

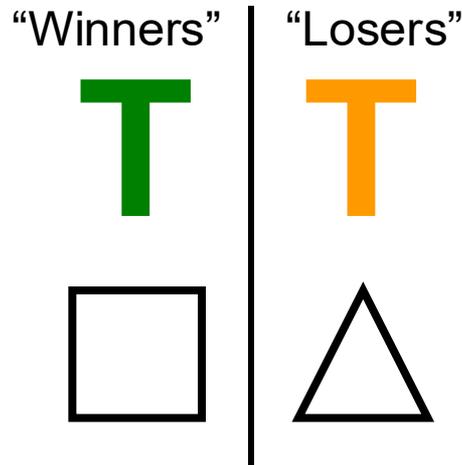


# Formation of category representations

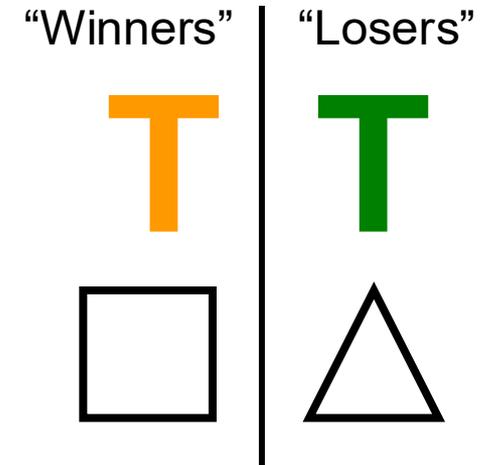
Andy Wills  
Exeter University

Malia Noury, Nick Moberly, Matt Newport

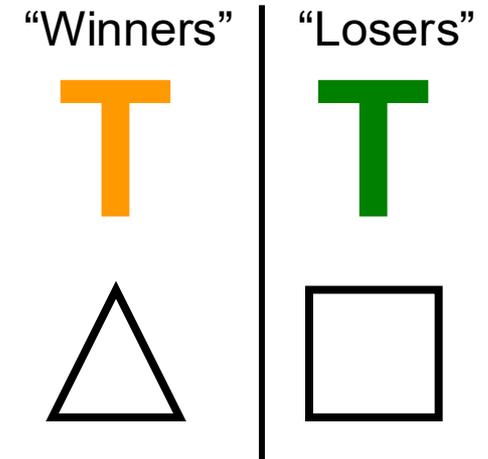
# Partial vs. full reversal



Partial reversal

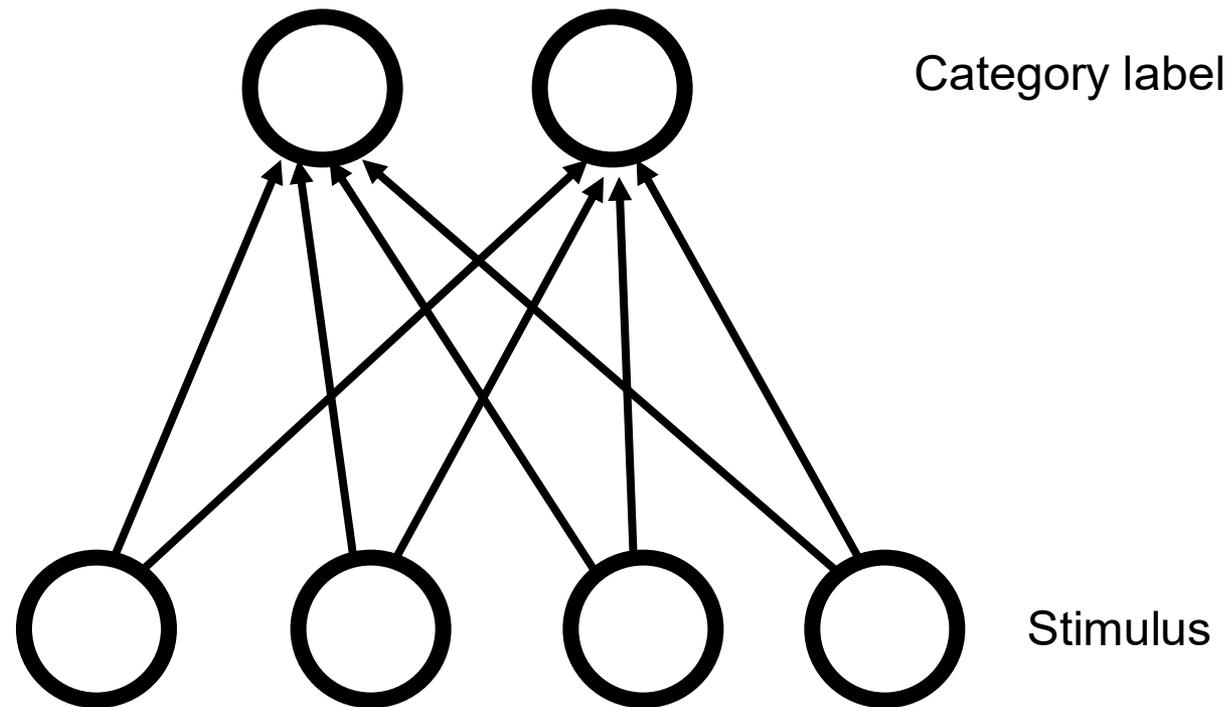


Full reversal

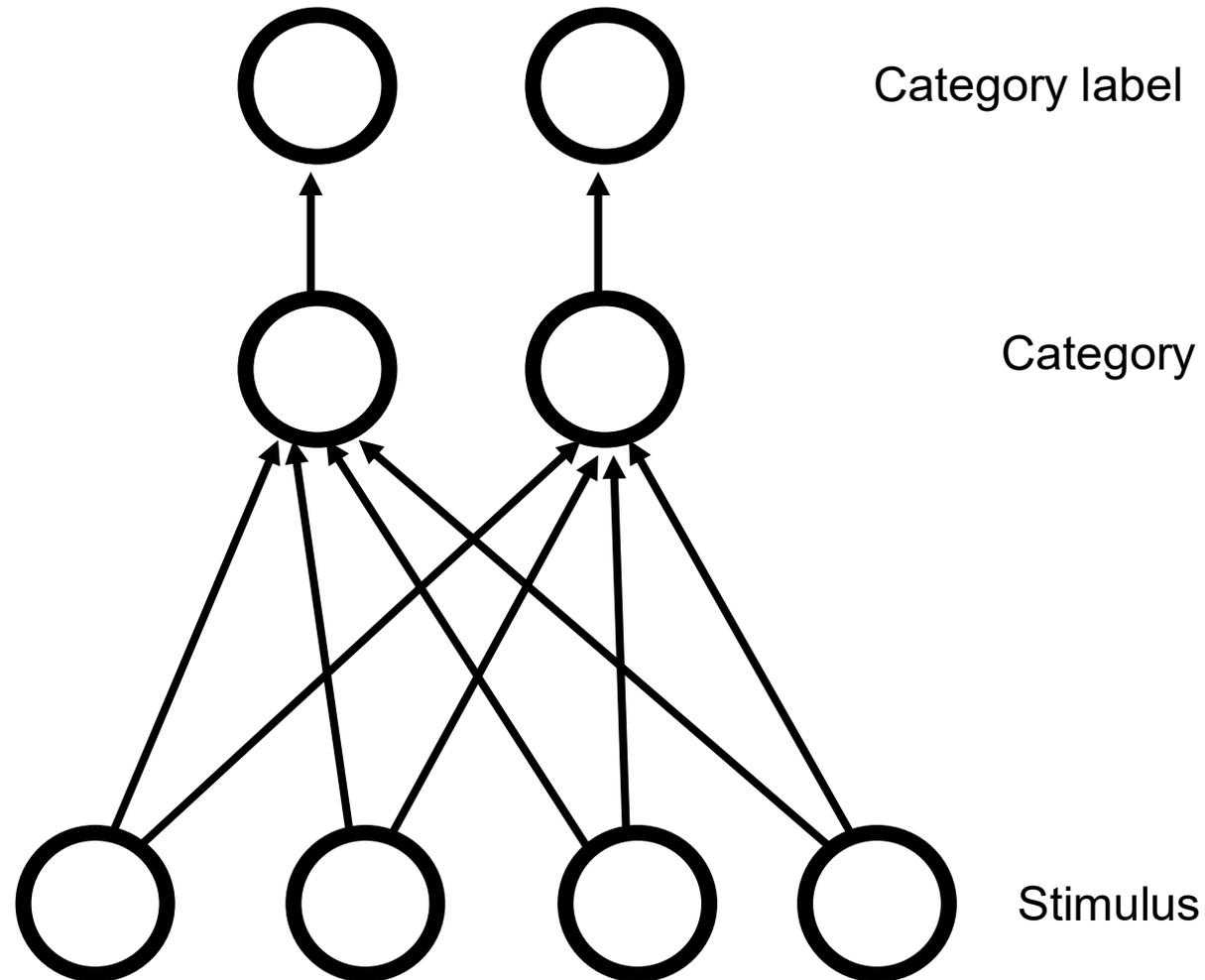


Bogartz (1965)  
**Sanders (1971)**  
Delamater & Joseph (2000)  
Honey & Ward-Robinson (2001)

# Formal models of categorization in humans



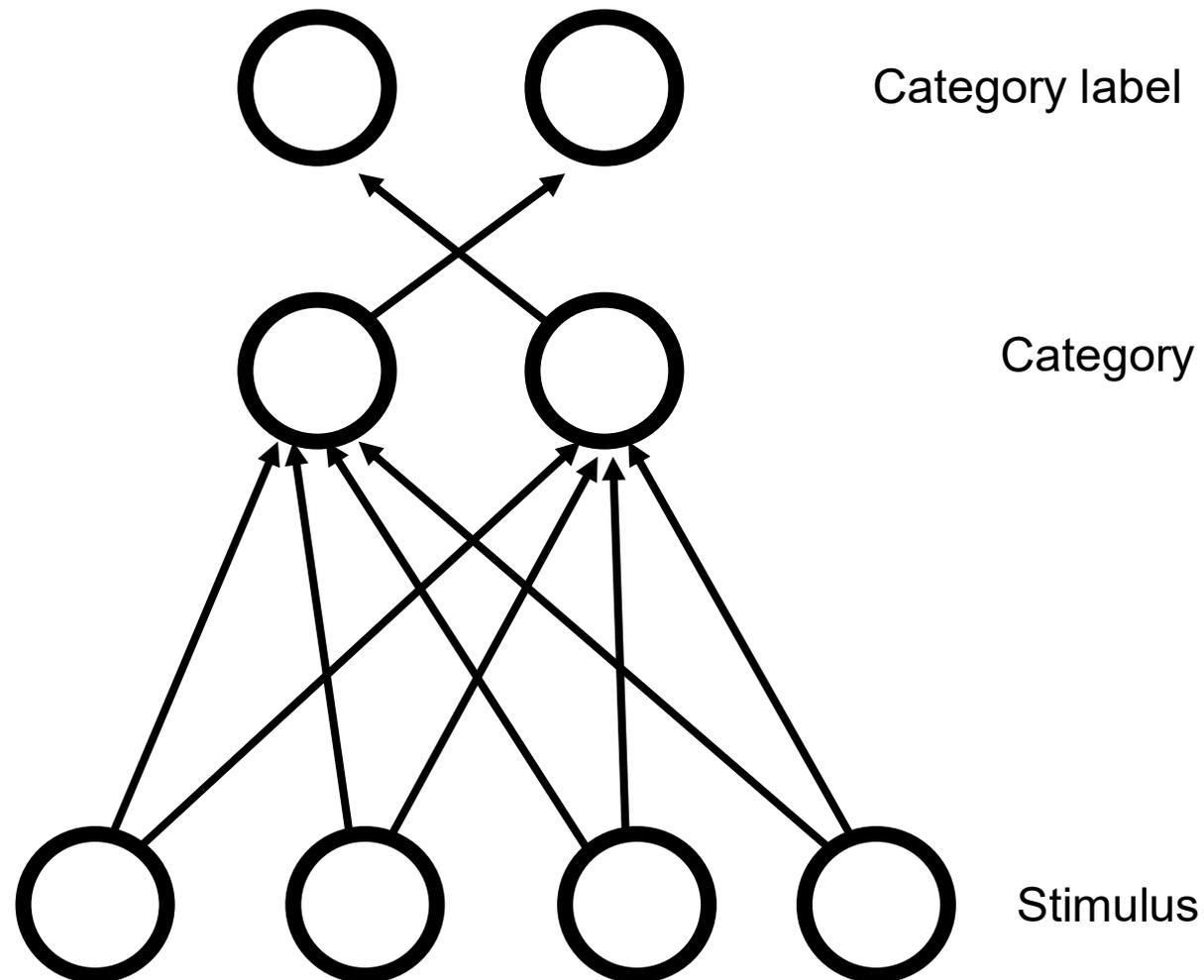
# Category representations



Kruschke (1996)

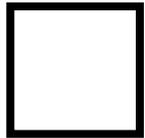
Honey (2000)

# Category representations



# Non-attentional optional shift

“Winners”



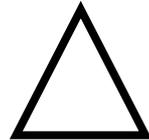
“Losers”



“Winners”



“Losers”



## Category A

!	+	↑	△	\$
?	+	↑	△	\$
!	X	↑	△	\$
!	+	↓	△	\$
!	+	↑	□	\$
!	+	↑	△	£
!	+	↑	□	£
!	+	↓	△	£
!	X	↑	△	£
?	+	↑	△	£
!	+	↓	□	\$
!	X	↑	□	\$
?	+	↑	□	\$
!	X	↓	△	\$
?	+	↓	△	\$
?	X	↑	△	\$

## Category B

?	X	↓	□	£
!	X	↓	□	£
?	+	↓	□	£
?	X	↑	□	£
?	X	↓	△	£
?	X	↓	□	\$
?	X	↑	□	\$
?	+	↓	□	\$
!	X	↓	□	\$
?	X	↑	△	£
?	+	↓	△	£
!	X	↓	△	£
?	+	↑	□	£
!	X	↑	□	£
!	+	↓	□	£

## Polymorphous concepts

No feature is necessary or sufficient

Coherent categories  
(maximizes within-category similarity and minimizes between-category similarity)

Feature-pairs stochastically independent

**Category A**

!	+	↑	△	\$
?	+	↑	△	\$
!	X	↑	△	\$
!	+	↓	△	\$
!	+	↑	□	\$
!	+	↑	△	£
!	+	↑	□	£
!	+	↓	△	£
!	X	↑	△	£
?	+	↑	△	£
!	+	↓	□	\$
!	X	↑	□	\$
?	+	↑	□	\$
!	X	↓	△	\$
?	+	↓	△	\$
?	X	↑	△	\$

**Category B**

?	X	↓	□	£
!	X	↓	□	£
?	+	↓	□	£
?	X	↑	□	£
?	X	↓	△	£
?	X	↓	□	\$
?	X	↓	△	\$
?	+	↓	□	\$
?	X	↑	△	£
?	X	↓	△	£
!	X	↓	△	£
?	+	↑	□	£
!	X	↑	□	£
!	+	↓	□	£

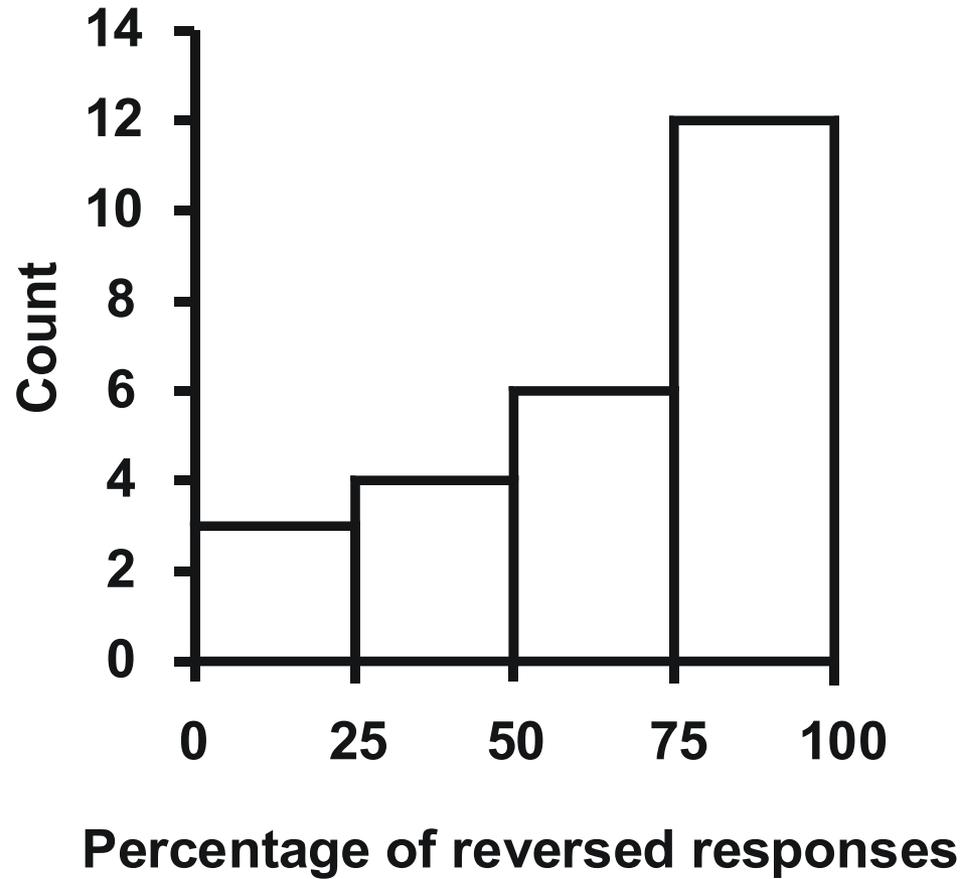
**Reversal**

✓  
✓  
✓  
✓  
✓  
✓  
✓

**%age of reversed responses in test phase**

Without category representations  
0% (Ignore phase 1)  
50%(Ignore phase 2)  
  
With category representations  
Up to 100%

**Mean: 64.5%**



**%age of reversed responses in test phase**

Category A					Category B					Reversal	
!	+	↑	△	\$	?	×	↓	□	£		
?	+	↑	△	\$	!	×	↓	□	£		
!	×	↑	△	\$	?	+	↓	□	£		
!	+	↓	△	\$	?	×	↑	□	£		
!	+	↑	□	\$	?	×	↓	△	£	✓	NR
!	+	↑	△	£	?	×	↓	□	\$	✓	NR
!	+	↑	□	£	?	×	↓	△	\$		
!	+	↓	△	£	?	×	↑	□	\$	✓	R
!	×	↑	△	£	?	+	↓	□	\$	✓	R
?	+	↑	△	£	!	×	↓	□	\$	✓	NR
!	+	↓	□	\$	?	×	↑	△	£	✓	R
!	×	↑	□	\$	?	+	↓	△	£	✓	R
?	+	↑	□	\$	!	×	↓	△	£	✓	NR
!	×	↓	△	\$	?	+	↑	□	£		
?	+	↓	△	\$	!	×	↑	□	£		
?	×	↑	△	\$	!	+	↓	□	£		

Category reps

Up to 50% (ignore both phases)

Up to 25% (ignore phase 1)

Down to 0% (ignore phase 2)

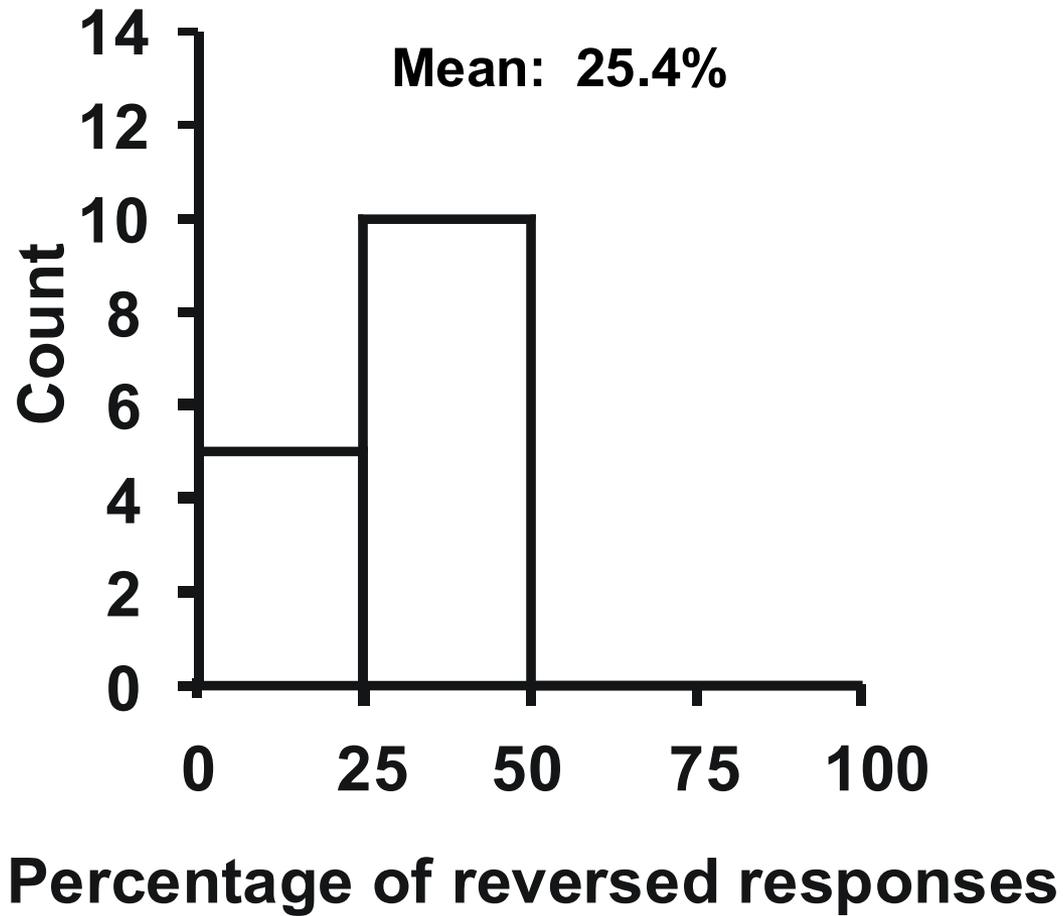
Feature-level

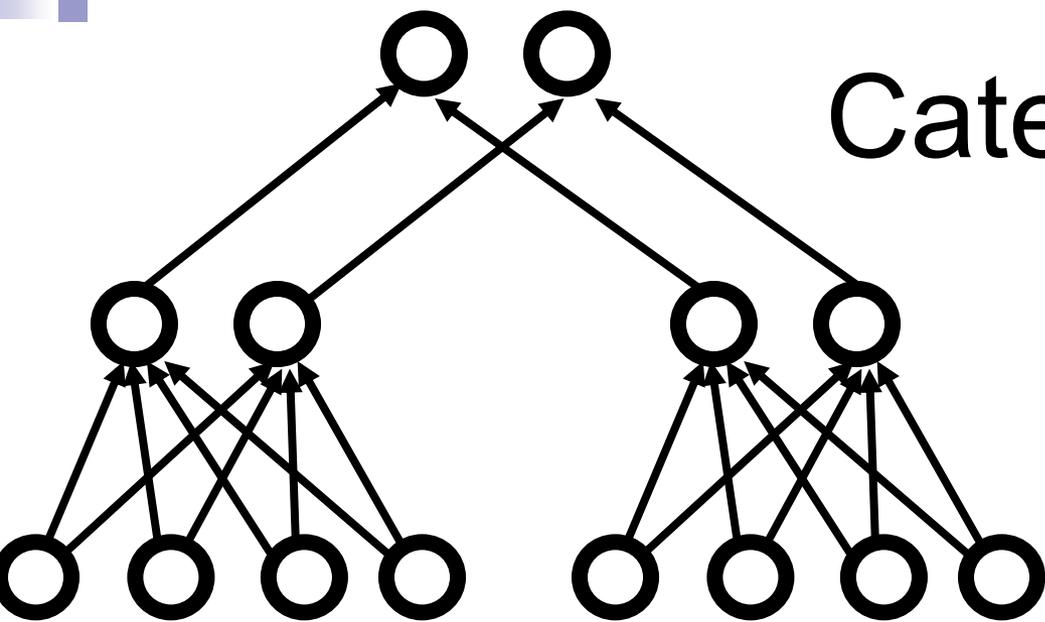
Up to 75% (reversal generalizes perfectly to non-diagnostic feature-pairs)

Neither

50% or less (similarity)

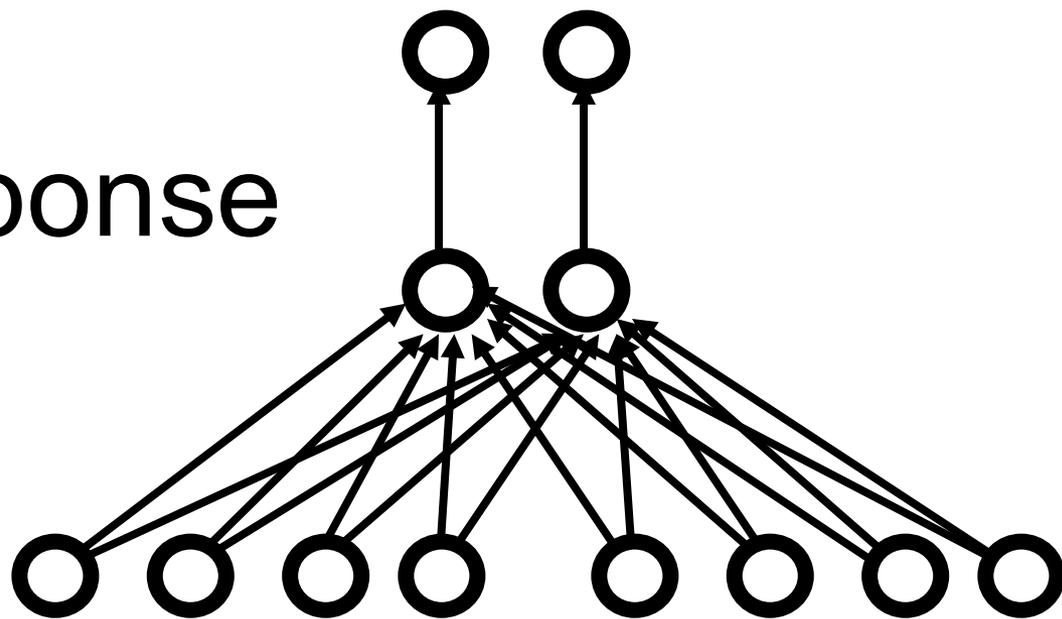
Both 0-75%

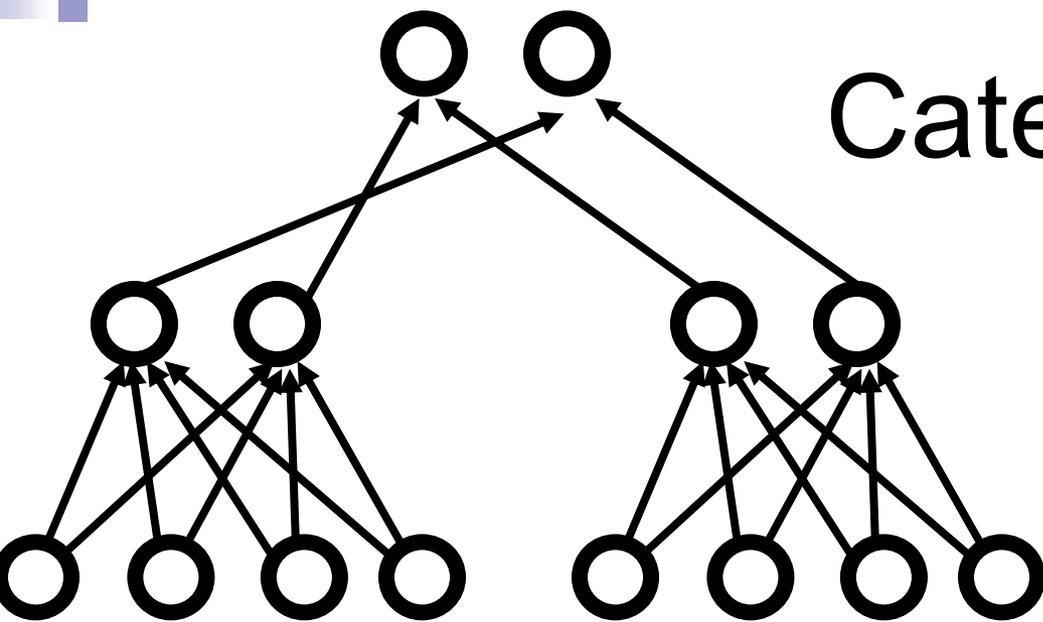




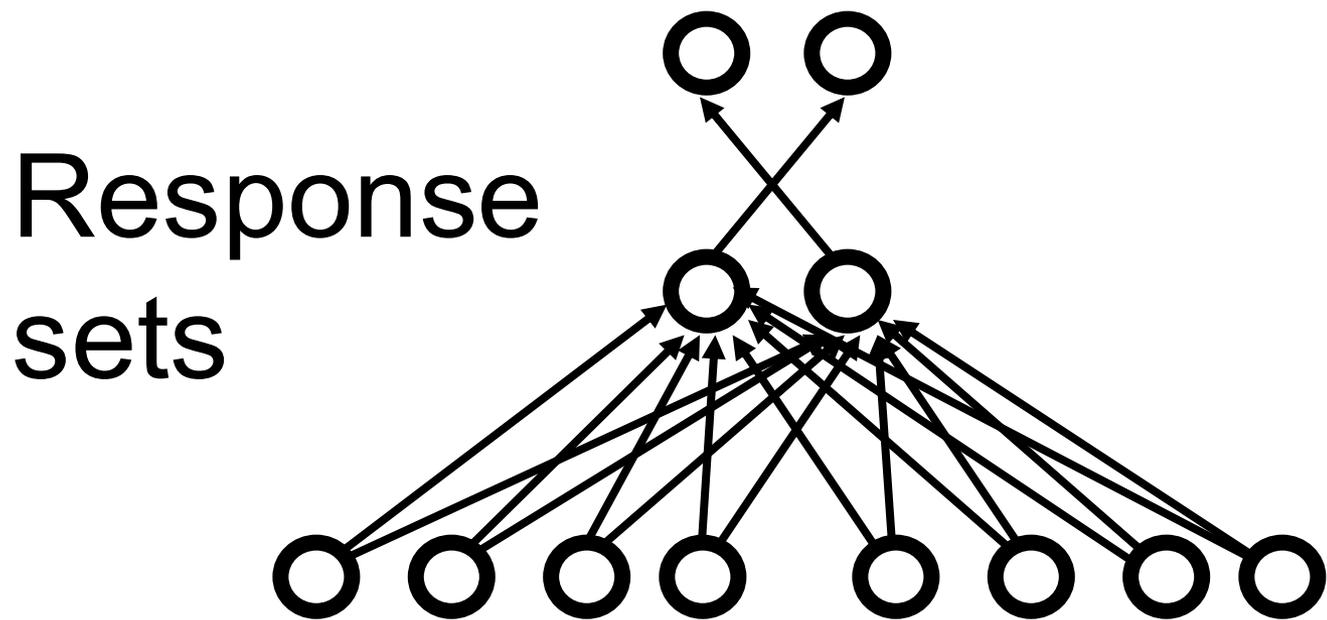
Categories

Response  
sets



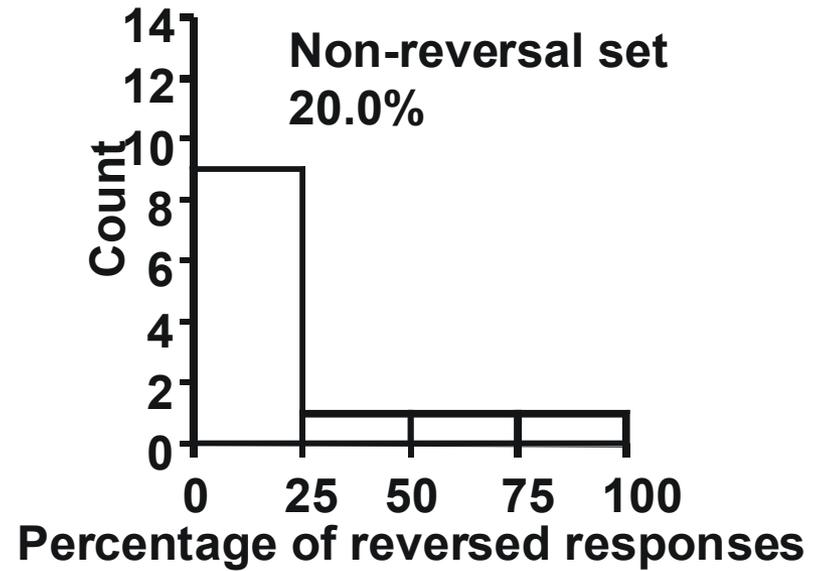
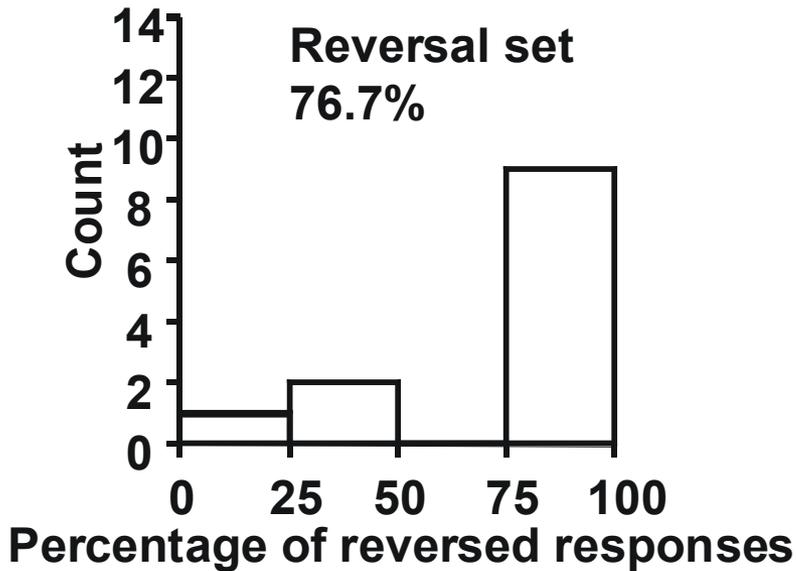


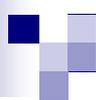
Categories



Response sets

Category A	Category B
	
	





# Summary

- These data are consistent with the idea that each distinct class of objects we have encountered evokes a distinct category representation that is dissociable from the category label it is given
- “a rose by any other name,” rather than “a rose is a rose.”