

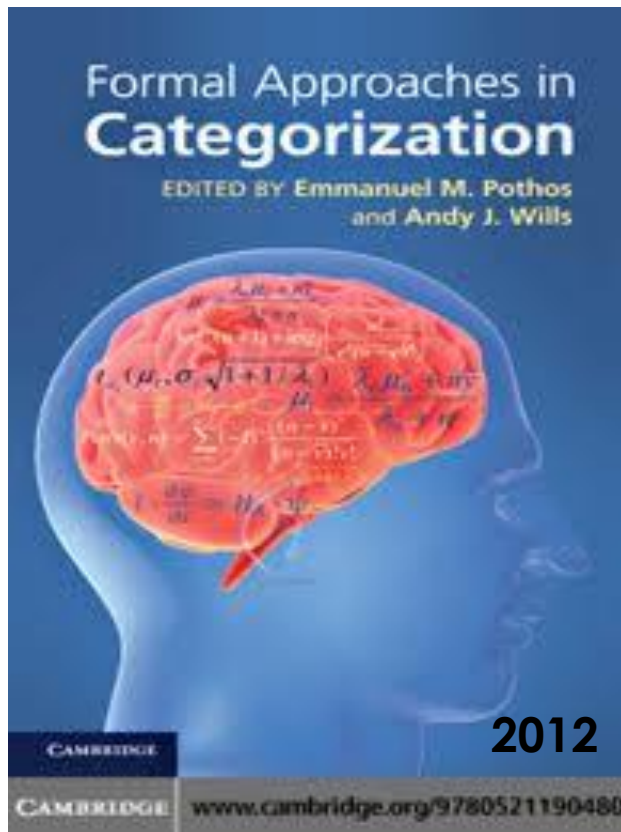
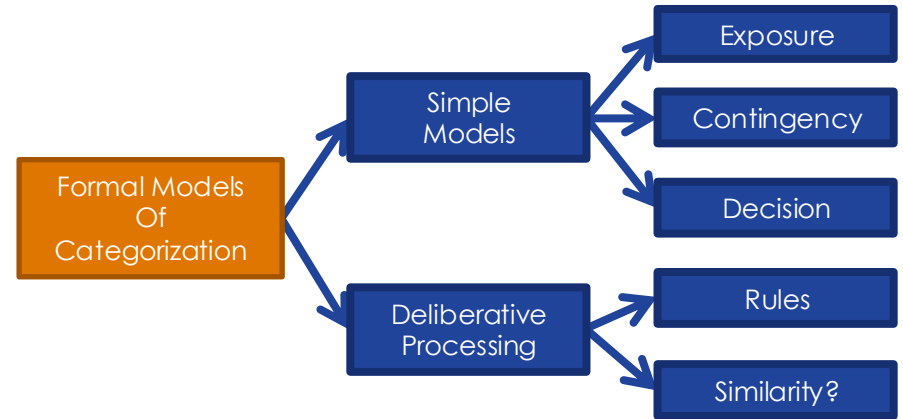
# Categorization

Andy Wills  
University of Exeter



Emmanuel Pothos

# Formal models



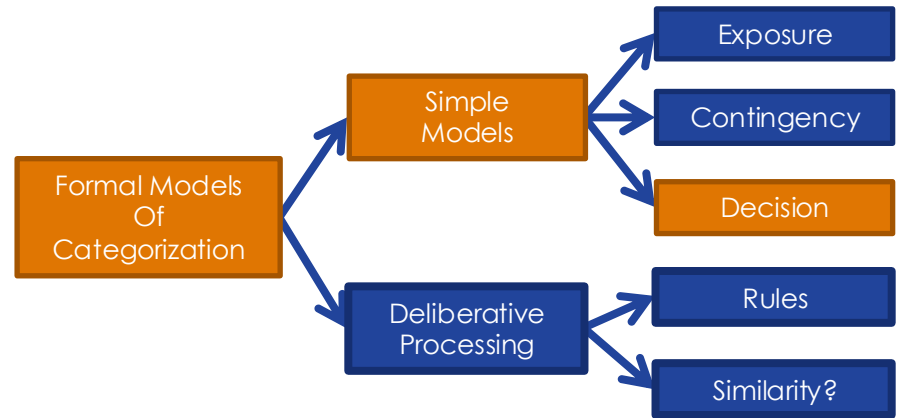
- Pothos & Wills (2011)
  - Contributions from key players: Ashby, McClelland, Kruschke, Nosofsky, Murphy, Medin..
- Wills & Pothos (2012)
  - Critique of progress in the field of formal modelling.
  - Roadmap for future progress in this area.

Wills, A.J. and Pothos, E.M. (2012).  
*Psychological Bulletin*, 138, 102-125.

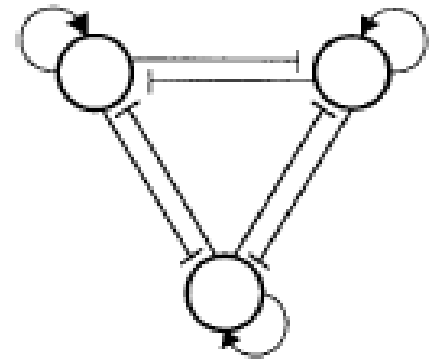


# Decision

- Most formal models of categorization use the Luce choice rule (aka ratio rule; softmax) to turn evidence terms into decisions.
- We showed, contrary to some previous beliefs, that such a system was not equivalent to a lateral inhibition decision systems.
- We also demonstrated that lateral inhibition provided a better account where  $N > 2$ .

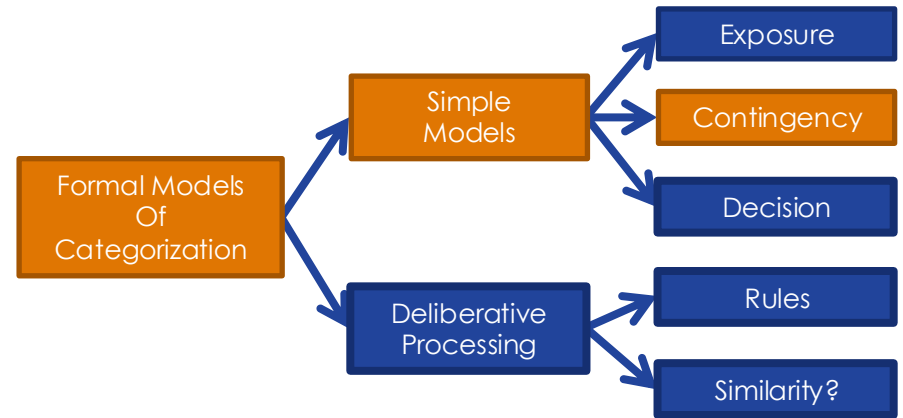


$$P(i) = \frac{v_i}{\sum_{j=1}^n v_j}$$





# Contingency



## □ Forward cue-competition

A+    AX+    X?  
B-    BY+    Y?

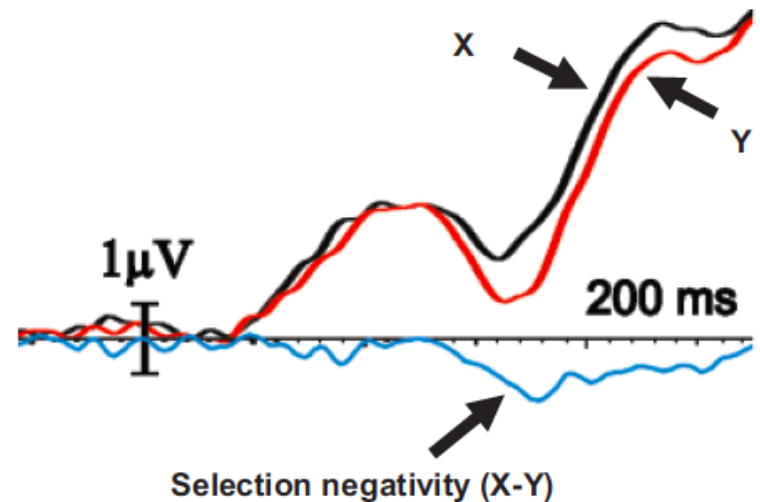
## □ One explanation – Prediction error drives selective attention.

## □ Supported this explanation through ERP and eye-tracking methods.

## □ Showed fast attentional differentiation.

- Not a consequence of deliberative processing, at least within-trial.

- Top-down influences?



Wills, A.J., Lavric, A, Croft, G. and Hodgson, T.L (2007).  
*Journal of Cognitive Neuroscience*. 19, 843-854.

# Media coverage

>> **LEARNING**

## Mom Was Right

So you goofed ...

**We learn more** from our mistakes than from our successes, the old cliché says—and now scientists know why. Researchers at the University of Exeter in England discovered a brain mechanism that alerts us to situations in which we previously went wrong.

In the study, students playing physicians had to diagnose a fictitious disease based on images from equally fictitious blood samples. When participants saw images that had

previously led them to an error warning signals in the brain a tenth of a second later—much did signals triggered by images led in a correct diagnosis. Experiments confirmed that slipups do indeed lead to learning, but this one is the brain's specific reaction to a

This early-warning signal is in situations ranging from the mundane. A child who touches a glowing burner in the brain will alert her to avoid the she made the last time.

**ALARM BELL:** An early warning system in the brain that helps avoid repeating mistakes has been identified by scientists. The alarm mechanism is triggered in just a tenth of a second and responds before there is any conscious consideration of the blunder. It will remind a footballer of a formidable opponent or warn a woman to avoid a certain type of

Metro

## We do learn by mistakes

**HUMANS** really do learn from their mistakes – in just 0.1 seconds, scientists claimed yesterday.

New research has identified a mechanism in the brain's lower region that helps us avoid past errors.

Using electrodes attached to volunteers'

Daily Mirror

## Scientific American Mind

Click to **LOOK INSIDE!**

LEWIS SMITH

## Why the Lion Grew Its Mane

A MISCELLANY OF RECENT SCIENTIFIC DISCOVERIES FROM ASTRONOMY TO ZOOLOGY



From making sense of the world's most mysterious places to exploring the future of space exploration, this collection of 120 articles is the ultimate guide to the latest in science and technology.

PAPADAKIS

**BBC**  
RADIO



## Brain signal stops error repetition

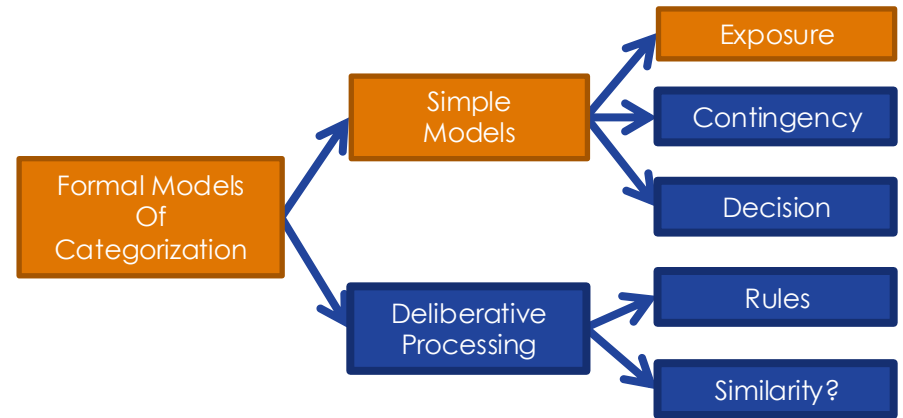
Scientists have discovered why we learn from our mistakes

Psychologists from the  
Financial Times



Alice Welham

# Exposure



- Unitization – Creation of new feature representations by fusion.
- Rob Goldstone, Phillippe Schyns, Richard Shiffrin – Unitization is function-dependent.
- We argued (on basis of the predictions of a simple connectionist model), that it is merely exposure dependent.
- Found evidence for our position from similarity judgments and eye-gaze measures

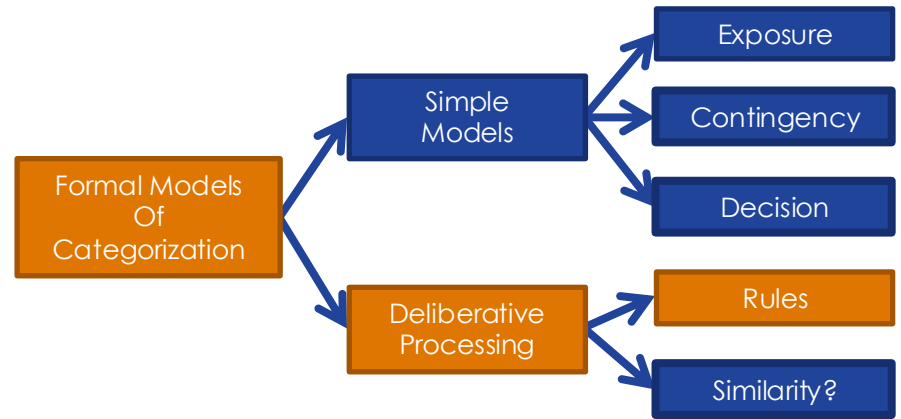


Welham, A.K. and Wills, A.J. (2011). *Memory and Cognition*. 39, 1518-1533



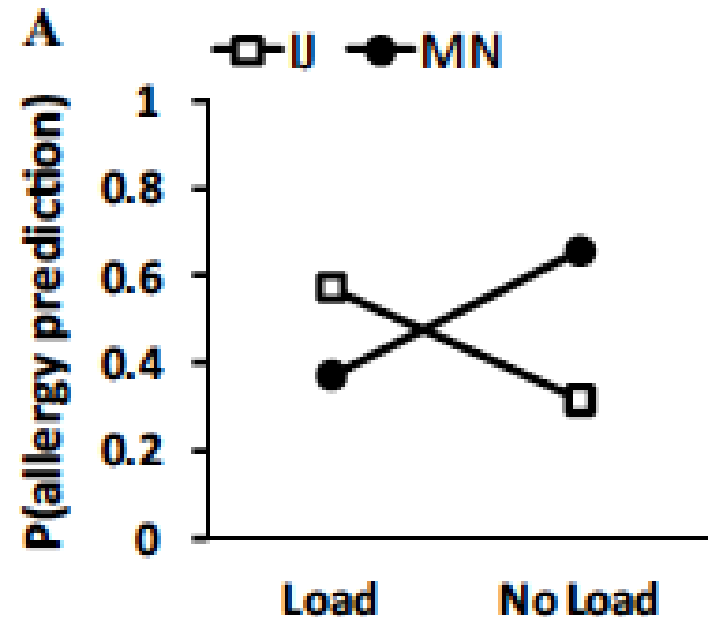
Steve Graham

# Rules



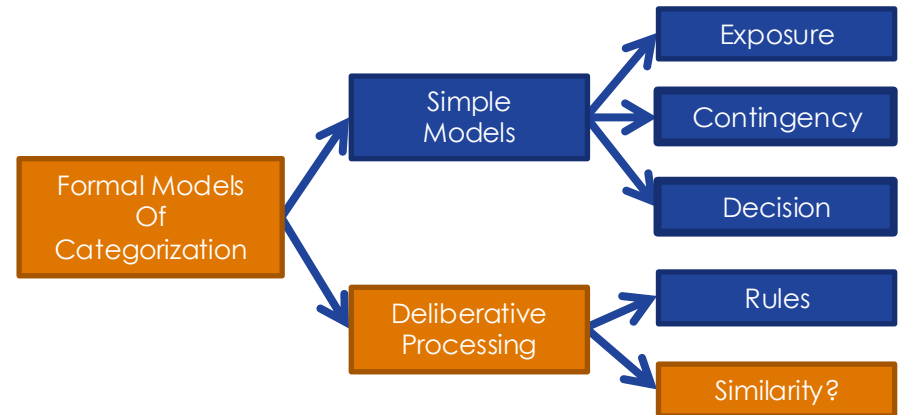
- Two ways of generalizing from known instances:
  - Overall similarity
  - Extraction and application of an underlying rule

A+	B+	AB-
C-	D-	CD+
I+	J+	
M-	N-	



Wills, A.J. et al. (2011). *Journal of Experimental Psychology: Animal Behavior Processes*, 37, 241-245.

# Overall similarity



- Is classification on the basis of overall similarity more or less requiring of cognitive resources than classification on the basis of a single dimension?



**Fraser  
Milton**



**Chris  
Longmore**



**Stephen  
Lea**



**Lisa  
Leaver**

**ADULT HUMANS**

**PIGEONS**

**SQUIRRELS**

# Brooks (1978) – Nonanalytic categorization

- Analytic categorization –Typically results in a subset of the stimulus properties having control over responding.
- Non-analytic categorization – Classification on the basis of overall similarity to known examples.
- Nonanalytic categorization more likely to occur when cognitive resources are limited.
- A striking *less is more* hypothesis
- A hypothesis also popular in reasoning (Sloman, 1996), and complex decision-making (Dijksterhuis et al., 2006)

# *Less is more, still.*

*“evidence suggests that in many situations, it is easier for people to base similarity and categorization judgments on more, rather than fewer, properties”*

Goldstone and Barsalou (1998, p. 239)

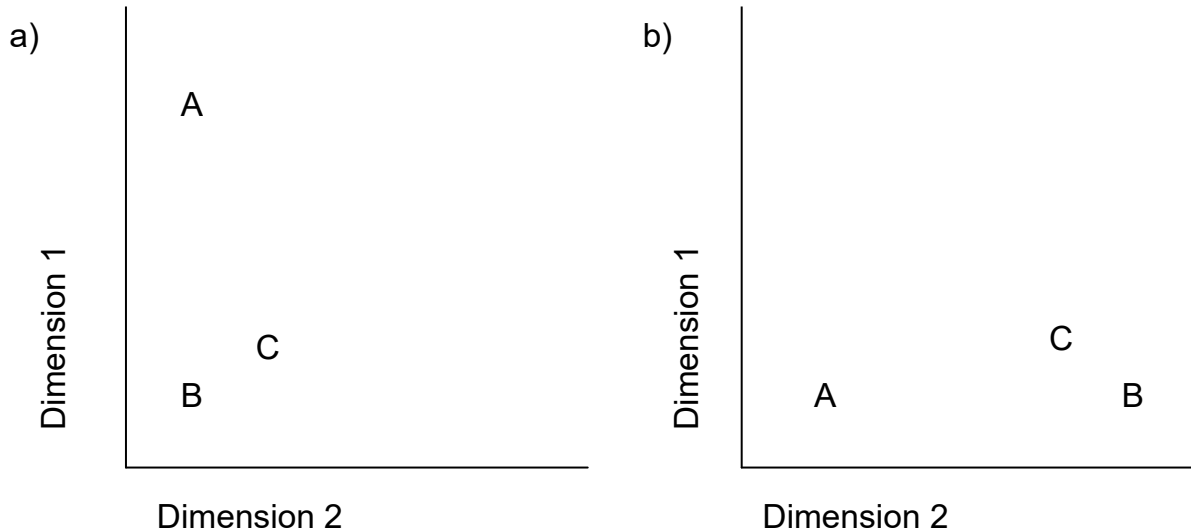
*“family resemblance categorization is the default mode of approach to tasks when full executive functioning is underdeveloped, inhibited, or lacking for any reason”*

Couchman et al. (2010, p.180).

# *Less is more* – Evidence?

- Triad procedure
  - Ward (1983), Smith & Kemler Nelson (1984)
  
- Ashby-Maddox procedure
  - Zeithamova & Maddox (2006)
  
- Criterial attribute procedure
  - Kemler Nelson (1984)
  - Smith & Shapiro (1989)

# Triad procedure

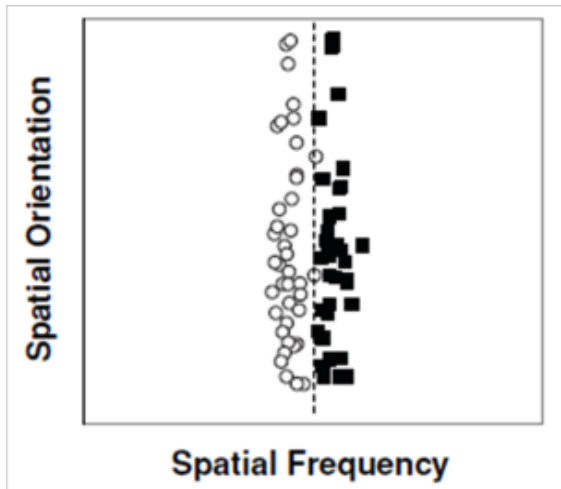


TRIAD  
EXAMPLE

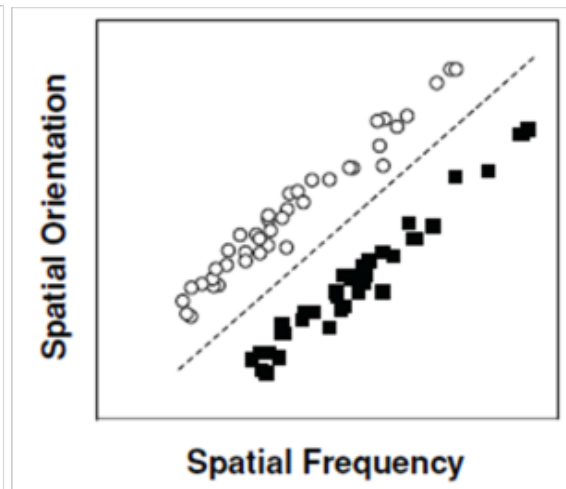
- Which two stimuli go together best?
  - Three possible answers: AB (*dimensional*), BC (*overall similarity*), AC (*haphazard*)
  - BC responding increases with time pressure / concurrent load.
- But...
  - Consistent AB and consistent BC responding requires use of both dimensions on all trials.
  - If time pressure / concurrent load increases *single-dimension* responding, then BC responding could increase (assuming AB responding dominates unspeeded / full attention)

# Ashby-Maddox procedure

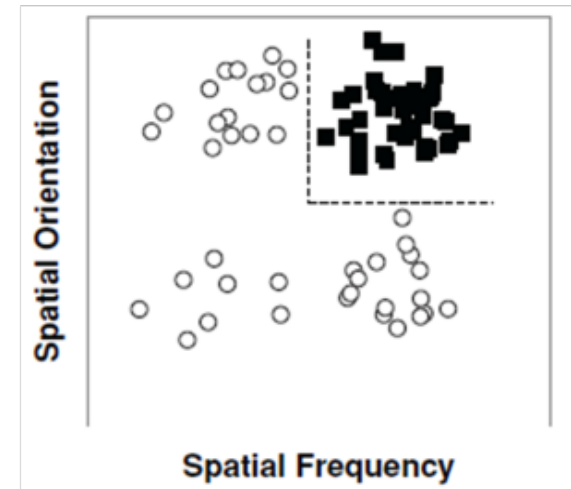
A. Rule-based  
(Single dimension)



B. Information  
Integration



C. Rule-based  
(Conjunction)



- Zeithamova and Maddox (2006) – Concurrent load affects single-dimension classification more than overall similarity (information integration) classification.
- But, concurrent load also affects conjunction categorization more than information integration.
- If reliable (see Newell et al, 2010; Newell, X, Wills & Milton, submitted), this procedure is probably more about verbalizability than single versus multiple dimensions.

# Criterial-attribute procedure

Category A

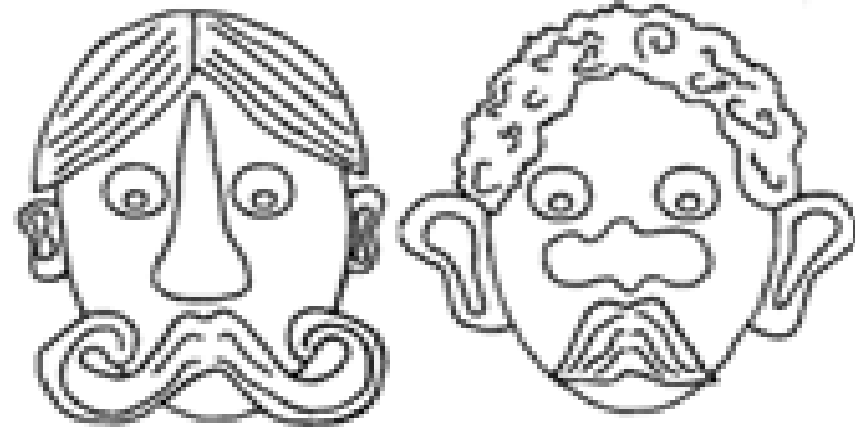
0 0 0 0  
0 1 0 0  
0 0 1 0  
0 0 0 1

Category B

1 1 1 1  
1 0 1 1  
1 1 0 1  
1 1 1 0

Test

0 1 1 1  
1 0 0 0



- Categorization can be learned by either detecting the criterial attribute, or calculating similarity over at least three of the dimensions.
- Test phase puts these two strategies in opposition.
- Concurrent load during training increases overall similarity classification in the (full attention) test – Smith & Shapiro (1989)

# Criterial-attribute procedure

Category A

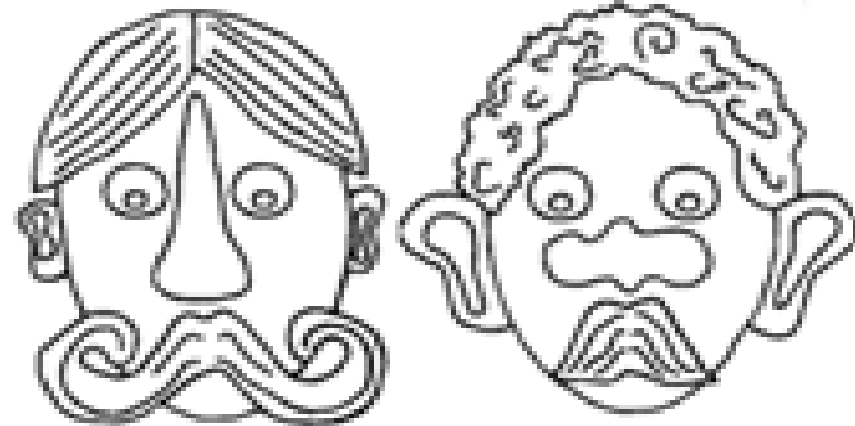
0 0 0 0  
0 1 0 0  
0 0 1 0  
0 0 0 1

Category B

1 1 1 1  
1 0 1 1  
1 1 0 1  
1 1 1 0

Test

0 1 1 1  
1 0 0 0



- A sufficient explanation
  - Searching for the one-dimensional hypothesis that meets the training criterion is in itself effortful, and is disrupted by concurrent load.
  - Any one of the non-criterial dimensions can support 75% accuracy - above chance, but below the learning criterion of the task.
  - This may make it particularly challenging to select the single one-dimensional hypothesis that permits 100% accuracy.
  - Having been unable to identify the criterial attribute, participants may resort to alternative approaches such as classifying by overall similarity, despite the additional effort this requires to implement.

# Summary of evidence

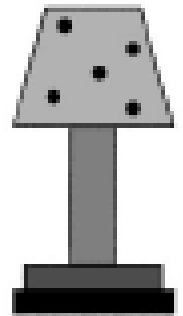
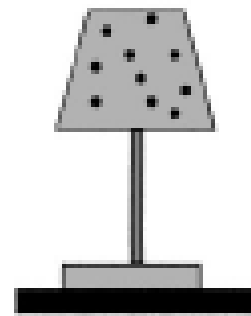
- Triad procedure
  - Interpretatively ambiguous
- Ashby-Maddox procedure
  - For current purposes, it confounds verbalizability with dimensionality.
- Criterial-attribute procedure
  - For current purposes, confounds effort in discovering a single-dimension rule that meets the training criterion with the effort of implementing a single-dimension versus an overall similarity classification.

# Making progress

- We need a procedure
  - Where both single-dimension and overall similarity strategies are verbalizable
  - Where effort of discovering an adequate strategy is not confounded with the effort of implementing it.
  - The latter implies an unsupervised procedure (rather than learning from feedback)
  - but the best-known – the triad procedure – is interpretatively ambiguous.
- The match-to-standards procedure provides a better alternative.

# Match-to-standards

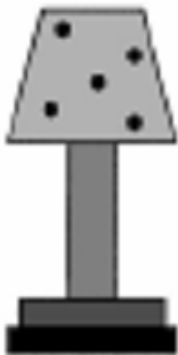
- Regehr & Brooks (1995); Milton & Wills (2004).
- Classification (sorting) without feedback.
- Classification of a set of 10 items (the 2 standards, plus the 8 one-away stimuli).
- 90+% of participants produce one of two kinds of sort, can describe their sort, and there's a high correspondence between the description and the sort.



# OS and UD sorting



0 0 0 0



Overall  
similarity

1 1 1 1  
0 1 1 1  
1 0 1 1  
1 1 0 1  
1 1 1 0

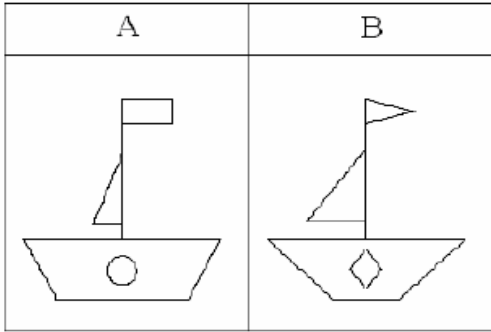
Unidimensional

1 1 1 1  
1 0 0 0  
1 0 1 1  
1 1 0 1  
1 1 1 0

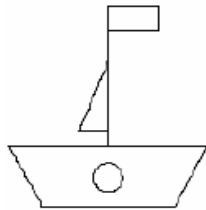
0 0 0 0  
1 0 0 0  
0 1 0 0  
0 0 1 0  
0 0 0 1

0 0 0 0  
0 1 1 1  
0 1 0 0  
0 0 1 0  
0 0 0 1

# Time pressure in MTS



Response-terminated



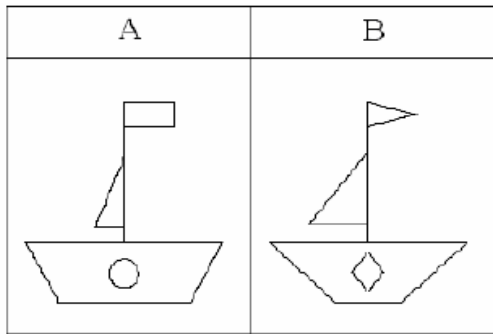
1024 or 4096 ms



Response-terminated

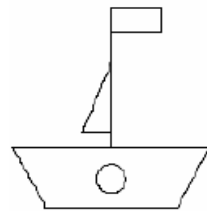
■ 12 blocks

# Concurrent load in MTS



Response-terminated

“3...5...2...8...6...7” 2480ms



1500ms



Response-terminated

■ 12 blocks

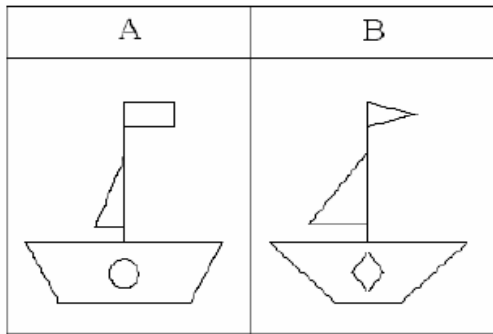
5

Response-terminated

# Possible reasons for difference?

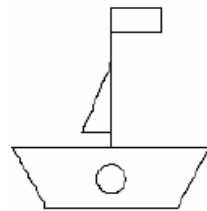
- Opposite results to the triad task.
- If this is due to unavoidable differences in procedure, then we'd favour conclusions based on MTS, for reasons give.
- But there are a number of areas where the procedures could be made more similar:
  - Short-term memory load versus operational load
  - Combination of time pressure and concurrent load
  - Memory versus perceptual representations

# Using operational load



Response-terminated

“3...5...2...8...6...7” 2480ms



1500ms



Response-terminated

■ 12 blocks

5

Response-terminated

# Using operational load

# Combining load and time

Using a more perceptual variant

# Expanding the evidence base

- Working memory capacity
- Instructions
- fMRI analysis
- Extreme time pressure
- [Comparative psychology – pigeons / squirrels]

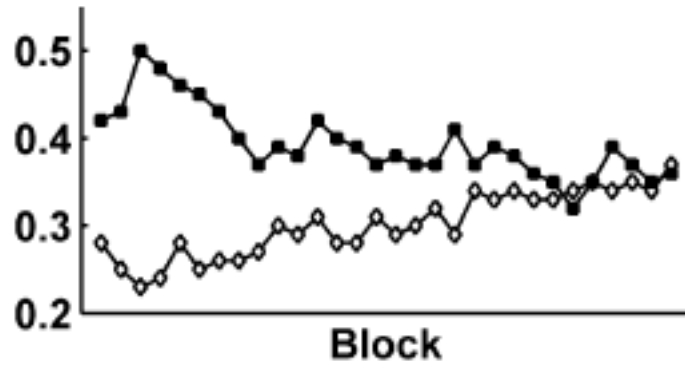
# Working memory capacity

# Instructions

# fMRI analysis

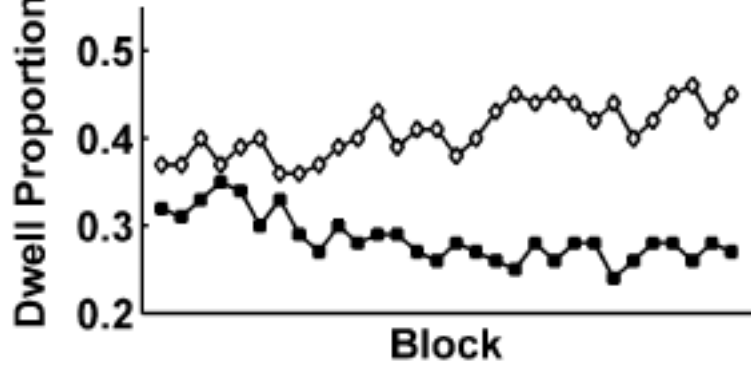
Configurations  
ations

**a**



- ◇ Non-obvious configurations
- Obvious configurations

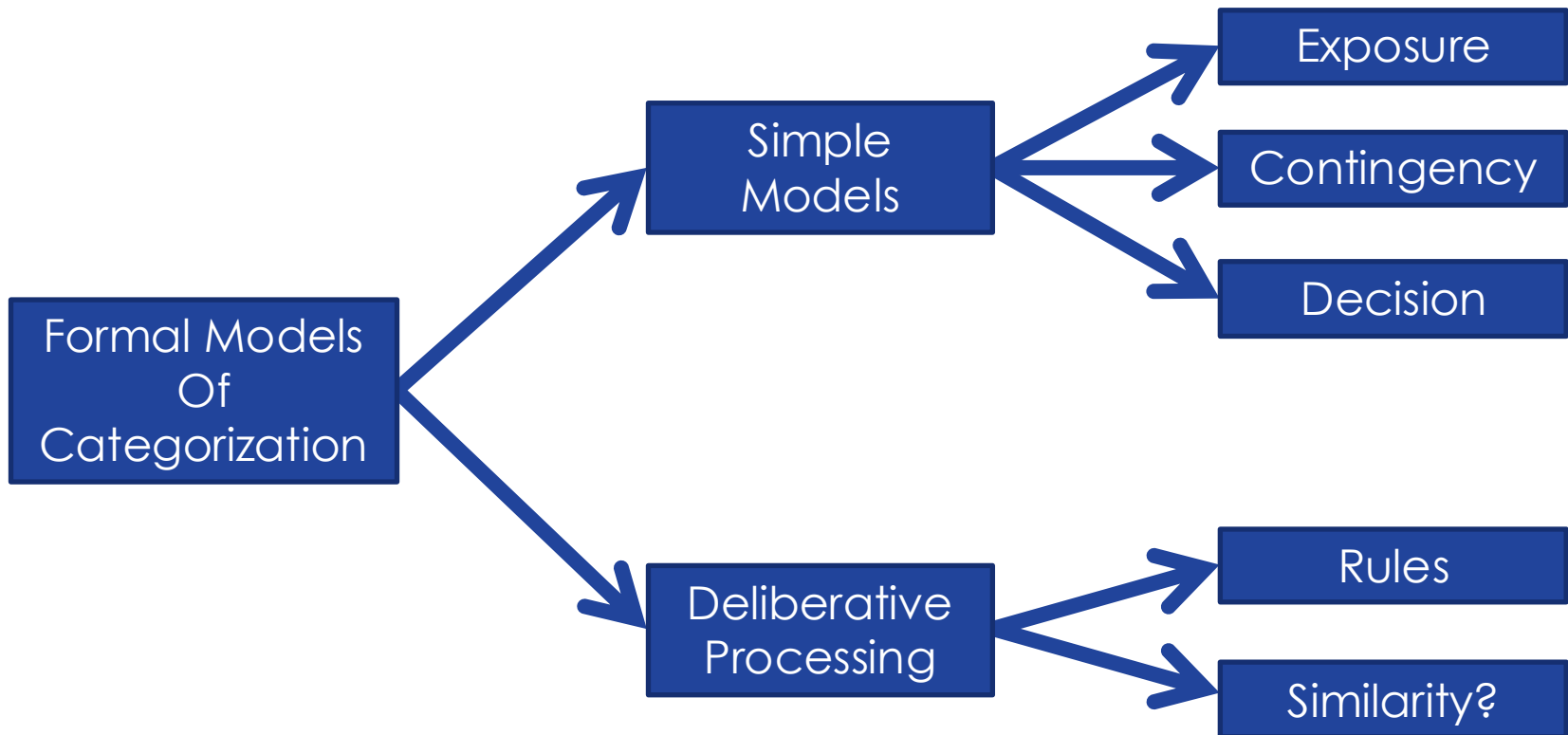
**b**



- ◇ Categorization
- Exposure

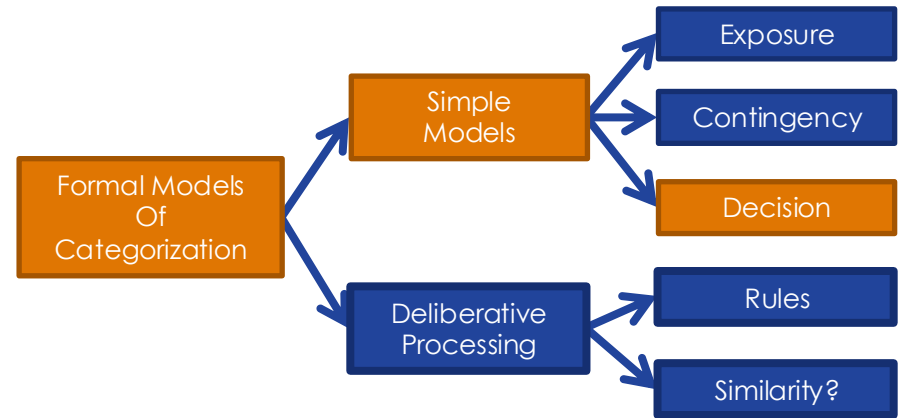


# Research hierarchy





# Decision



## ■ Semantic dementia patient JB

- Over 18 months went from preserved colour naming / recognition to floor.
- Free classification of colours remained good, did not diverge from controls, and varied no more than controls.
- Central to these analyses were formal models and analyses of classification that I developed (inc. arbitrary coherent categories model).



# Extreme time pressure