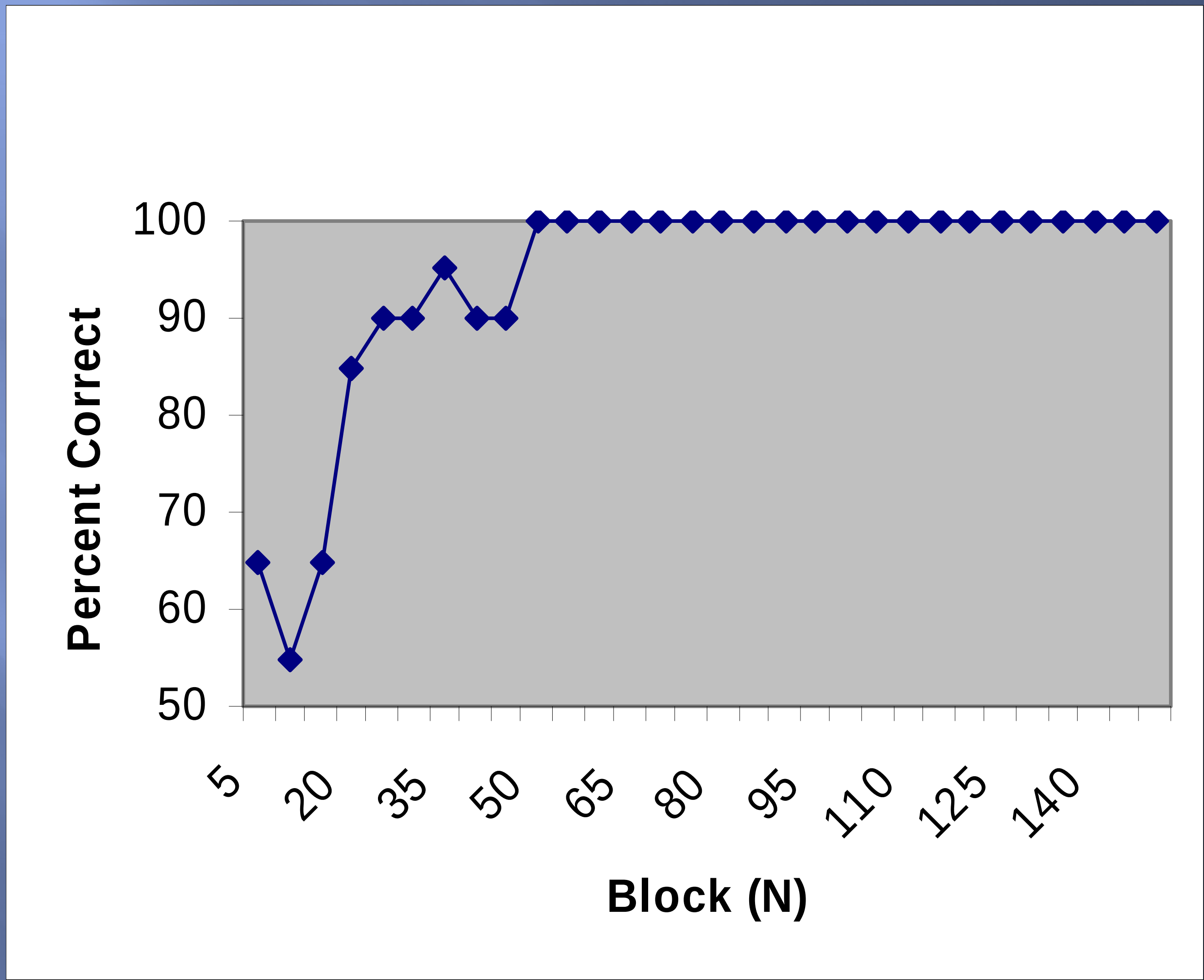


Nosofsky & Palmeri (1997, Psych Review)

Munsell Color Patches – 3 Subjects – 1800 Trials

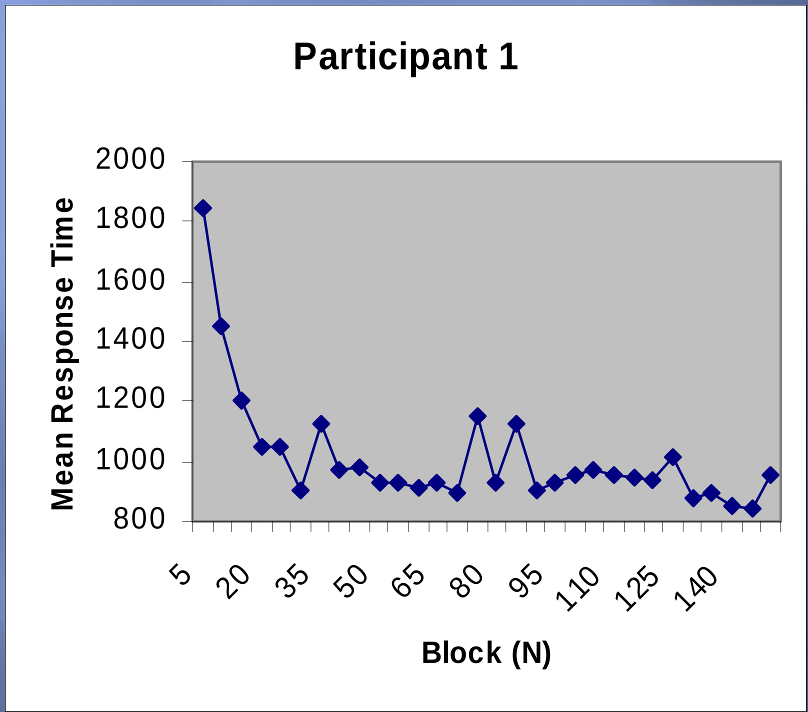


SPEED Accuracy

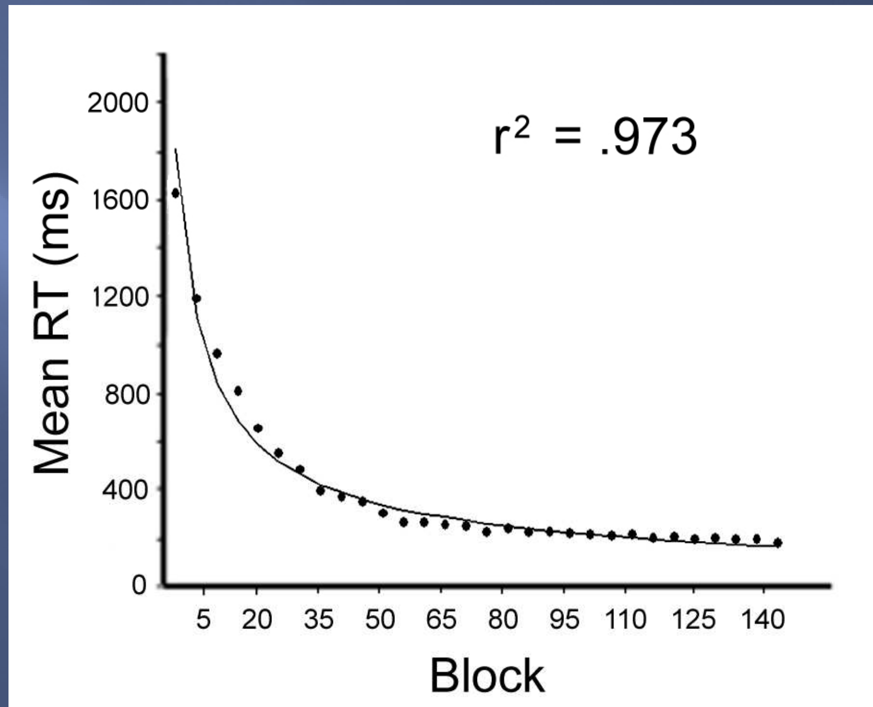


Mean Response Time

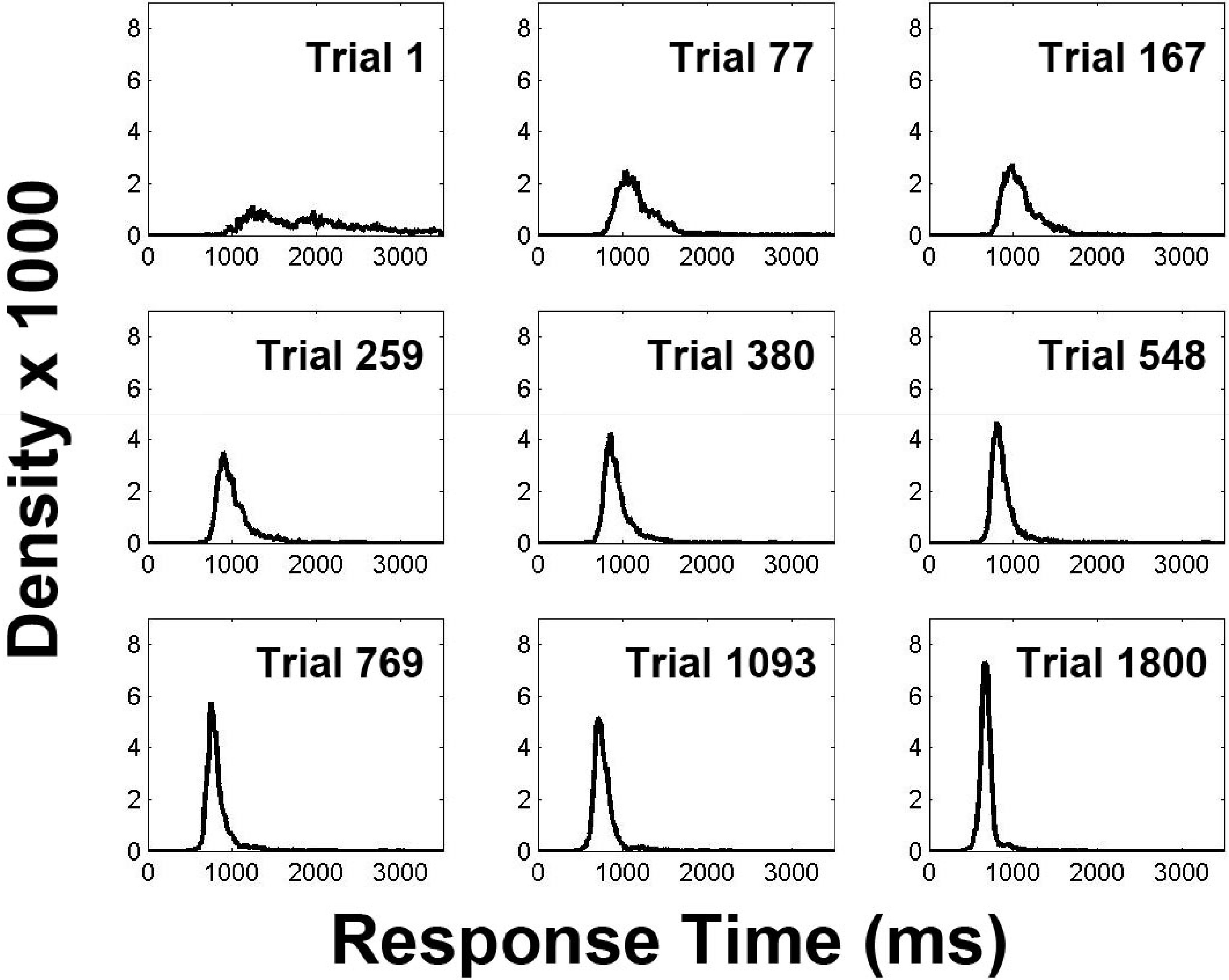
Nosofsky & Palmeri (1997)



SPEED



SPEED RT Density Functions



Future Directions

- fMRI
- Model automaticity development in:
  - -- neuropsychological populations
  - -- subjects under influence of drugs
- Automaticity in rule-based tasks

Conclusions

- Two category learning systems

- Explicit, logical reasoning system

  -- Uses working memory & executive attention

  -- Frontal cortex

- Procedural learning system

  -- Striatum

- Learning systems train long-term cortical representations

ACKNOWLEDGMENTS

Collaborators:

Learning

Leola Alfonso-Reese, Michael Beran, Robert Cook, Shawn Ell, Vince Filoteo, Todd Maddox, Alan Pickering, David Smith, And Turken, Elliott Waldron, many others

Automaticity

John Ennis, Brian Spiering

Funding:

Public Health Services Grant MH3760-2