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

Orientation

Bar Width

A

B
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
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!
Effective learning requires:

- consistent feedback immediately after response
- consistent mapping from category to response location
- no active feedback processing
Is the information-integration task inherently more difficult?
THE TWO CATEGORY LEARNING SYSTEMS OF COVIS


• 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

Low Speed Cell

High Speed Cell

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

The COVIS Explicit System

ACC = Anterior Cingulate
PFC = Lateral Prefrontal Cortex
MDN = Medial Dorsal Nucleus of the Thalamus
GP = Globus Pallidus
CD = Head of the Caudate Nucleus
VTA = Ventral Tegmental Area
SN = Substantia Nigra pars compacta
HC = Hippocampus

- Excitatory projection
- Inhibitory projection
- Dopamine projection

Orientation
Bar Width

PFC
MDN
GP
HC
ACC
Association Cortex
Sensory Cortex
VTA/SN
The COVIS Procedural-Learning System

The Striatal Pattern Classifier (Ashby & Waldron, 1999)
• 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

![Bar chart showing proportion correct for Rule-Based and Information-Integration methods with Delay and Imm conditions.](image)

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

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
"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)
“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)

## A Double Dissociation?

<table>
<thead>
<tr>
<th>Category Learning</th>
<th>Categorization Expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Patients with Basal Ganglia Dysfunction</strong> <em>(Parkinson’s disease, Huntington’s disease)</em></td>
<td>Impaired</td>
</tr>
<tr>
<td><strong>Patients with certain visual cortex lesions</strong> <em>(category-specific agnosia)</em></td>
<td>Unimpaired if stimuli are perceived normally?</td>
</tr>
</tbody>
</table>
BUILDING A MODEL OF AUTOMATICITY

Motor Cortex

Visual Cortex
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

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

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 $\varepsilon(t)$ is white noise.
Modeling Activation in Other Units

Globus Pallidus

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

Thalamus:

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

Premotor Area

\[
\frac{dE_J(t)}{dt} = \left[ \alpha_E T_J(t) + \sum_K v_{K,J}(n) I_K(t) \right] \left[ 1 - E_J(t) \right] - \beta_E E_K(t) - \gamma_E [E_J(t) - E_{base}] + \sigma_E \epsilon(t) E_J(t) \left[ 1 - E_J(t) \right],
\]
Excitatory projection (glutamate)
Inhibitory projection (GABA)
Dopamine projection

SPEED

Hebbian learning

3-factor learning

Response
A - B Difference

Premotor Area (Cortex)

Sensory Association Cortex

Stimulus

Thalamus{$VAM_l$}

Globus Pallidus

Striatum

$SN_{bc}$
\[ v_{K,J}(n+1) = v_{K,J}(n) + \alpha_v I_k(t) \left[ E_J(t) - \theta_{\text{NMDA}} \right] \begin{cases} 1 & \text{if} \, v_{K,J}(n) > \theta_{\text{NMDA}} \\ 0 & \text{otherwise} \end{cases} \]

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

**LTP**

- Presynaptic activation
- Postsynaptic activation (above NMDA threshold)

**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

dopamine below baseline (error)

activation above NMDA threshold

activation above NMDA threshold
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

![Graph showing dopamine release over trials](image-url)
SPEED - After Over-Training

A Stimulus

Sensory Association Cortex

\( \Delta_{A,B}(t) \)

\( E_A(t) \)

\( T_A(t) \)

\( G_A(t) \)

Premotor Area (Cortex)

Thalamus\(_{VA/VL}\)

Globus Pallidus

\( I_K(t) \)

Striatum

SN\(_{pc}\)

D(n)
Experimental Tests
Tactile Category Learning

Romo, Merchant et al.
Model Fits

Proportion of Low Responses

- Monkeys
- SPEED

Speed (mm/s)
SPEED’s Single Cell Responses – Putamen

Low Speed Cell

High Speed Cell

S ON-OFF

12
14
16
18
20
22
24
26
28
30

n / s

200 ms / div

mm / s
SPEED’s Responses – Premotor Cortex

Low Speed Cells    High Speed Cells

Romo et al., 1997
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

Block (N) vs Percent Correct

- Block (N): 5, 20, 35, 50, 65, 80, 95, 110, 125, 140
- Percent Correct: 50, 60, 70, 80, 90, 95, 100

The graph shows the trend of percent correct over different block sizes, indicating an improvement in accuracy as the block size increases.
Mean Response Time

Nosofsky & Palmeri (1997)

Block (N)

Mean Response Time

Participant 1

Mean RT (ms)

Block

SPEED

$r^2 = .973$
SPEED RT Density Functions

Trial 1

Trial 77

Trial 167

Trial 259

Trial 380

Trial 548

Trial 769

Trial 1093

Trial 1800

Density x 1000

Response Time (ms)
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