A Neurobiological Theory of Automaticity in Perceptual Categorization

Brian J. Spiering
Laboratory for Computational Cognitive Neuroscience
University of California, Santa Barbara, CA

Bear or Dog?

Early Notions of Expertise

"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 Expertise

"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.

Outline

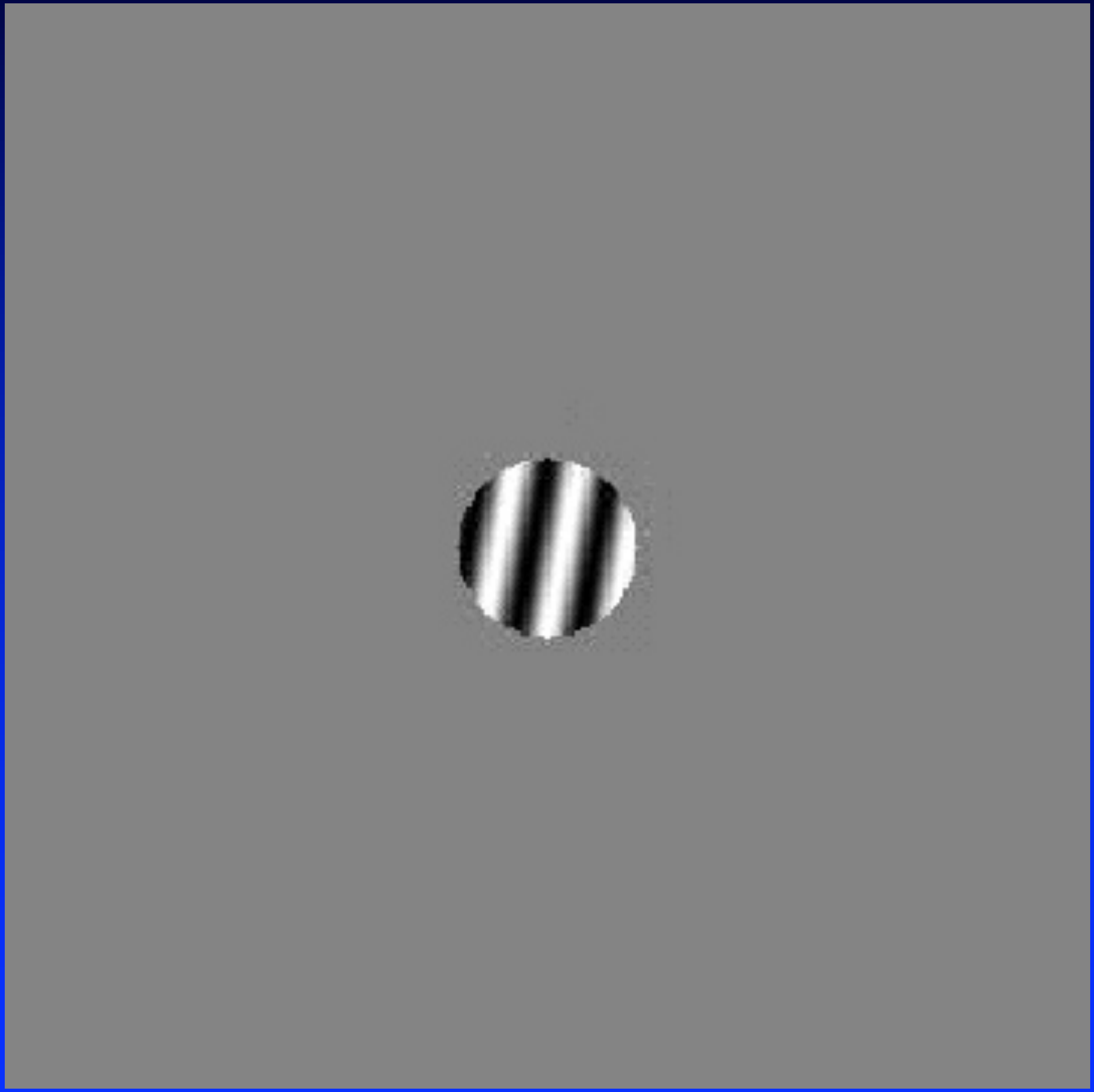
1) Category learning

2) Expertise model

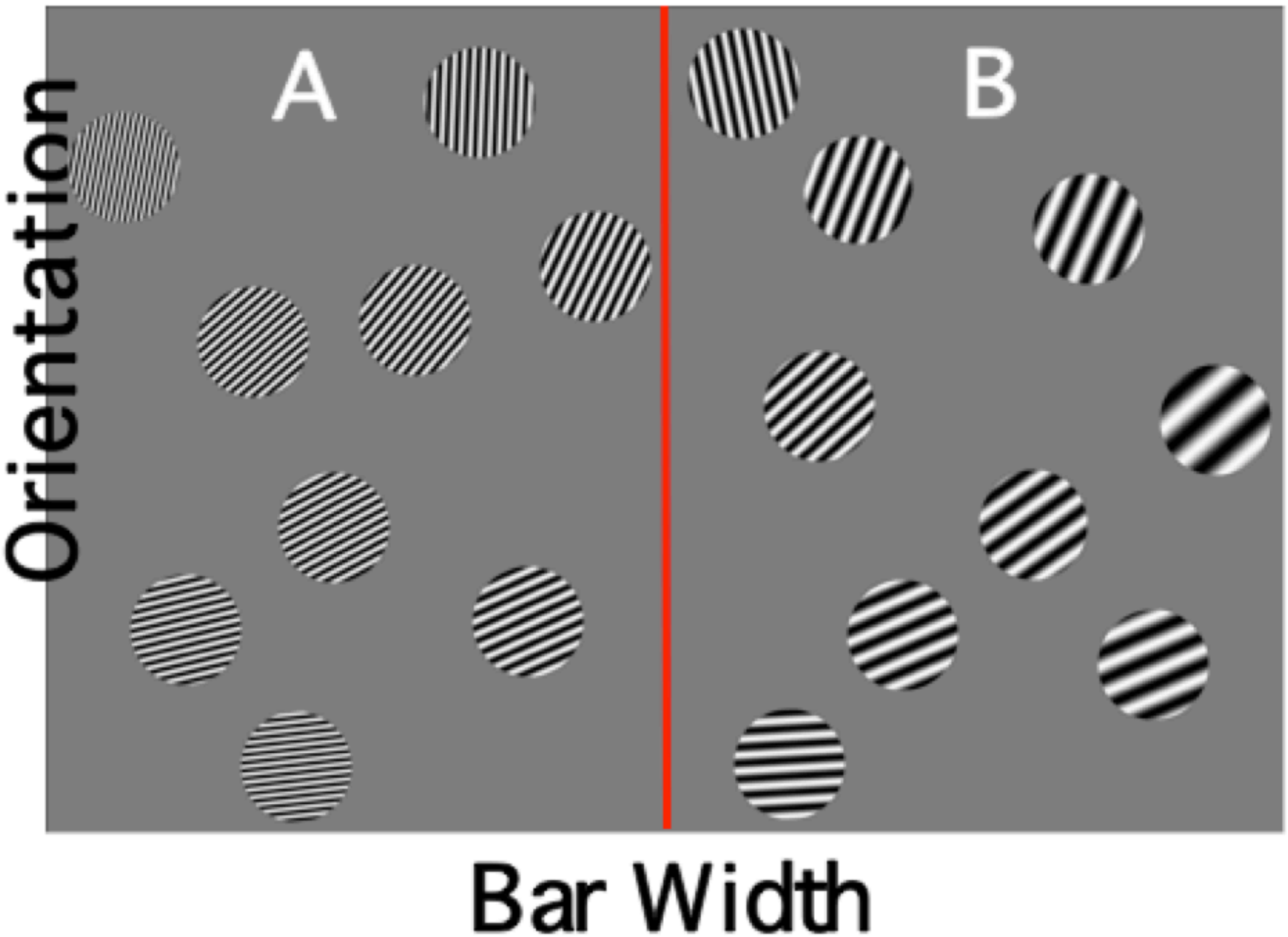
3) Tests of the model

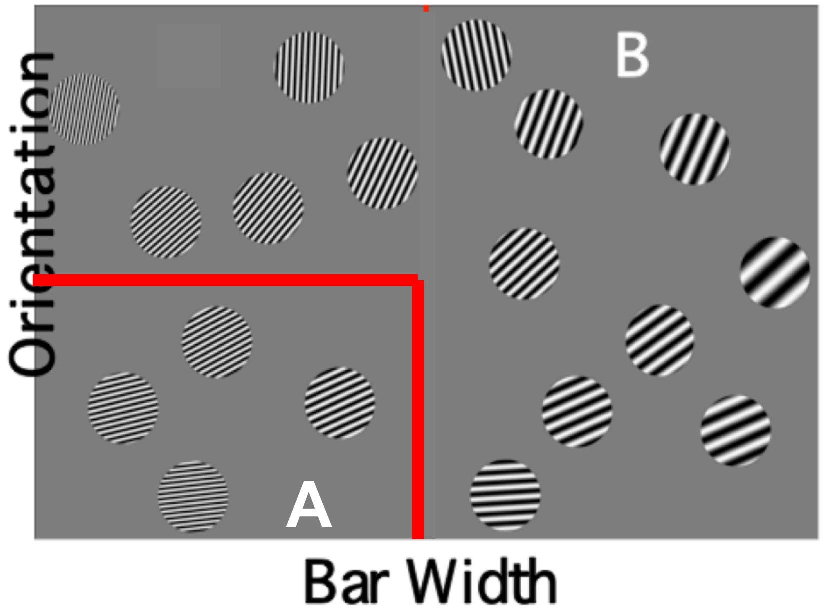
4) Future directions

Stimulus on a Single Category-Learning Trial



Rule-Based Category Learning





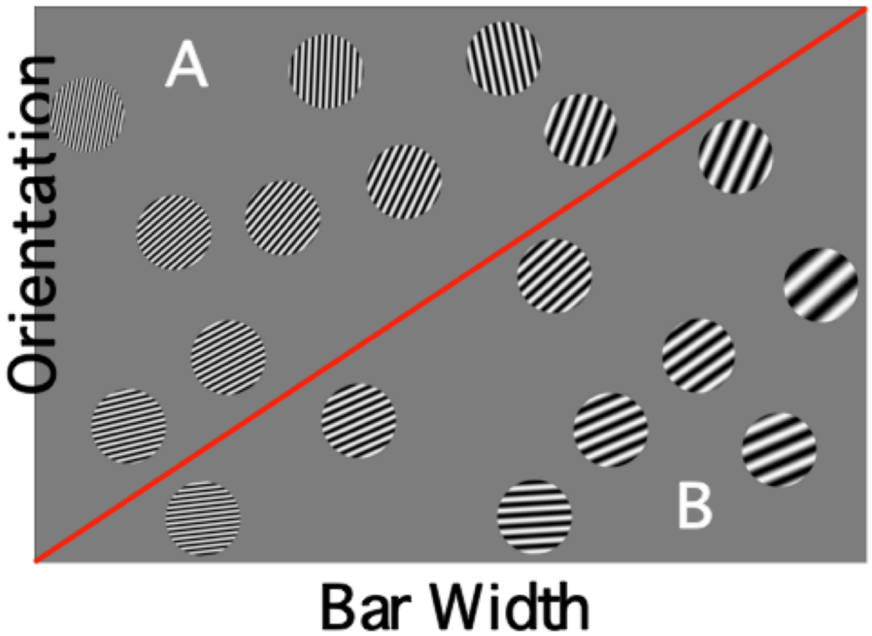
Rule-Based Category Learning

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

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

A Real-Life Information-Integration Task?

Does this mammogram show a tumor?

A Real-Life Information-Integration Task?

Tumor!

The Two Category Learning Systems of COVIS

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

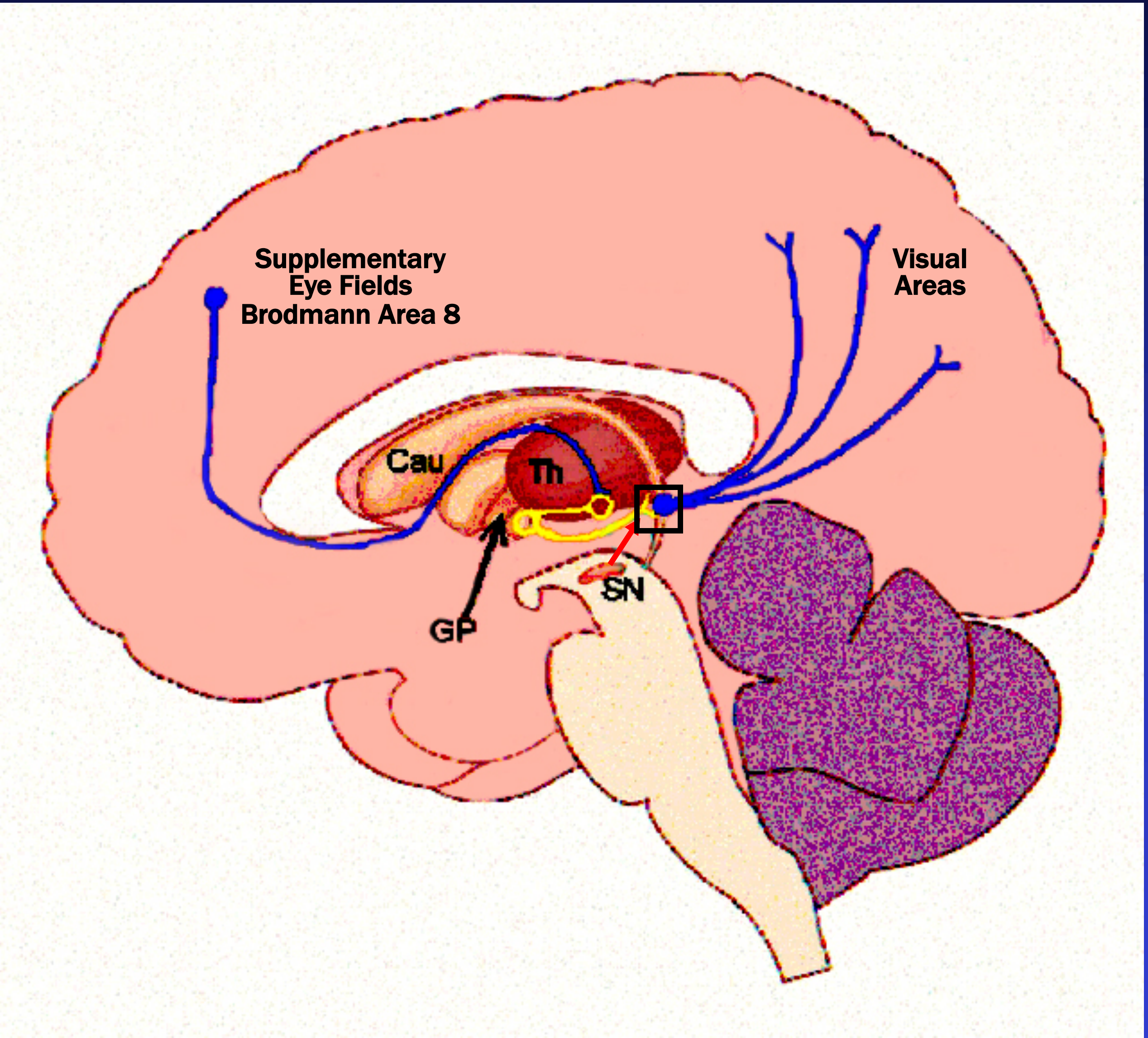
- Explicit, logical-reasoning system
  -- quickly learns explicit rules
- Procedural- or habit-learning system
  -- slowly learns almost any similarity-based rule
- Simultaneously active in all tasks (at least initially)

The COVIS Explicit System

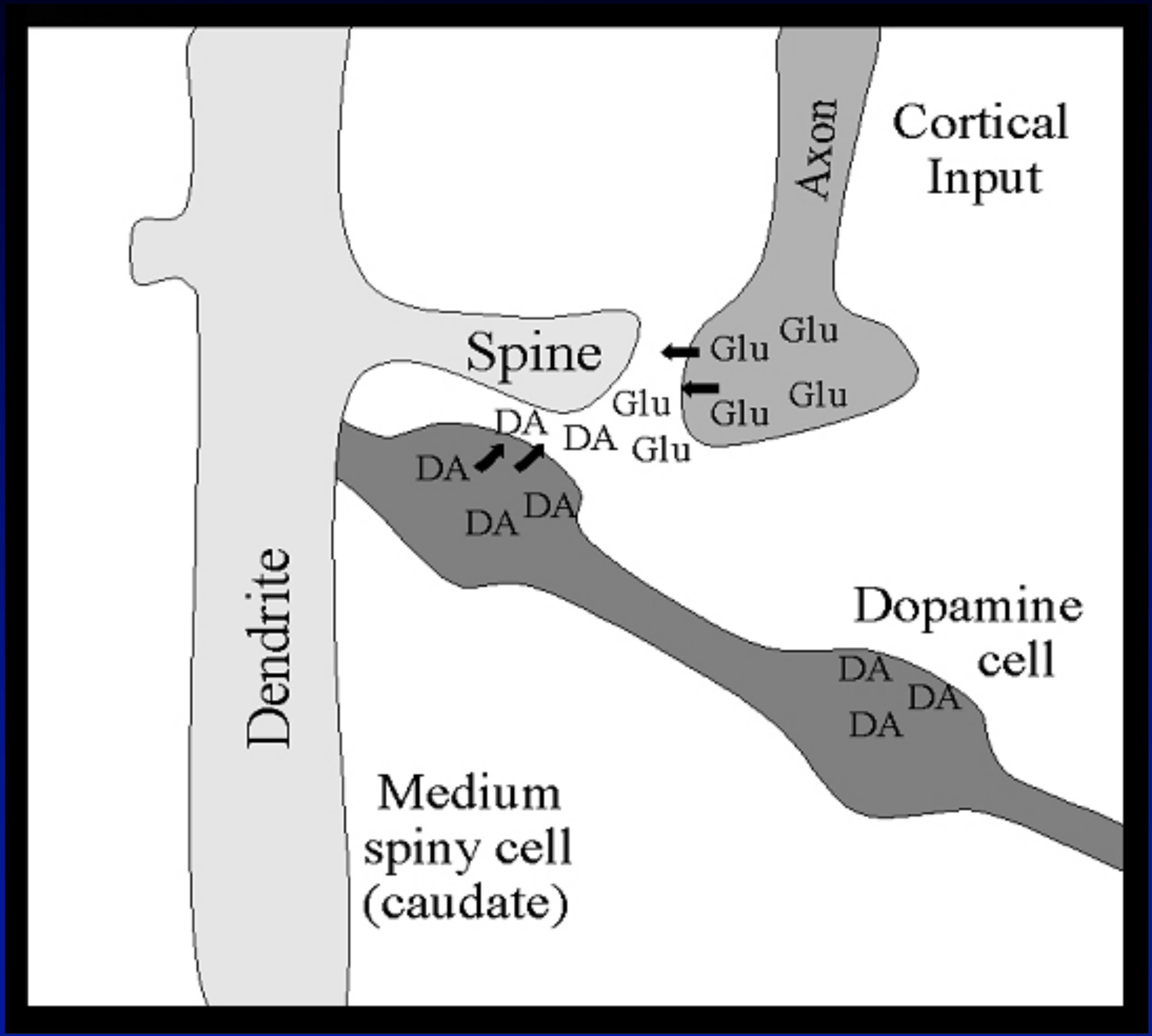
- Logical reasoning system
- Uses working memory and executive attention
- Prefrontal cortex, anterior cingulate, head of the caudate nucleus, thalamo-cortical loops, medial temporal lobe structures
- Working memory & attentional switching component – FROST (Ashby, Ell, Valentin, & Casale, 2005, *J. of Cognitive Neuroscience*)

The COVIS Procedural-Learning System

The Striatal Pattern Classifier (Ashby & Waldron, 1999)



A Closer View of a Cortical-Striatal Synapse



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

Neuroimaging experiments

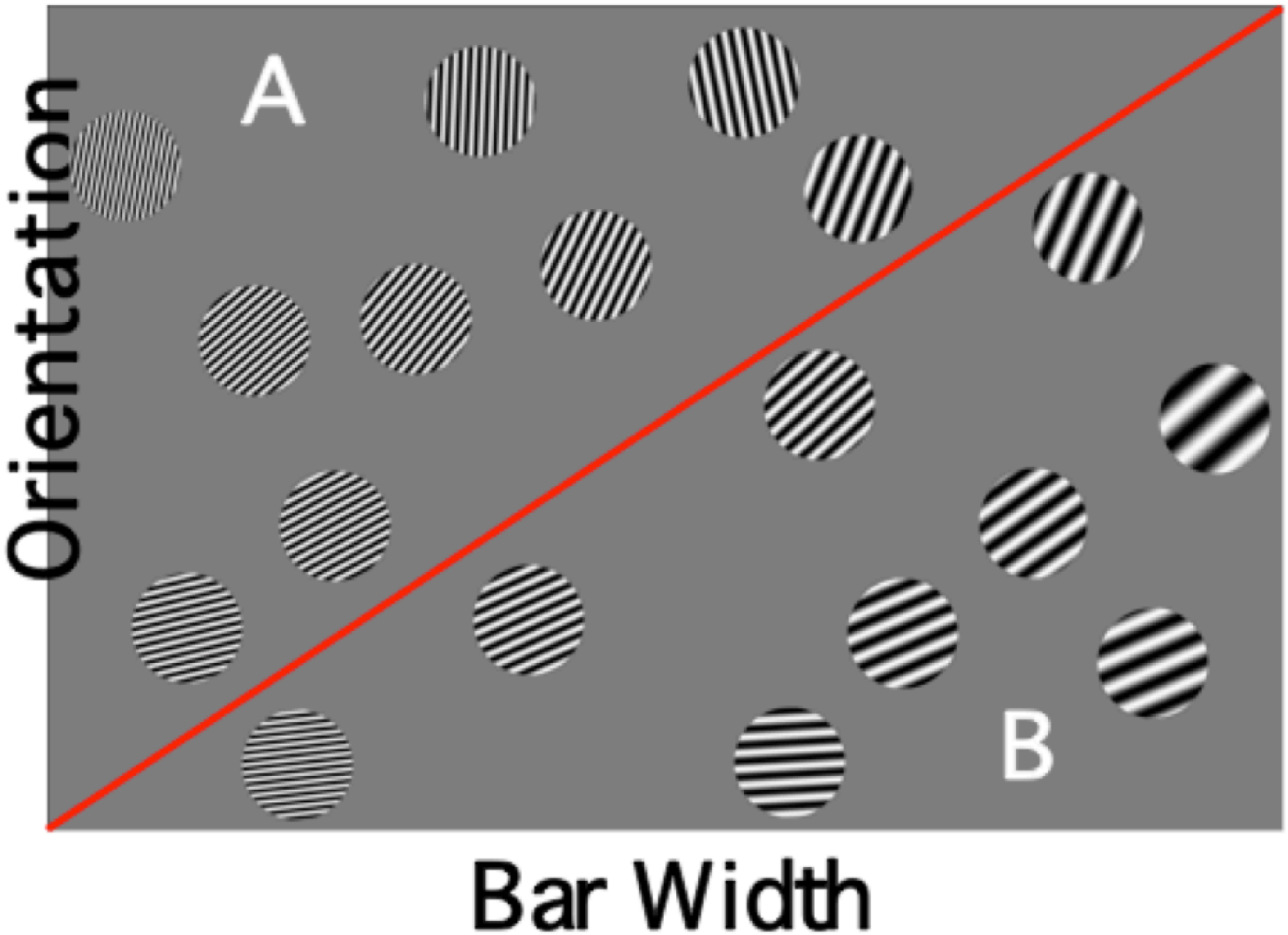
Konishi et al., 1999; Lombardi et al., 1999; Nomura et al., in press; Poldrack, et al., 2001; Rao et al., 1997; Rogers, Andrews, Grasby, Brooks, & Robbins, 2000; Seger & Cincotta, 2002; Volz et al., 1997

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

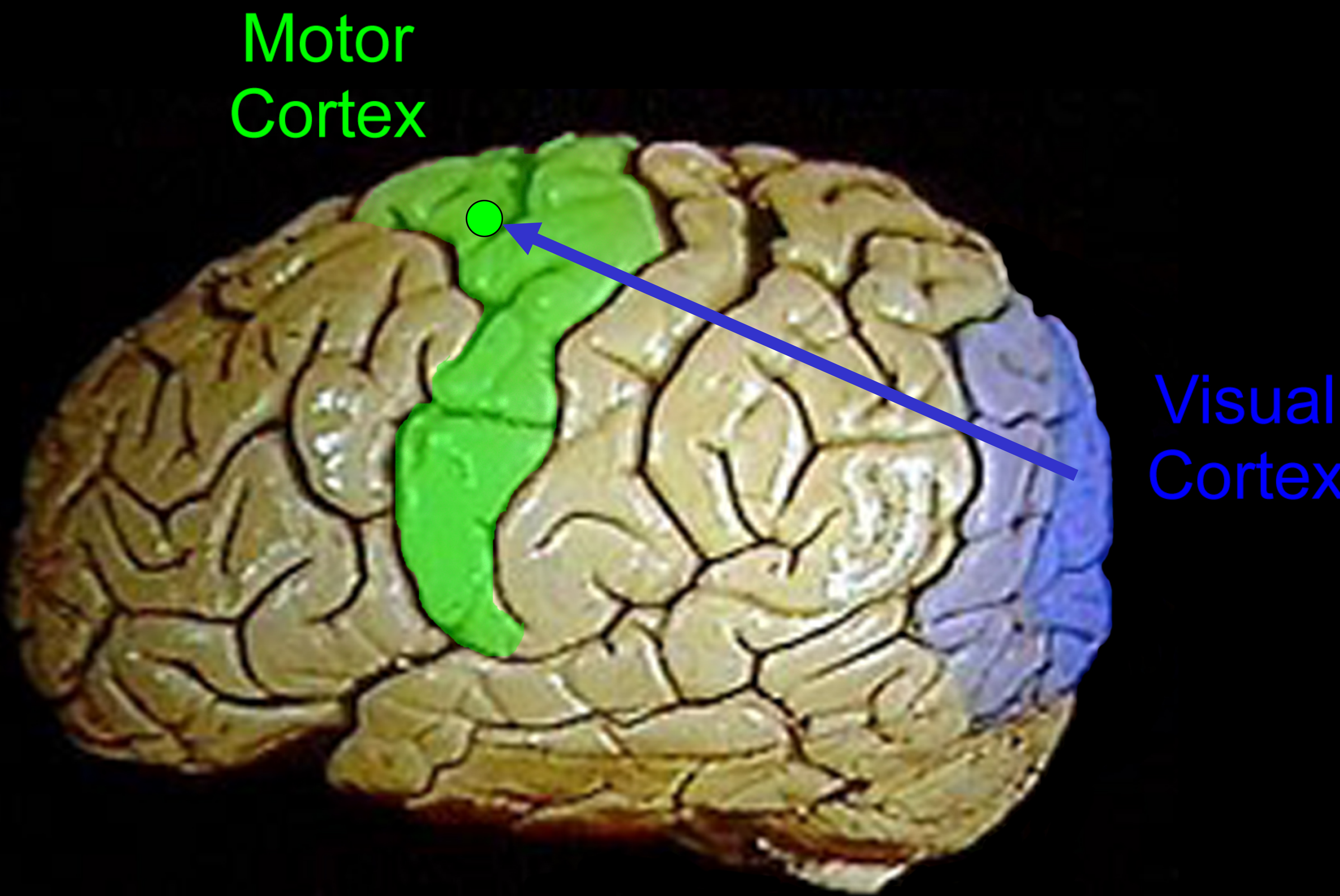
Information-Integration Category Learning



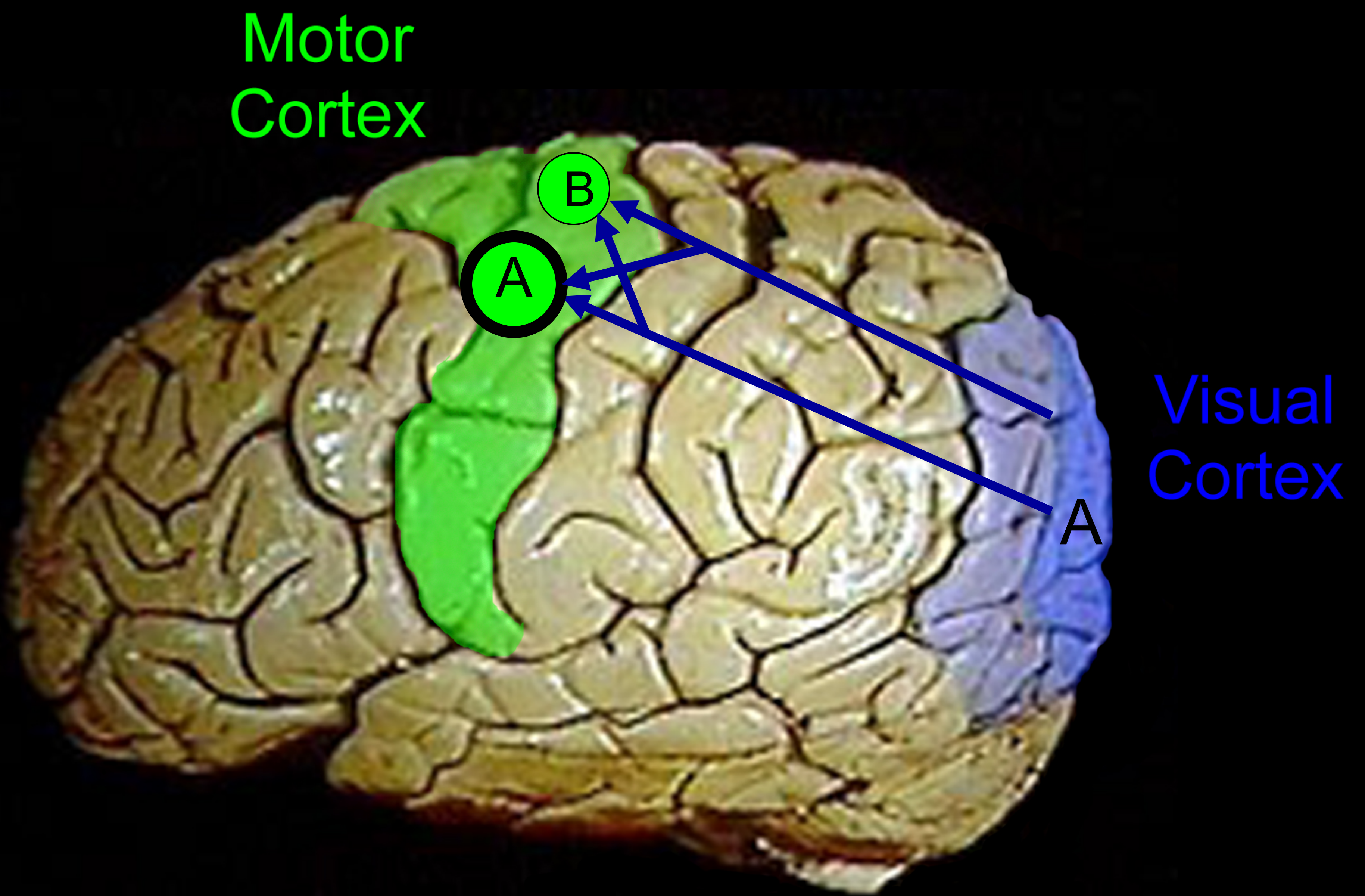
A Double Dissociation?

	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

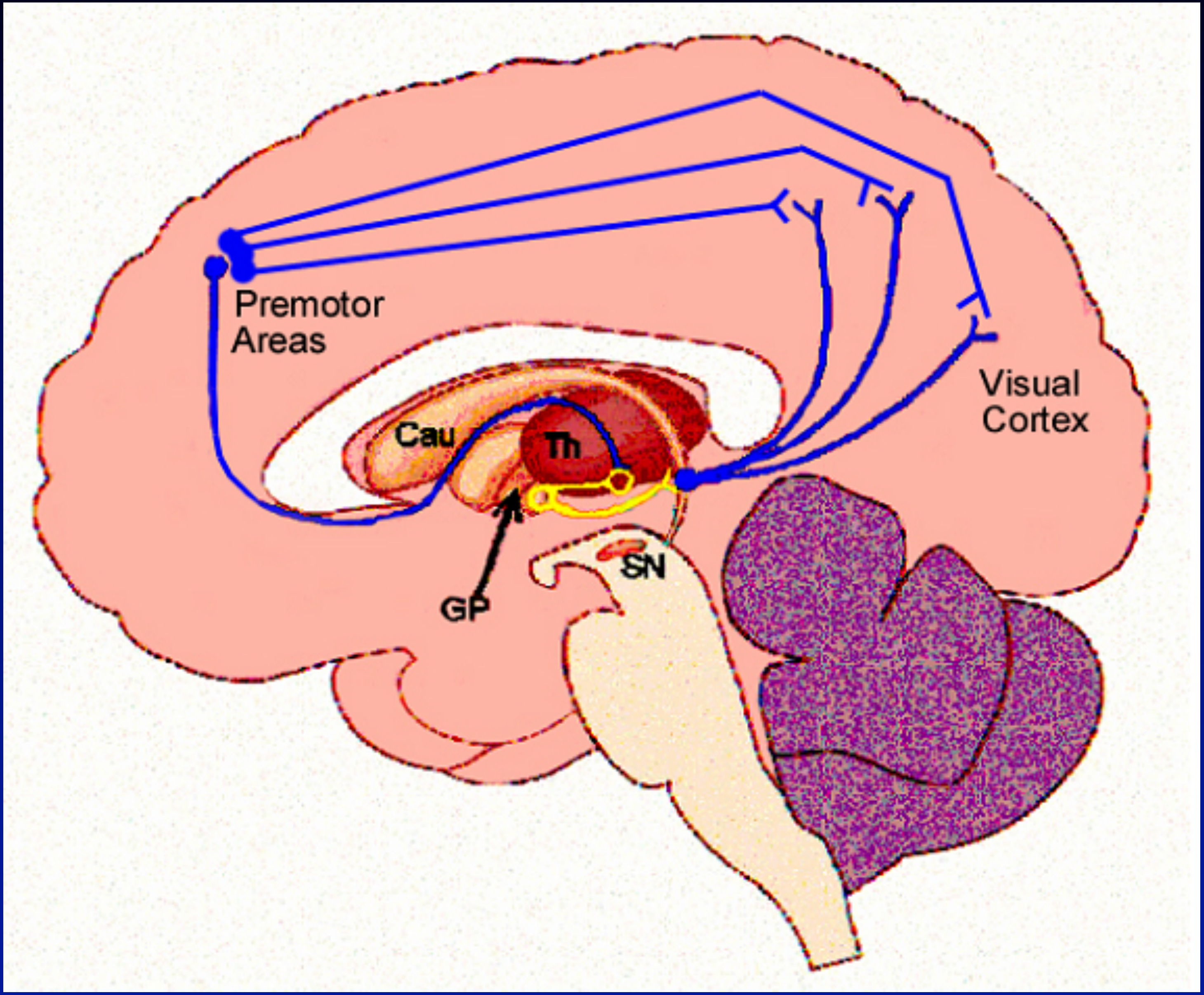
Building a Model of Expertise



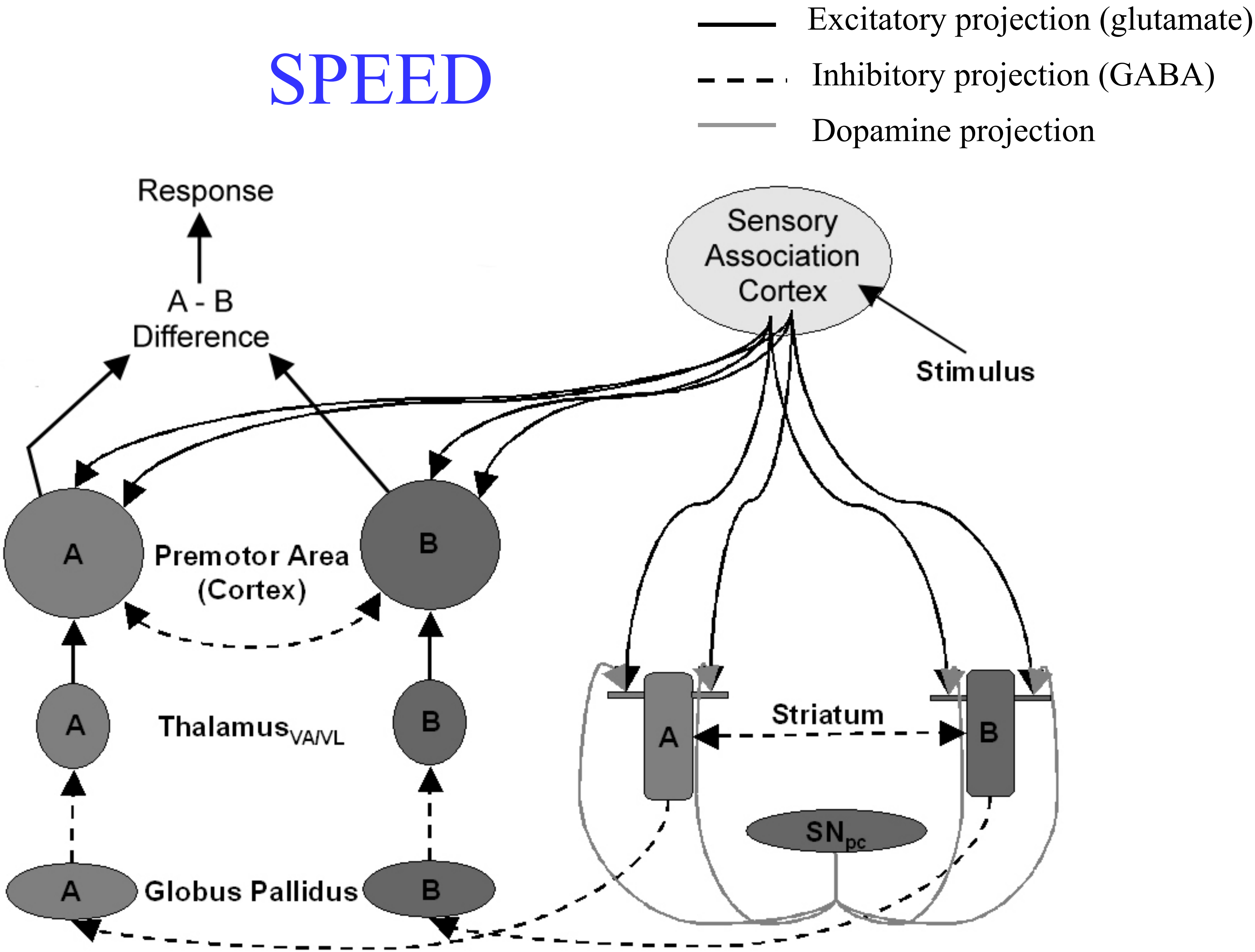
Building a Model of Expertise



A Model of Categorization Automaticty in II Tasks



SPEED


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

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 ?(t) is white noise.

Modeling Activation in Other Units

Globus Pallidus:

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

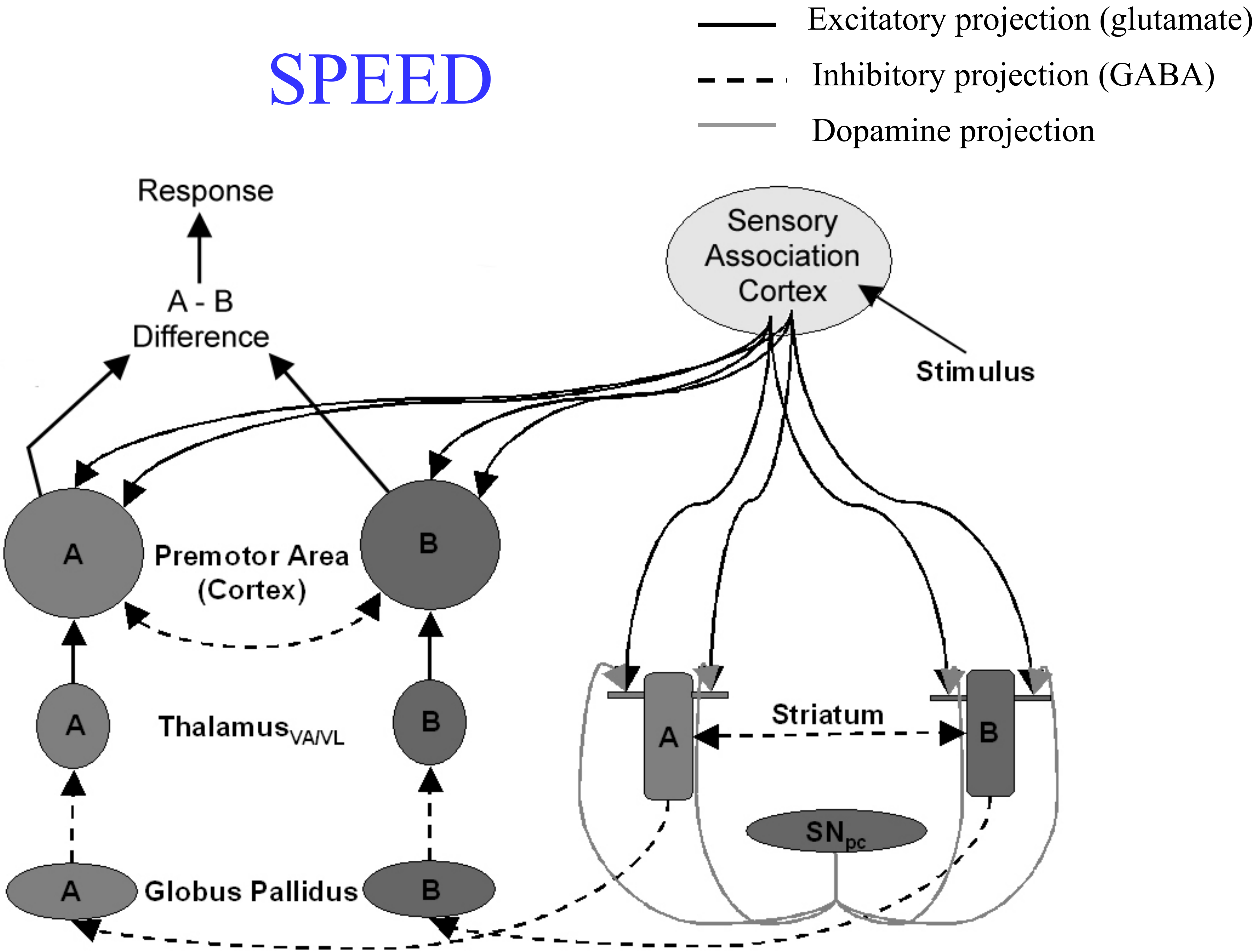
Thalamus:

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

Premotor Area:

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

SPEED



Excitatory projection (glutamate)
Inhibitory projection (GABA)
Dopamine projection
Response
A - B
Difference
Sensory
Association
Cortex
Stimulus
A
Premotor Area
(Cortex)
B
A
$\text{Thalamus}_{VA/VL}$
B
Striatum
A
B
SN_{pc}
A
Globus Pallidus
B

Cortical-Cortical Learning (Hebbian)

$v_{K,J}(n)$ = strength of synapse between visual cortical cell K and premotor cell J on trial n

LTP

presynaptic activation

postsynaptic activation (above NMDA threshold)

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

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

LTD

postsynaptic activation (below NMDA threshold)

Cortical-Striatal Learning (Reward Mediated)

LTP

activation above NMDA threshold

dopamine above baseline (Correct Response)

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

dopamine below baseline (error)

activation below NMDA threshold

activation above NMDA threshold

LTD

Modeling Dopamine Release

If the response on trial *n* was correct:

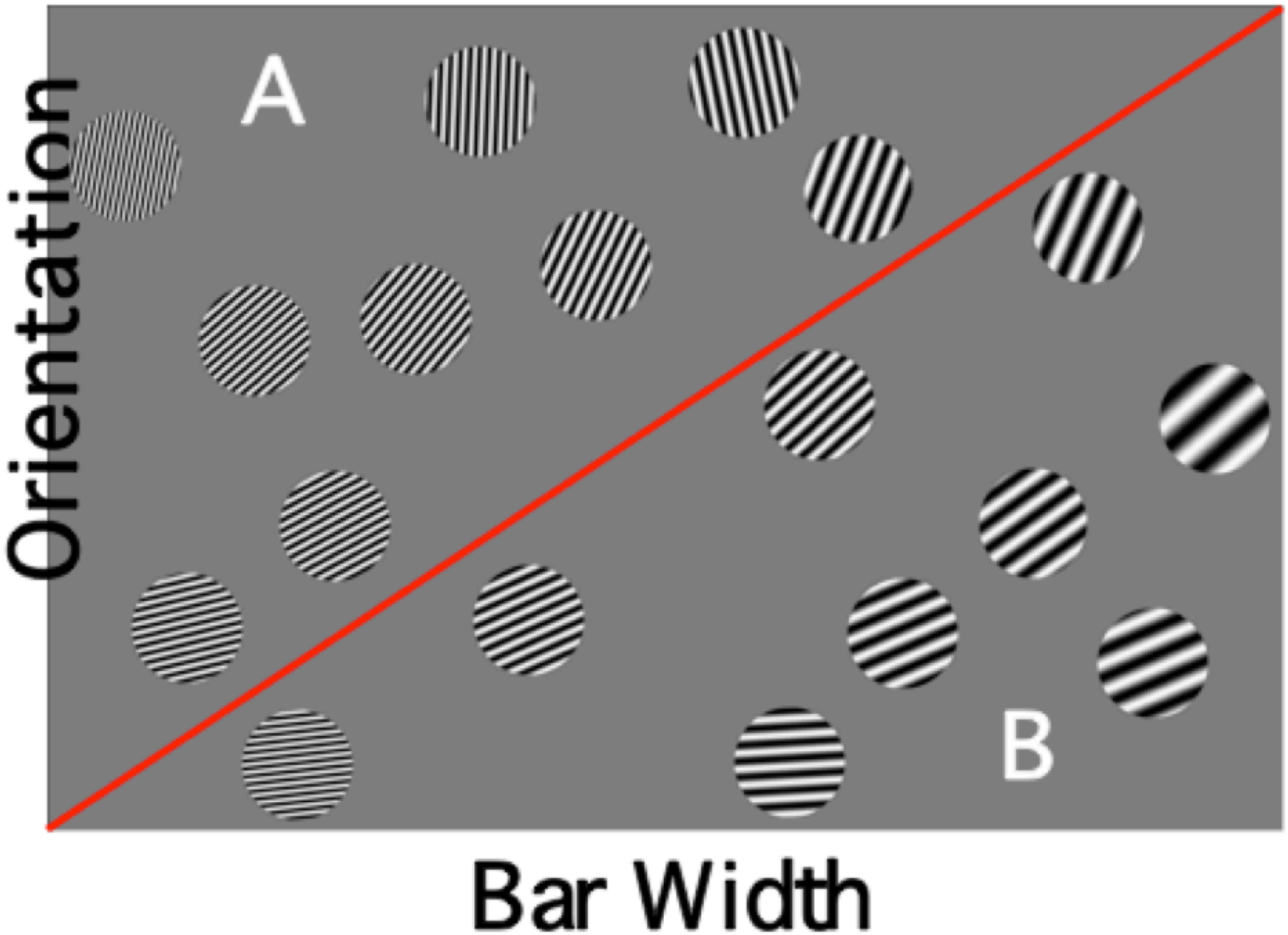
$$D(n) = D_{base} + [1 - P(C)](1 - D_{base}),$$

where $P(C)$ is the current probability of a correct response.

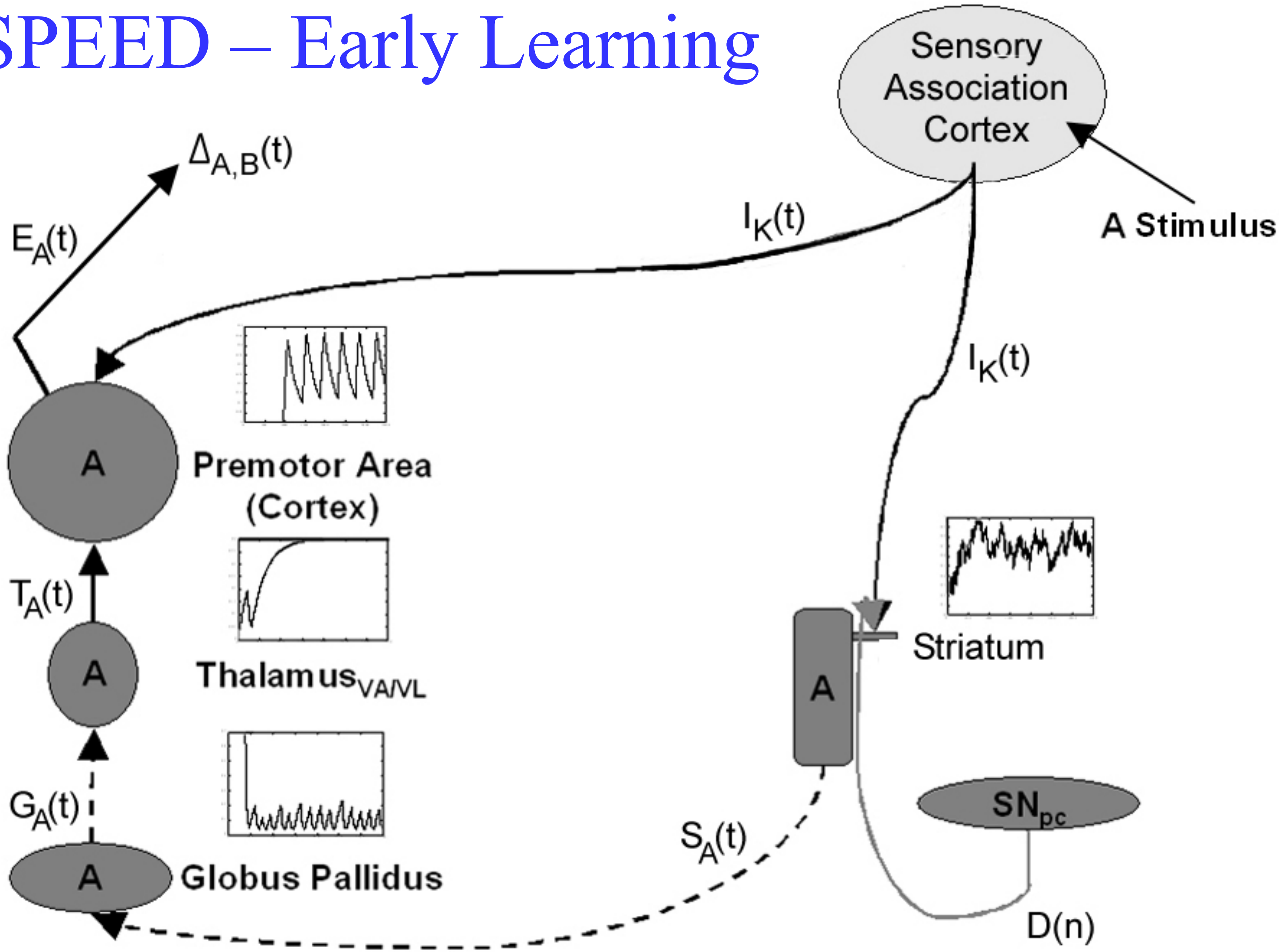
If the response on trial *n* was incorrect:

$$D(n) = D_{base} - P(C)D_{base}$$

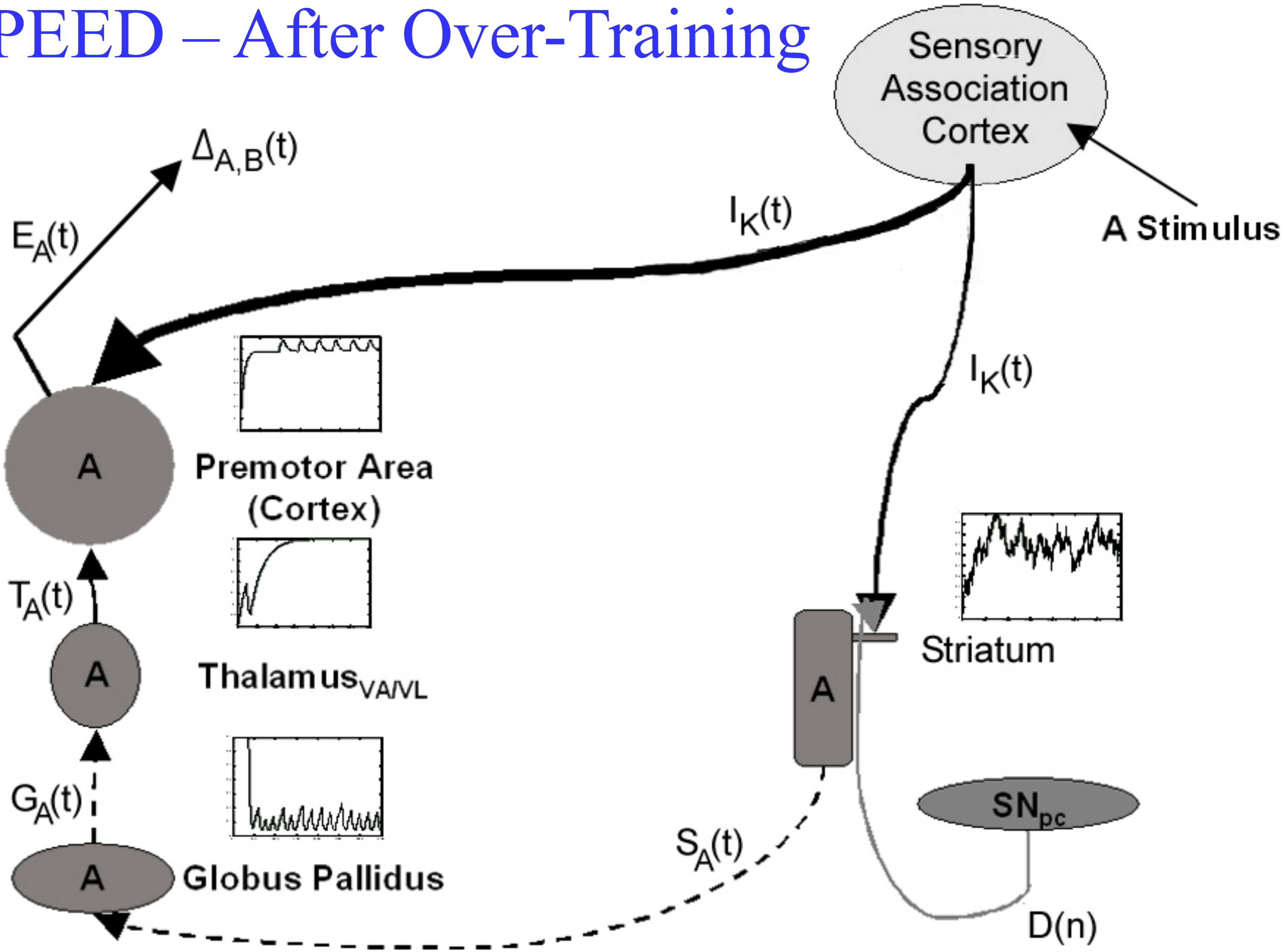
Information-Integration Category Learning



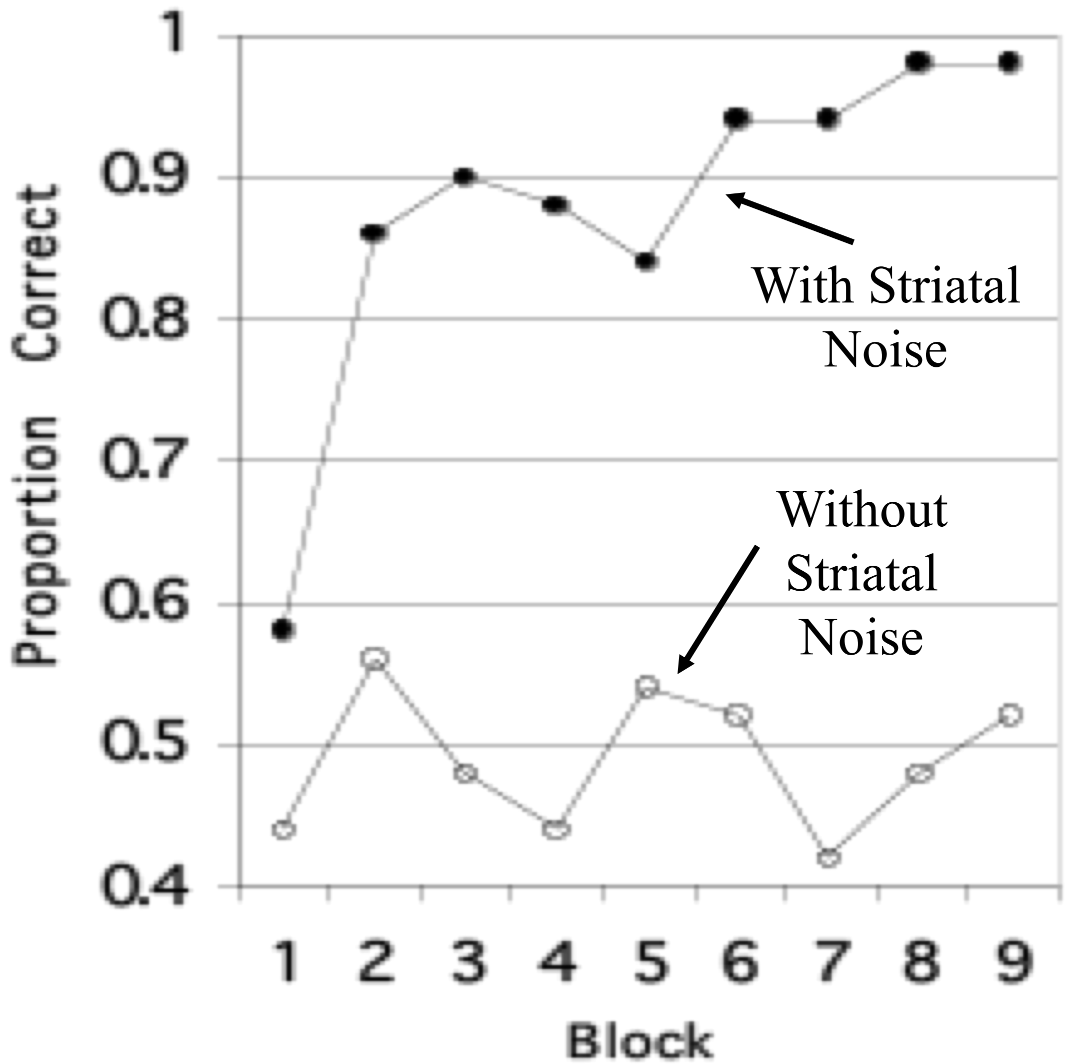
SPEED – Early Learning



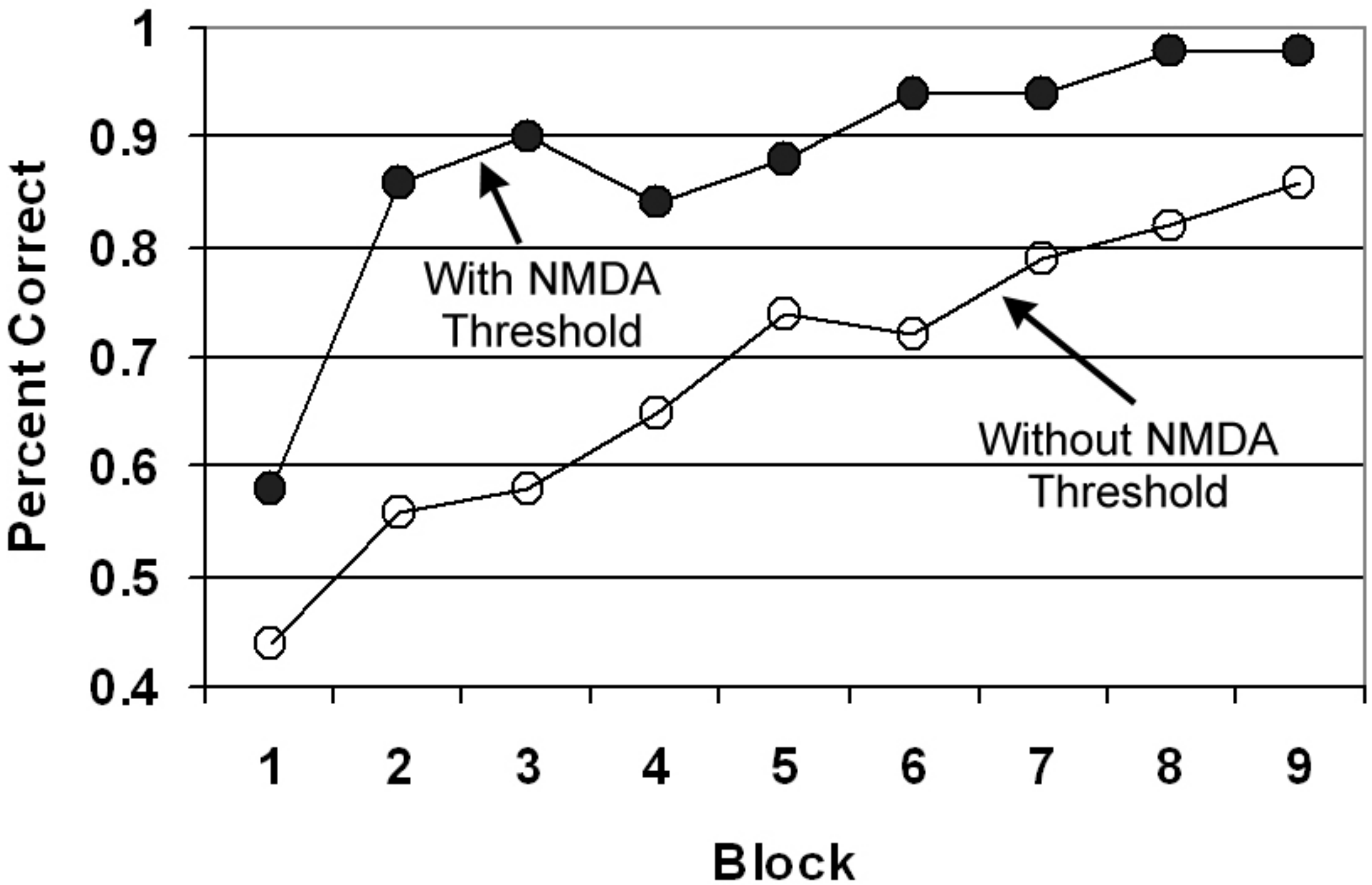
SPEED – After Over-Training



II Learning With and Without Striatal Noise



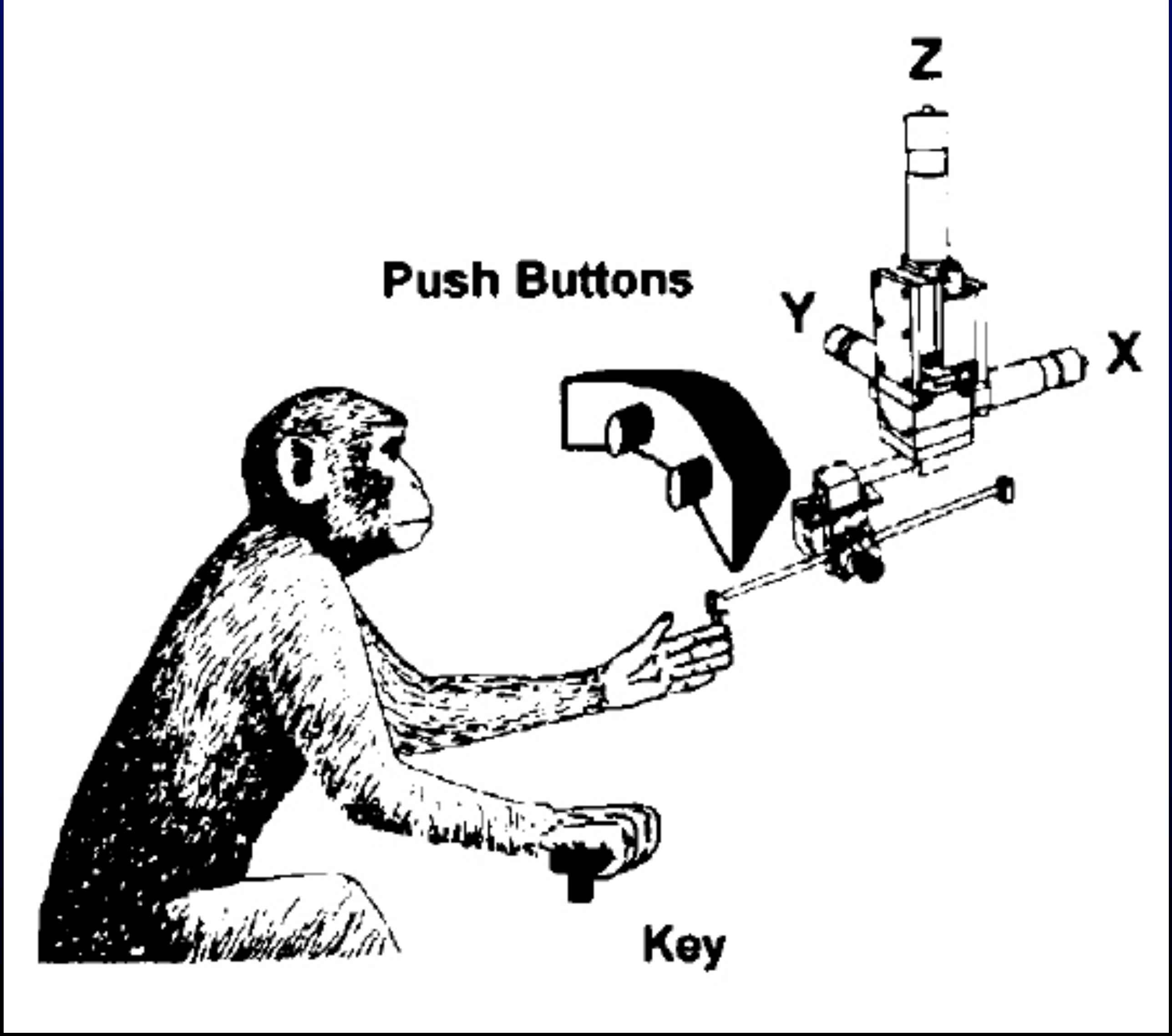
II Learning With and Without NMDA Threshold



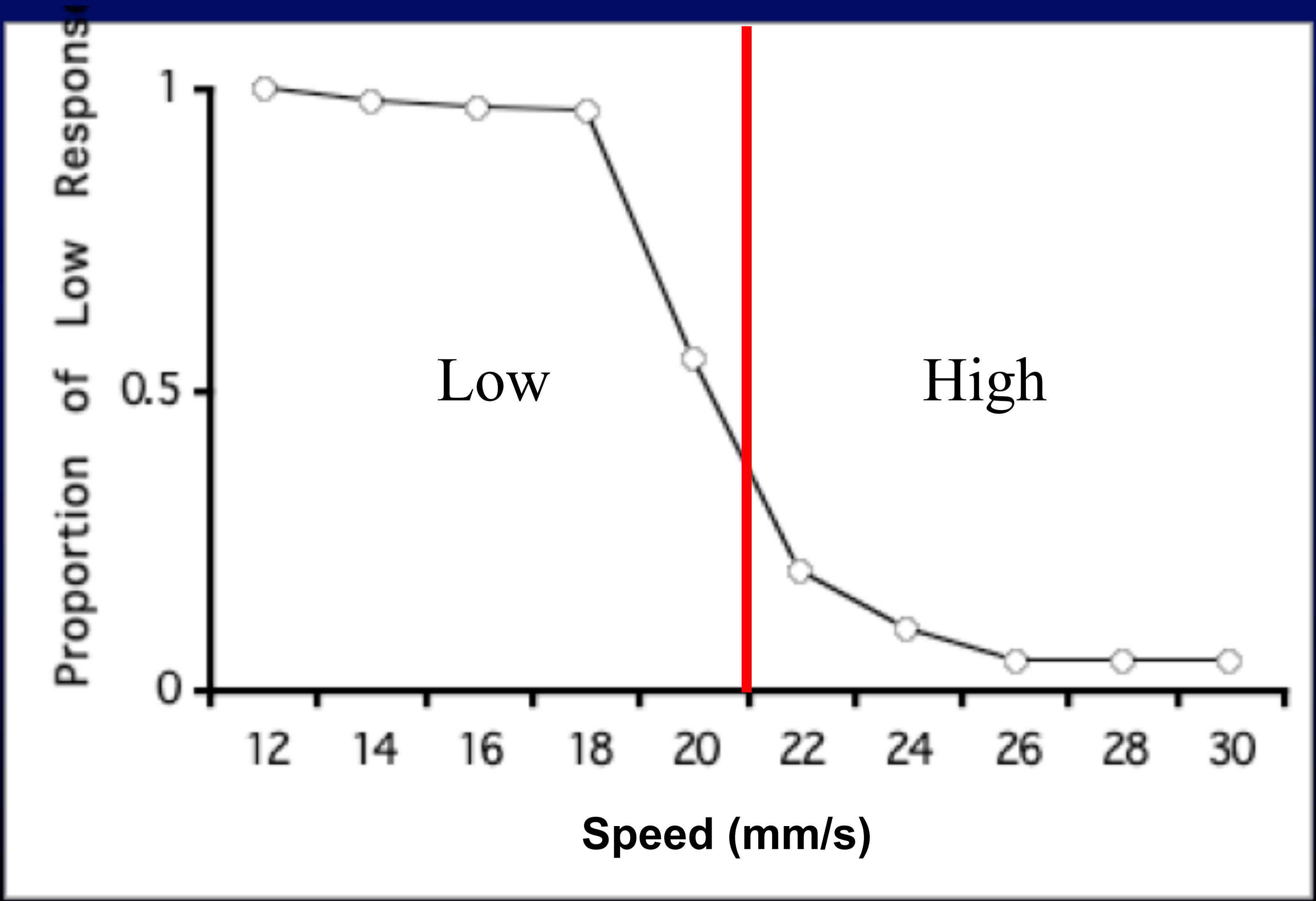
Experimental Tests

Tactile Category Learning

Romo, Merchant et al.



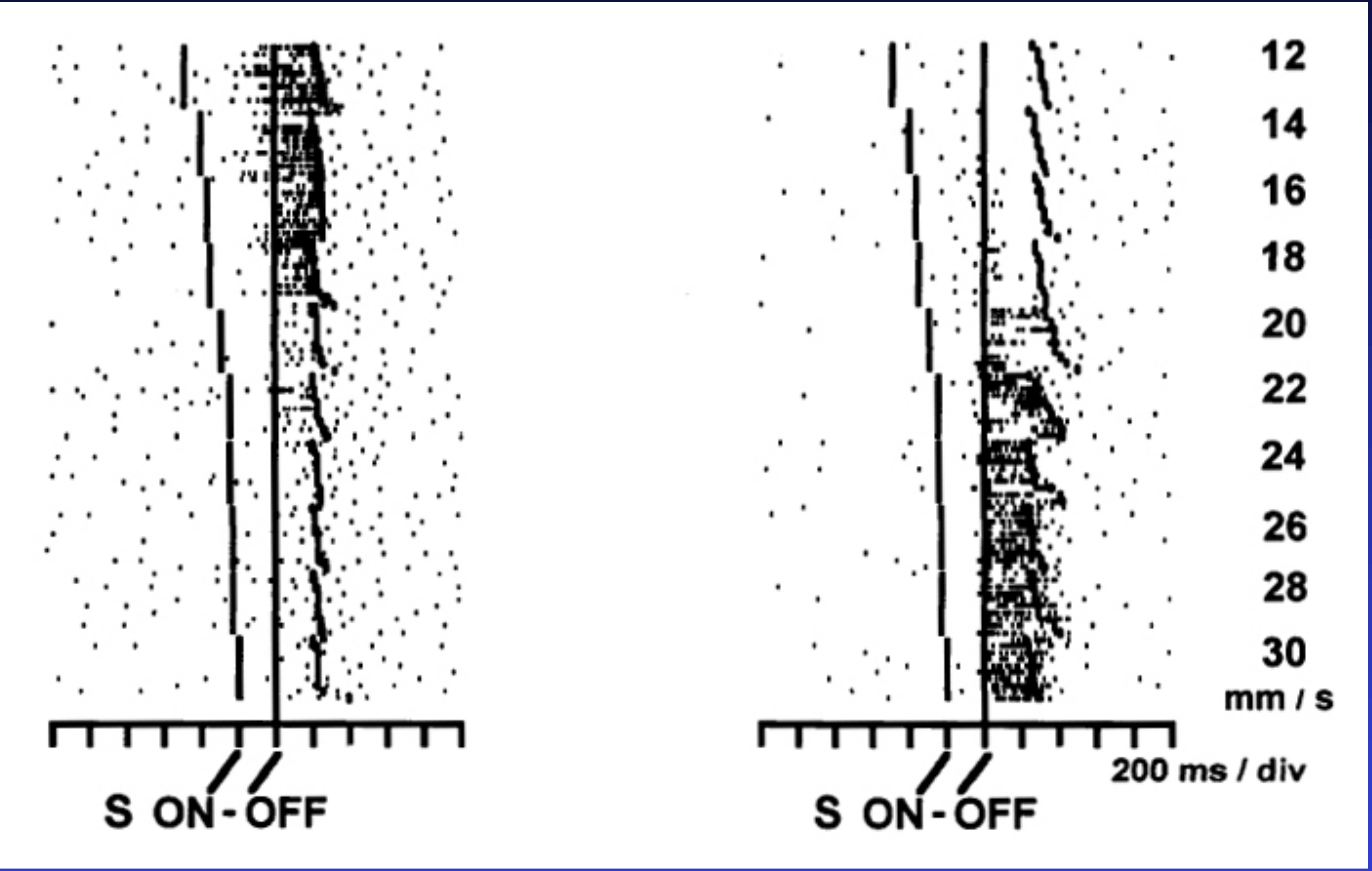
Proportion of "Low" Responses



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

Single Cell Responses – Left Putamen

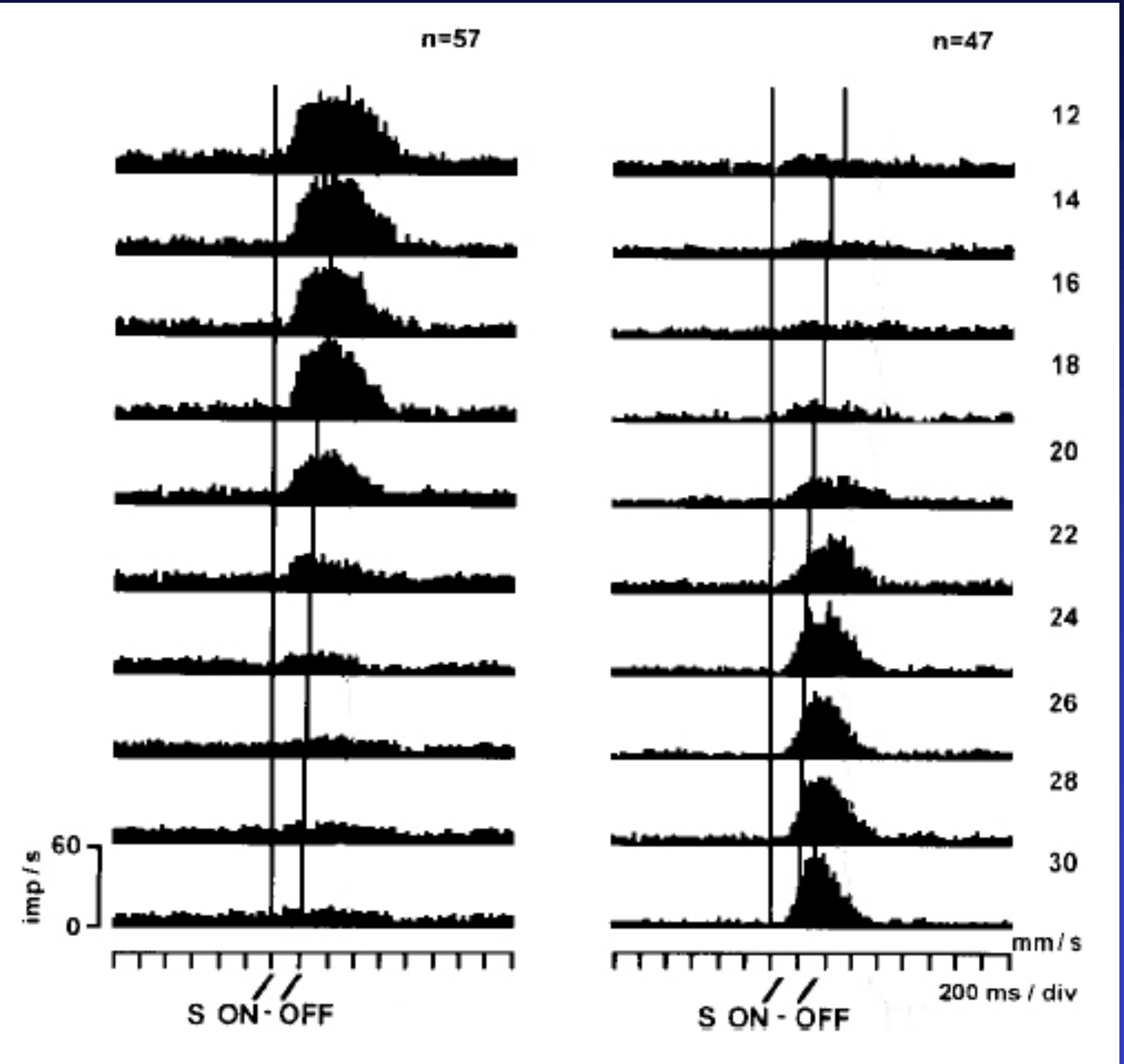
Low Speed Cell | High Speed Cell



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

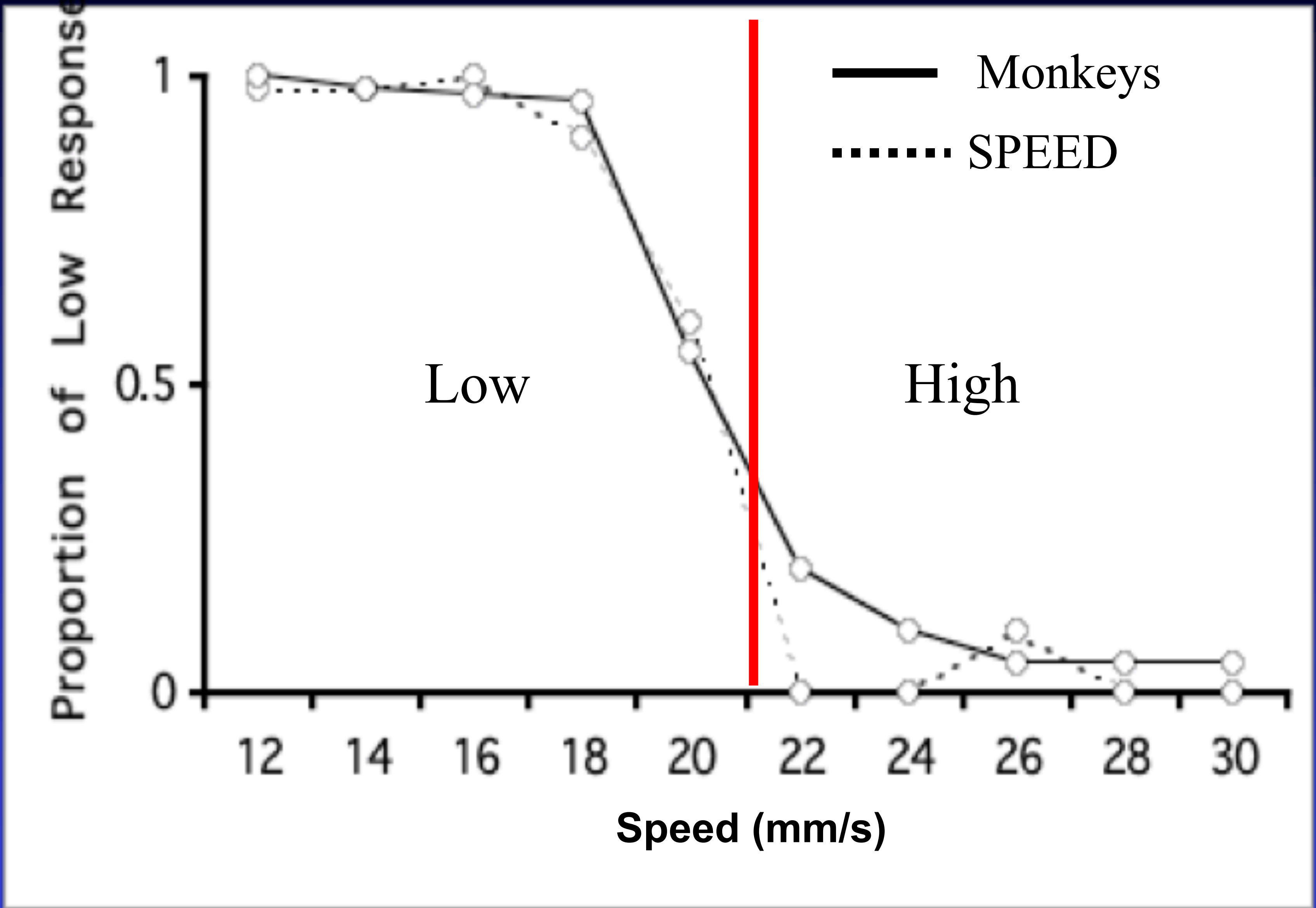
Population Responses – Premotor Cortex

Low Speed Cells High Speed Cells



Romo et al. (1997, Cerebral Cortex)

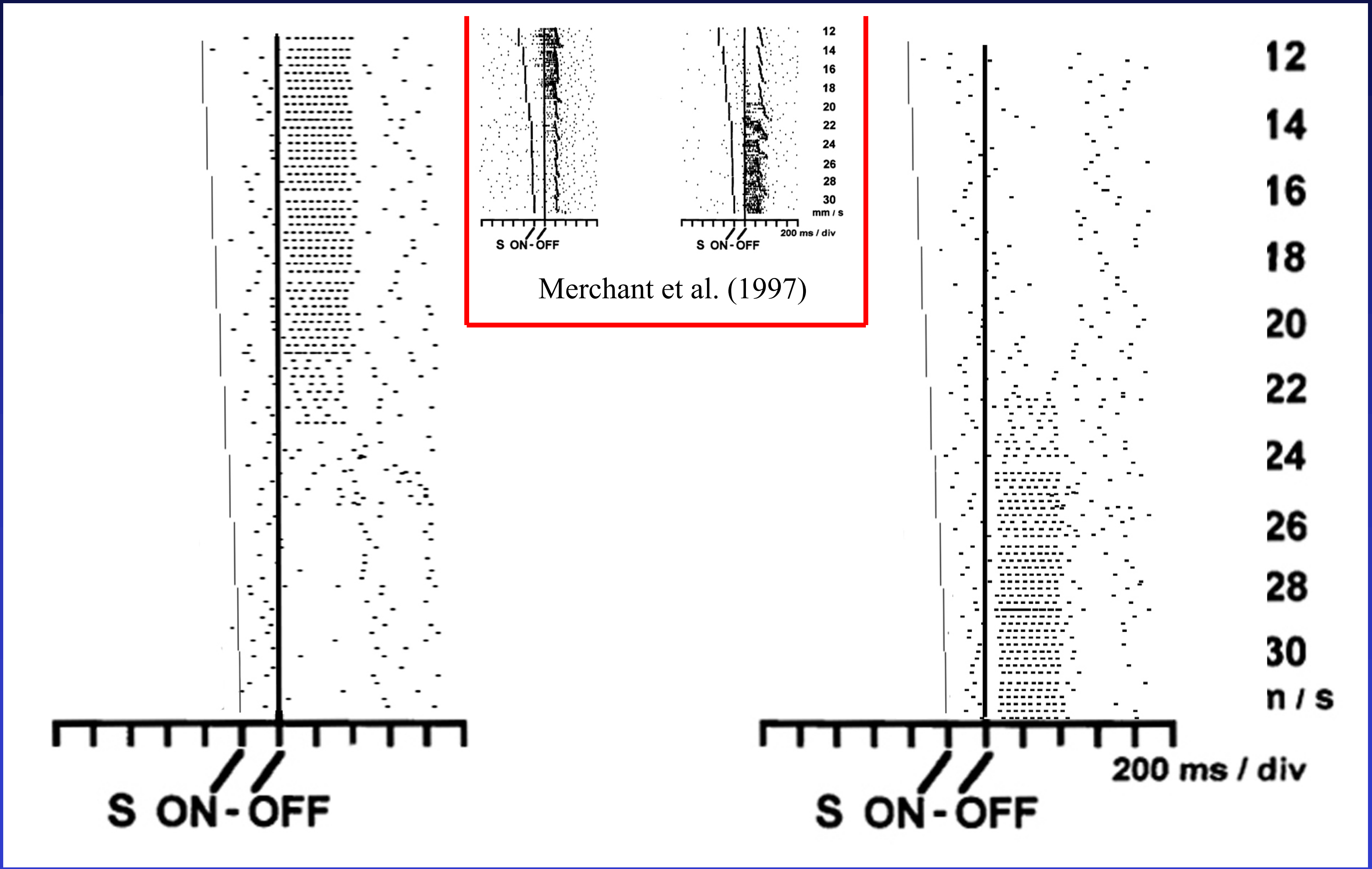
Proportion of "Low" Responses



SPEED's Single Cell Responses –Putamen

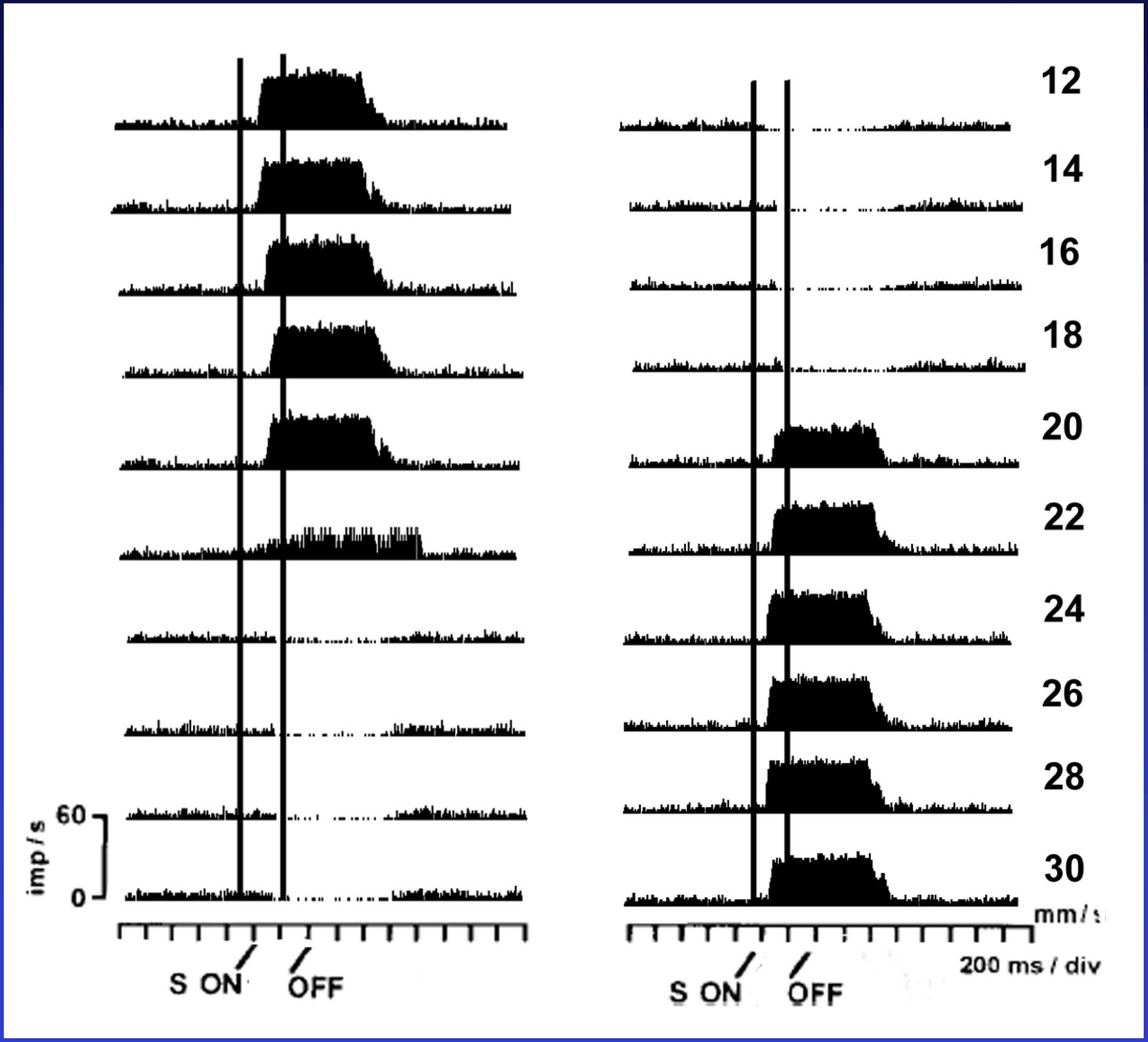
Low Speed Cell

High Speed Cell

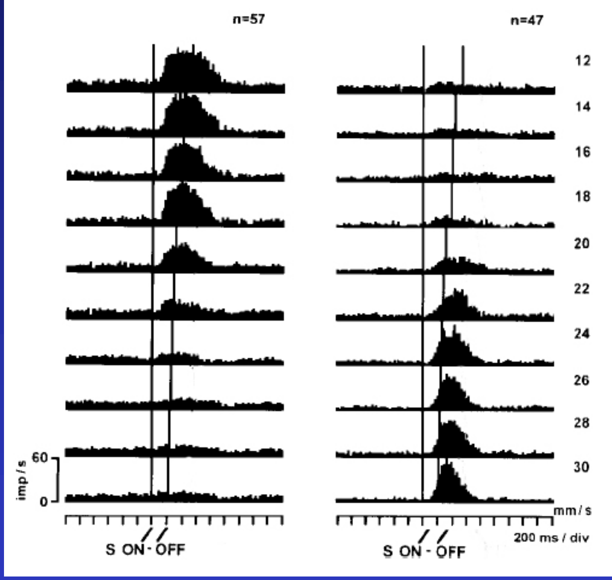


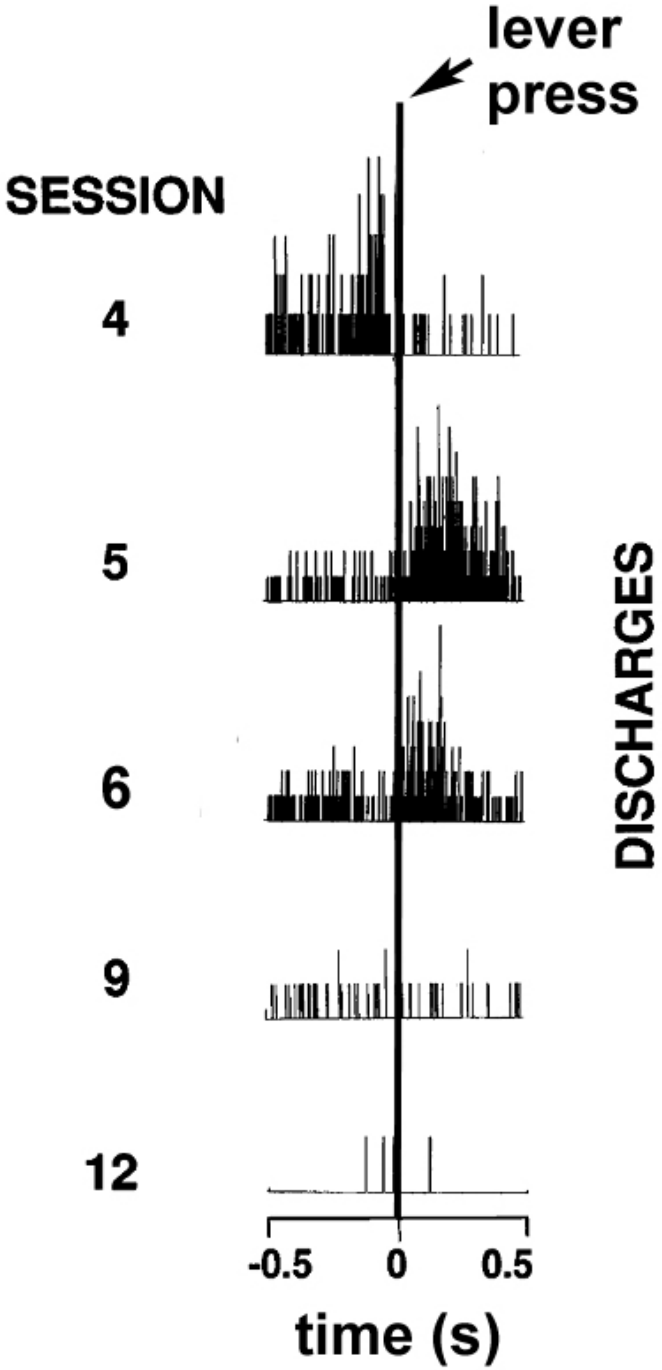
SPEED's Responses – Premotor Cortex

Low Speed Cells High Speed Cells



Romo et al., 1997





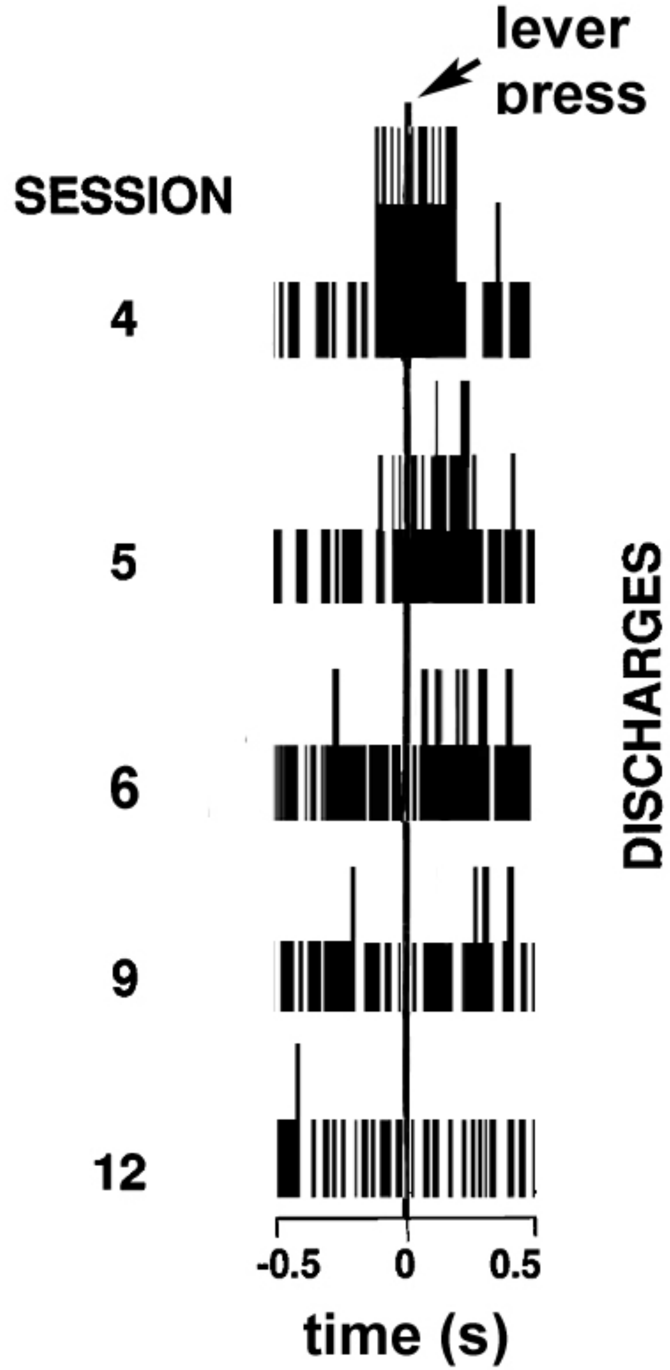
Carelli, Wolske, & West (1997, Journal of Neuroscience)

Lever press to tone

70 trials/day

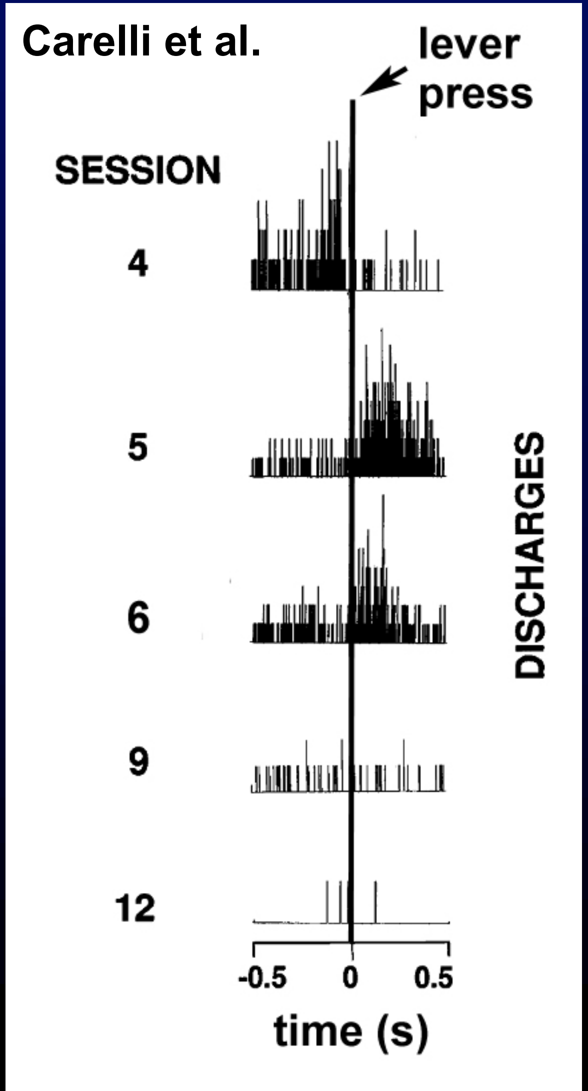
18 days

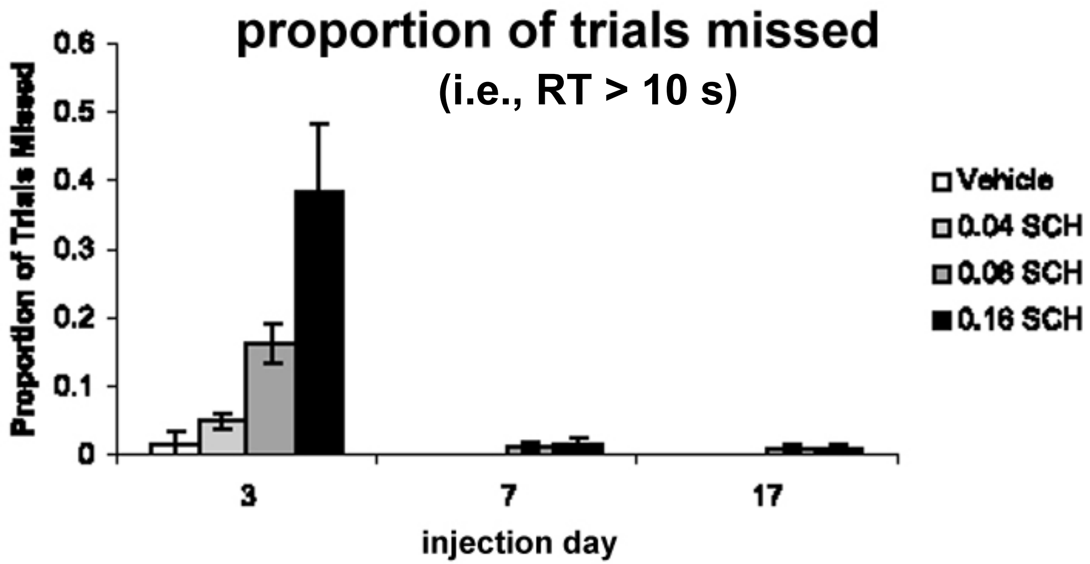
Striatal Response

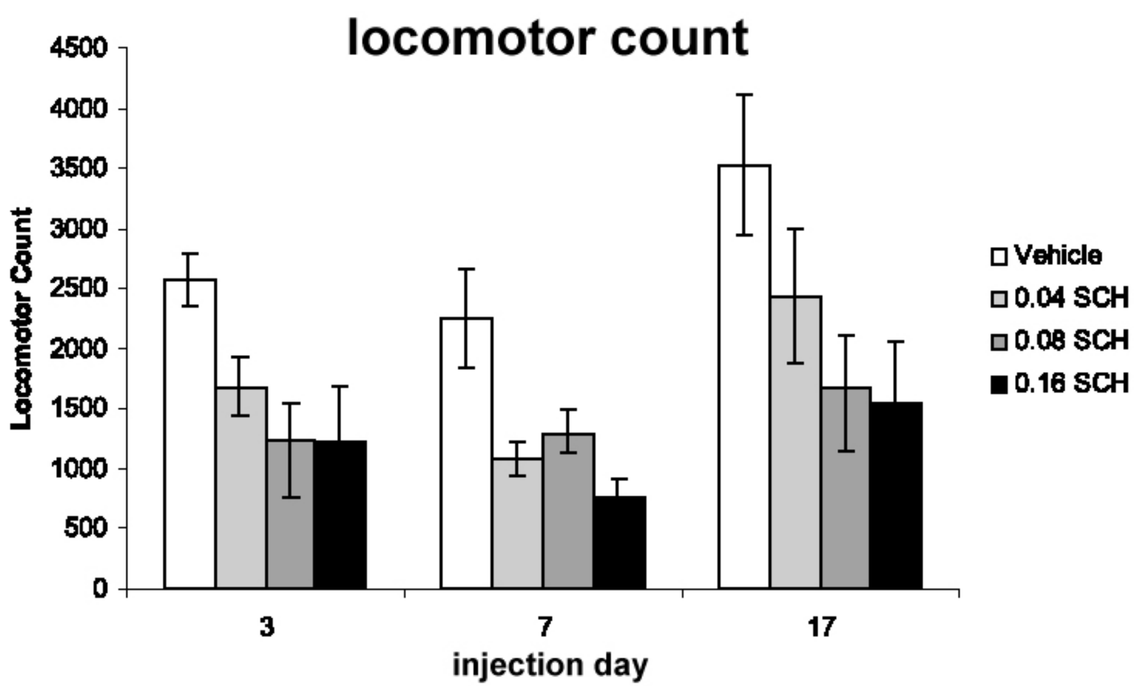


SPEED’s Striatal Responses

Carelli et al. (1997, Journal of Neuroscience)







Choi, Balsam, & Horvitz (2005, J. of Neuroscience)

Food pellet dropped into compartment

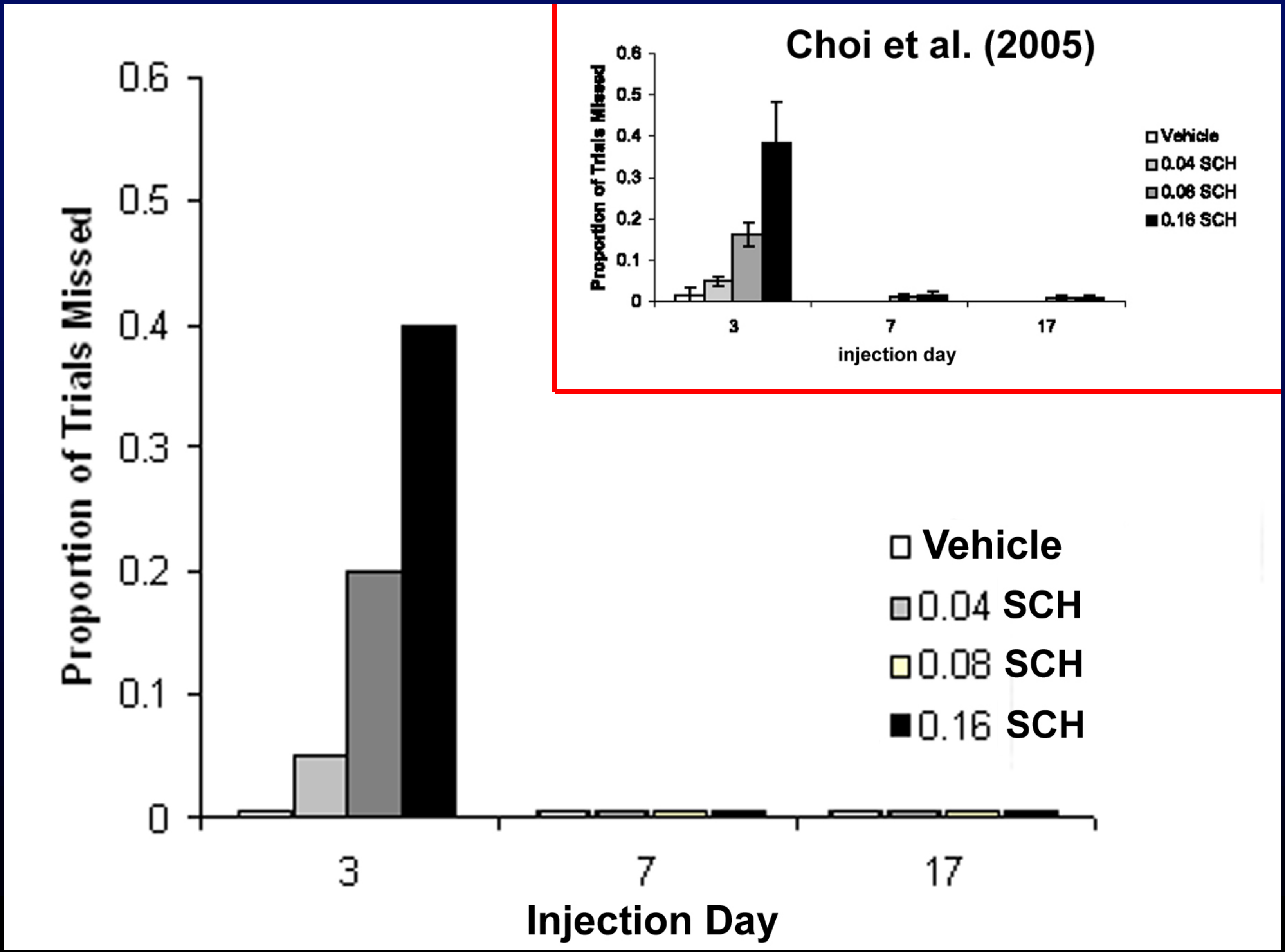
Minimum 30 s between trials

28 trials/day

17 days

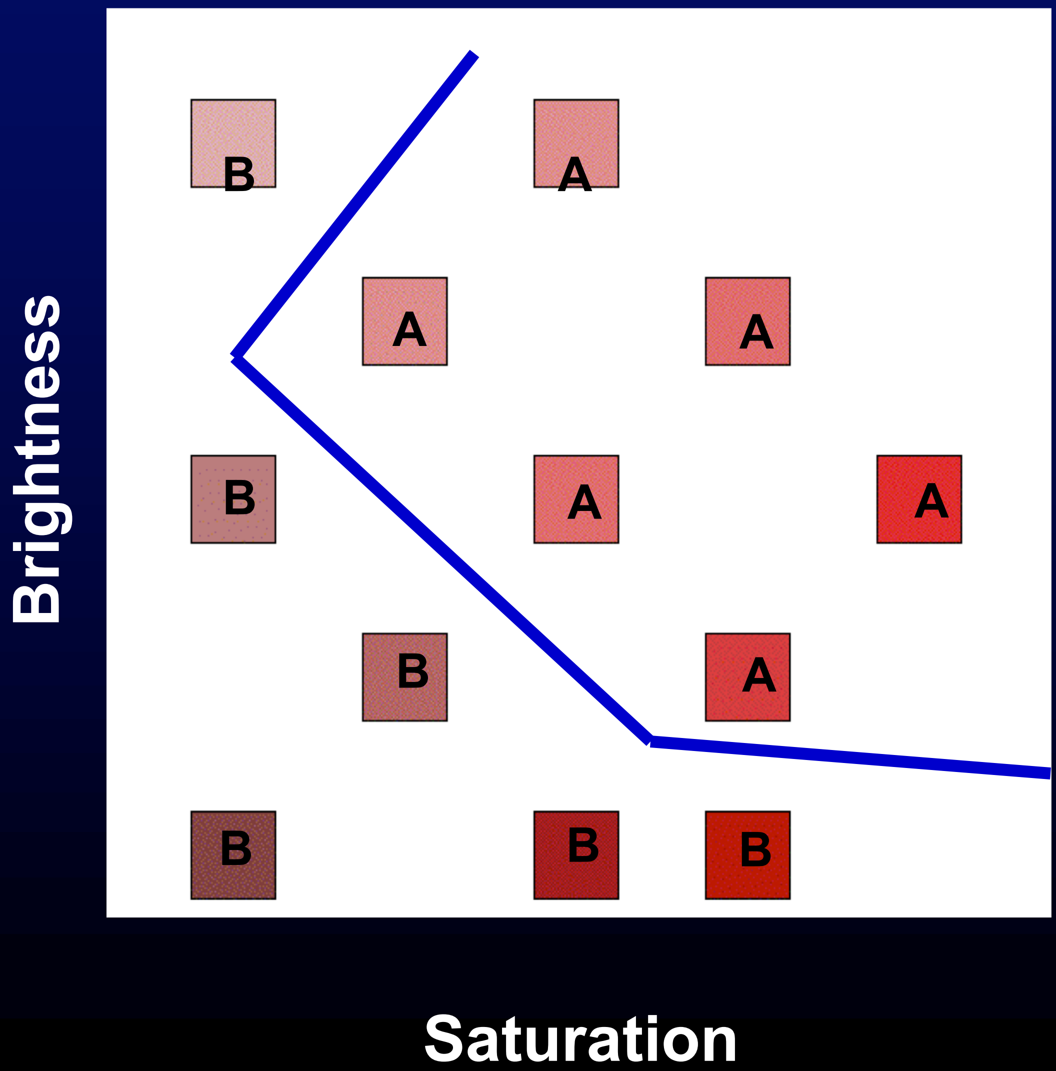
Injected with dopamine (D1) antagonist

SPEED's Performance in the Choi et al. (2005) Experiment

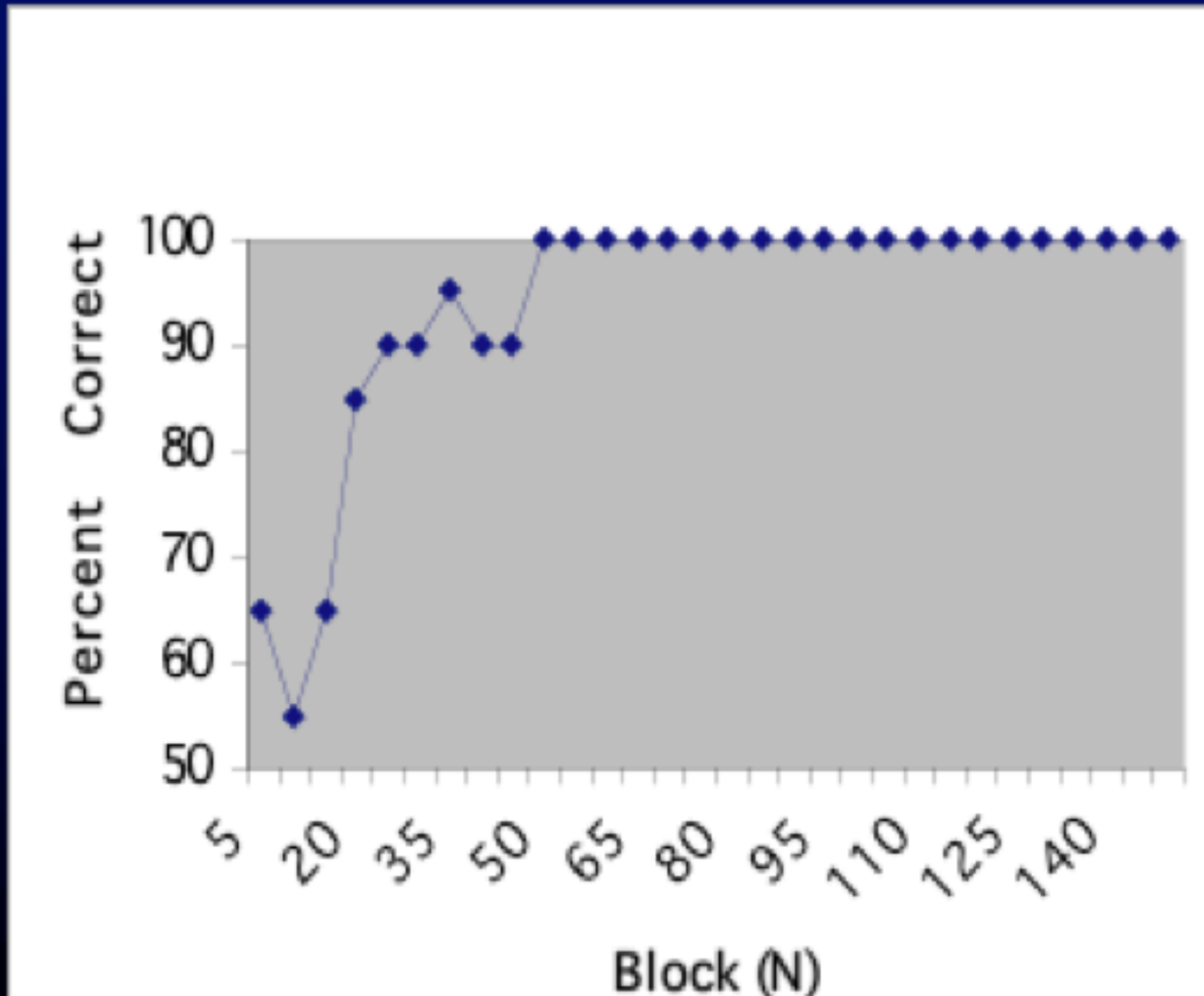


Nosofsky & Palmeri (1997; Psychological Review)

Munsell Color Patches – 3 Subjects – 1800 Trials

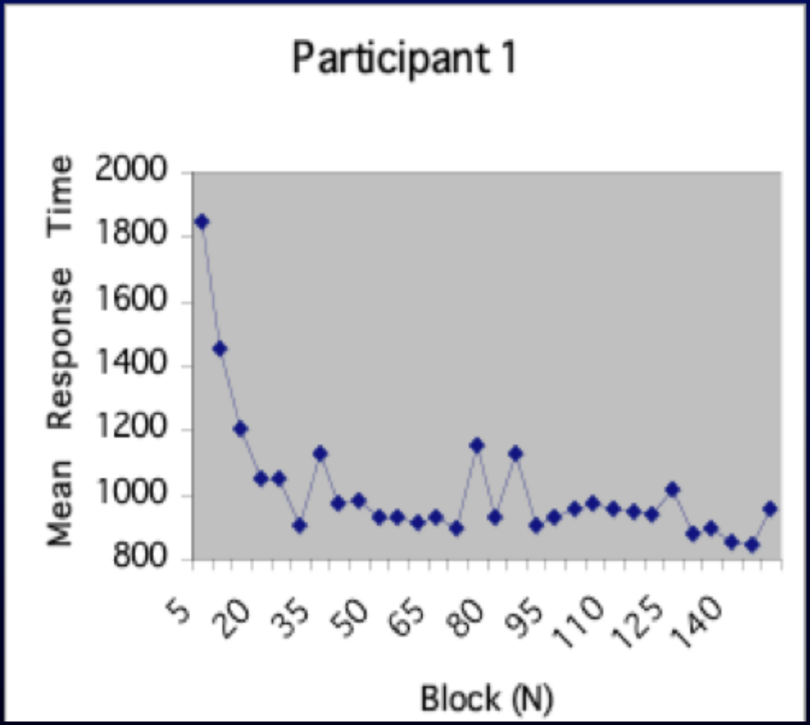


SPEED Accuracy

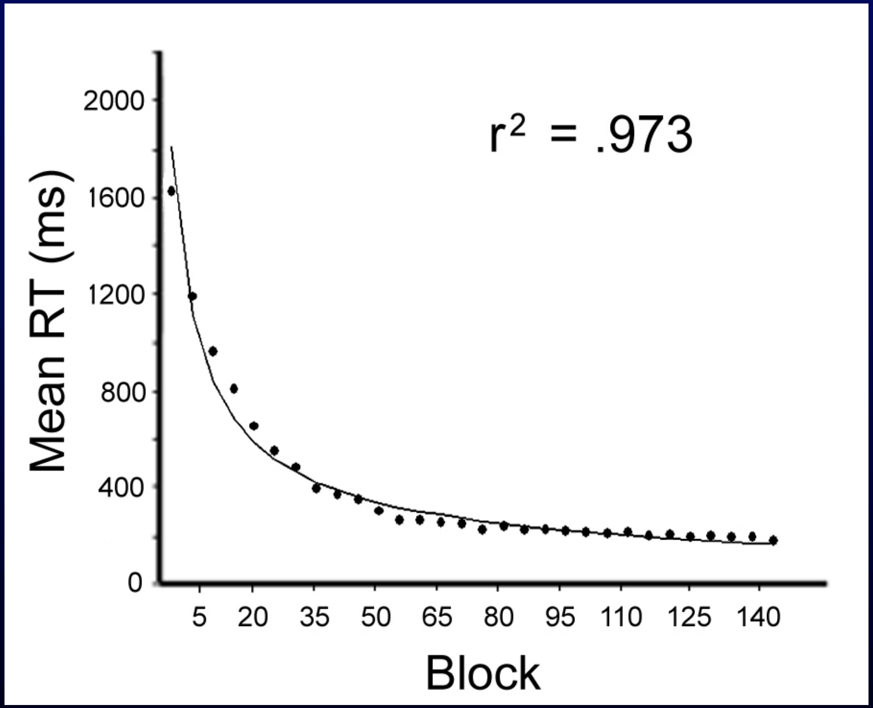


Mean Response Time

Nosofsky & Palmeri (1997)



SPEED



Future Directions

- fMRI
- Model expertise development in:

  -- neuropsychological populations

  (e.g., Parkinson's disease)

  -- subjects under influence of drugs

  (e.g., dopamine antagonists)

- Expertise in other (non procedural-learning) systems

  -- e.g., explicit rule-based categorization systems

Conclusions

- Two paths from stimulus to response
  - -- Cortical-striatal-pallidal-thalamic-cortical
  - -- Cortical-cortical
- Model accounts for many automaticity phenomena
  - -- Single-cell category learning data
  - -- The striatum goes off line with extended practice
  - -- Learning is dopamine dependent early but not late
  - -- Power law speed up with practice

Acknowledgments

Collaborators:

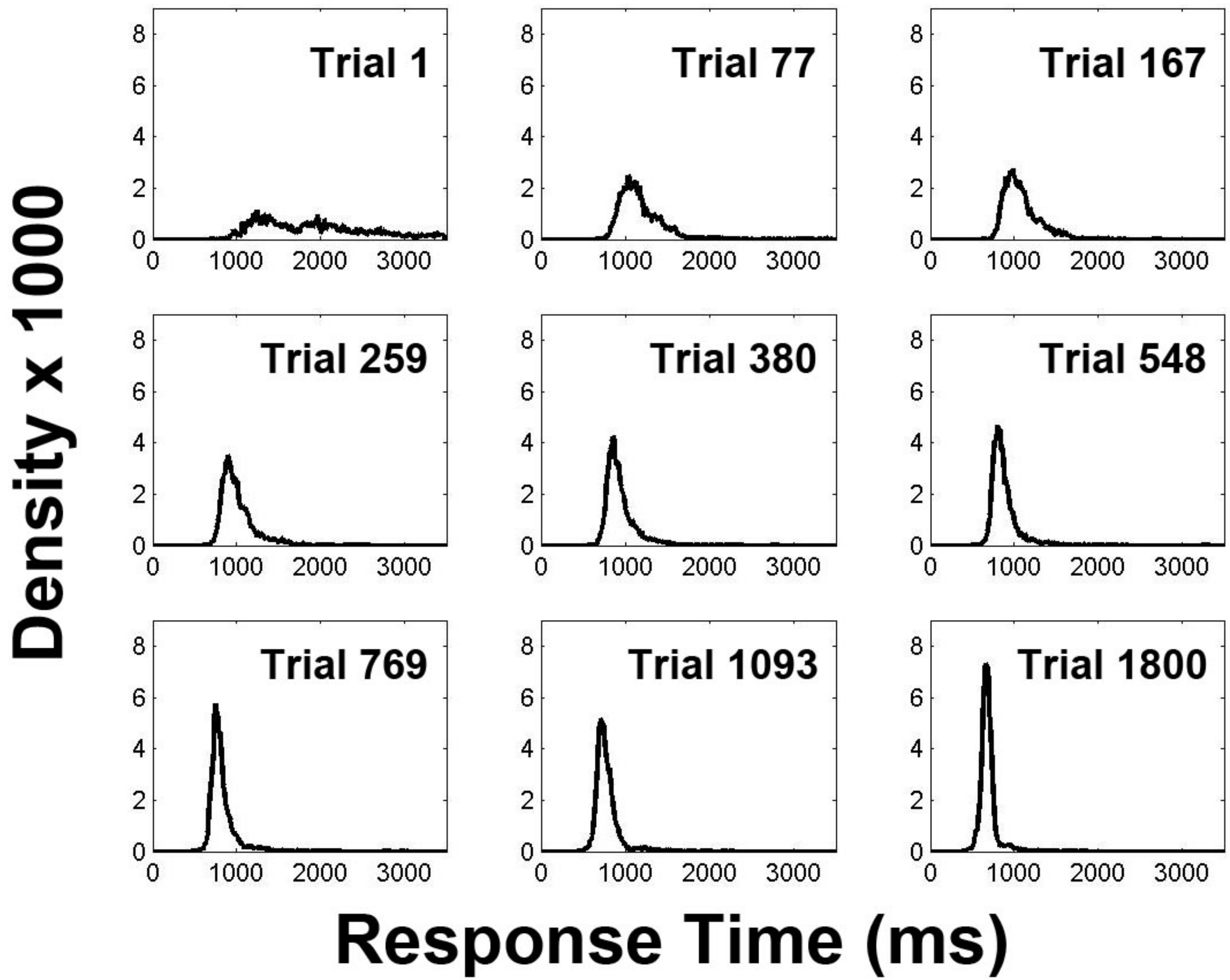
F. Greg Ashby

John Ennis

Funding:

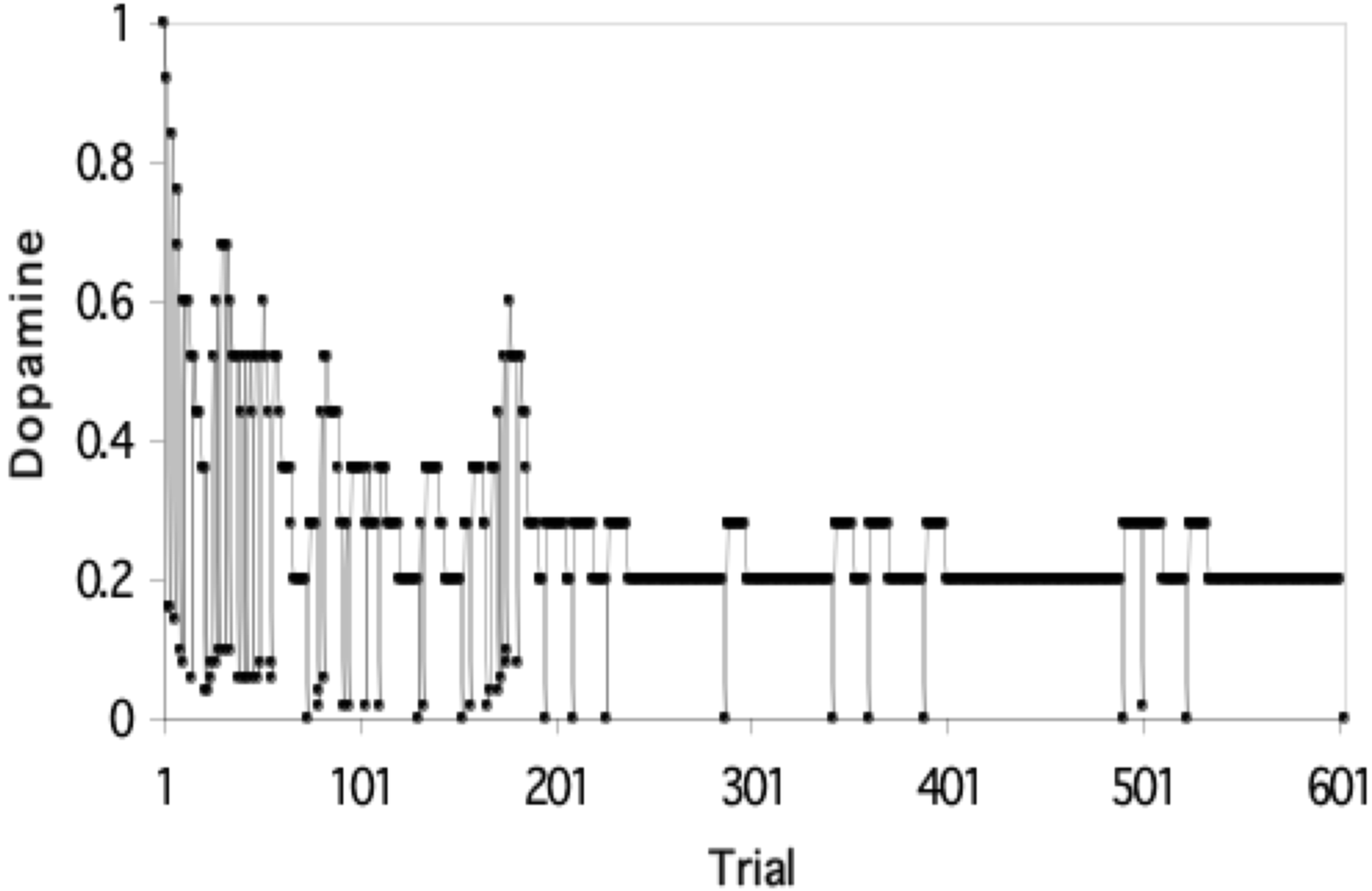
Public Health Services Grant MH3760-2

SPEED RT Density Functions



The Caudate Nucleus

Dopamine Release



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