Pattern Recognition

- Why is pattern recognition important?
  - Humans' ability to recognize patterns is what separates us most from machines
- Models of pattern recognition
  - Templates
  - Features
  - Structural analysis

The Mystery of Pattern Recognition


ABCDE
ABCDE
ABCDE
ABC
ABCDE
ABCDE
ABCDE
ABCDE
ABCDE
ABCD
ABCDE
ABCDE

Plumes of smoke pour from the World Trade Center buildings in New York Tuesday. Some people have seen the face of Satan in the plume of smoke. AP PHOTO

Satan seen in smoke plume

Associated Press

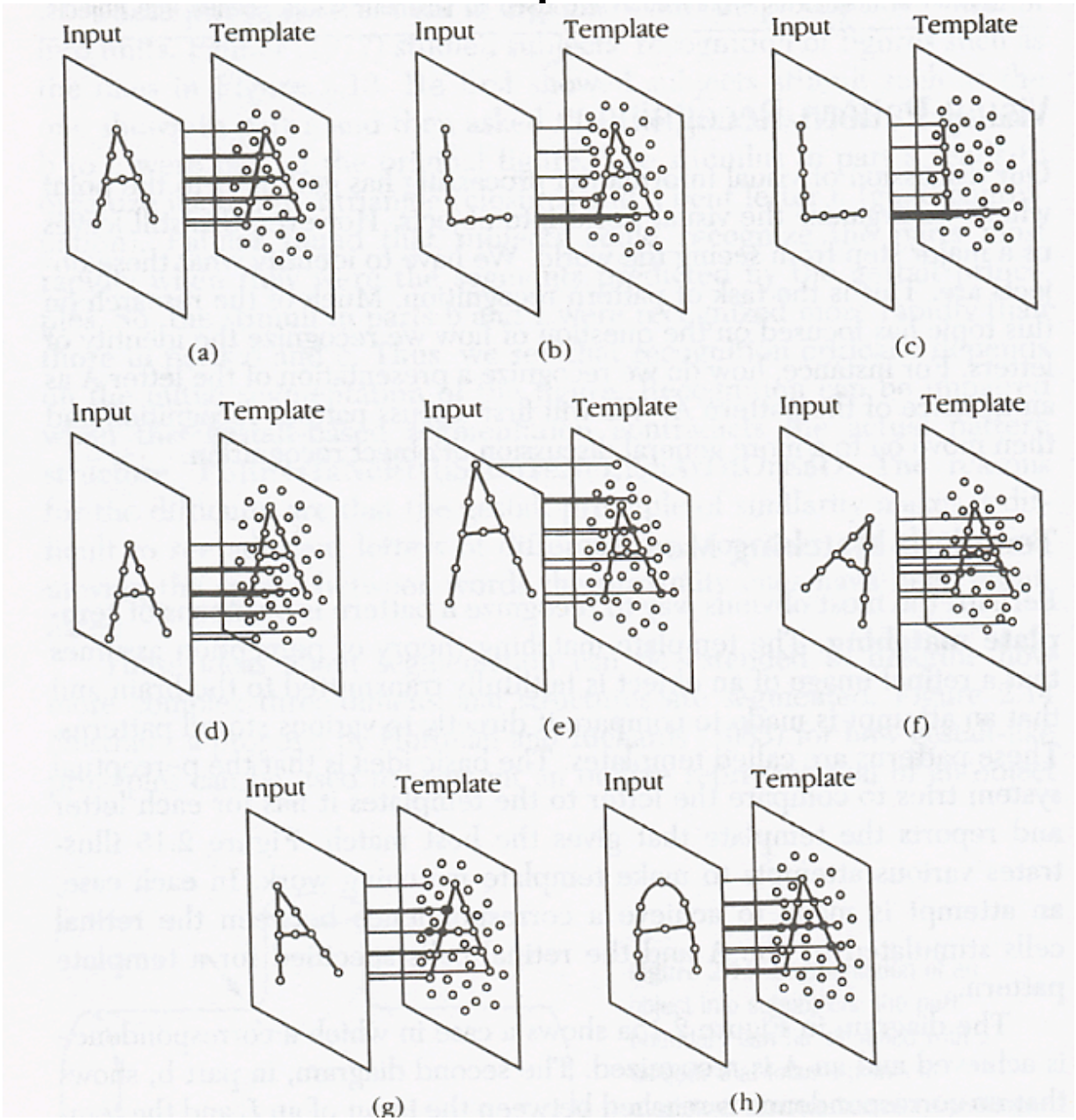
NEW YORK — Several newspapers that printed an Associated Press photo of fire consuming a World Trade Center tower received calls from readers Wednesday saying they saw a strange shape in the smoke.

Some readers said they could identify eyes, a nose, a mouth and horns in the black and gray clouds — and they wondered if the photo had been manipulated to include a satanic face.

Vin Alabiso, an AP vice president and the executive photo editor, said the photo was untouched. Readers were reacting to natural indentations in the smoke clouds, he said Wednesday.

"AP has a very strict written policy which prohibits the alteration of the content of a photo in any way," he said. "The smoke in this photo combined with light and shadow has created an image which readers have seen in different ways."

Templates

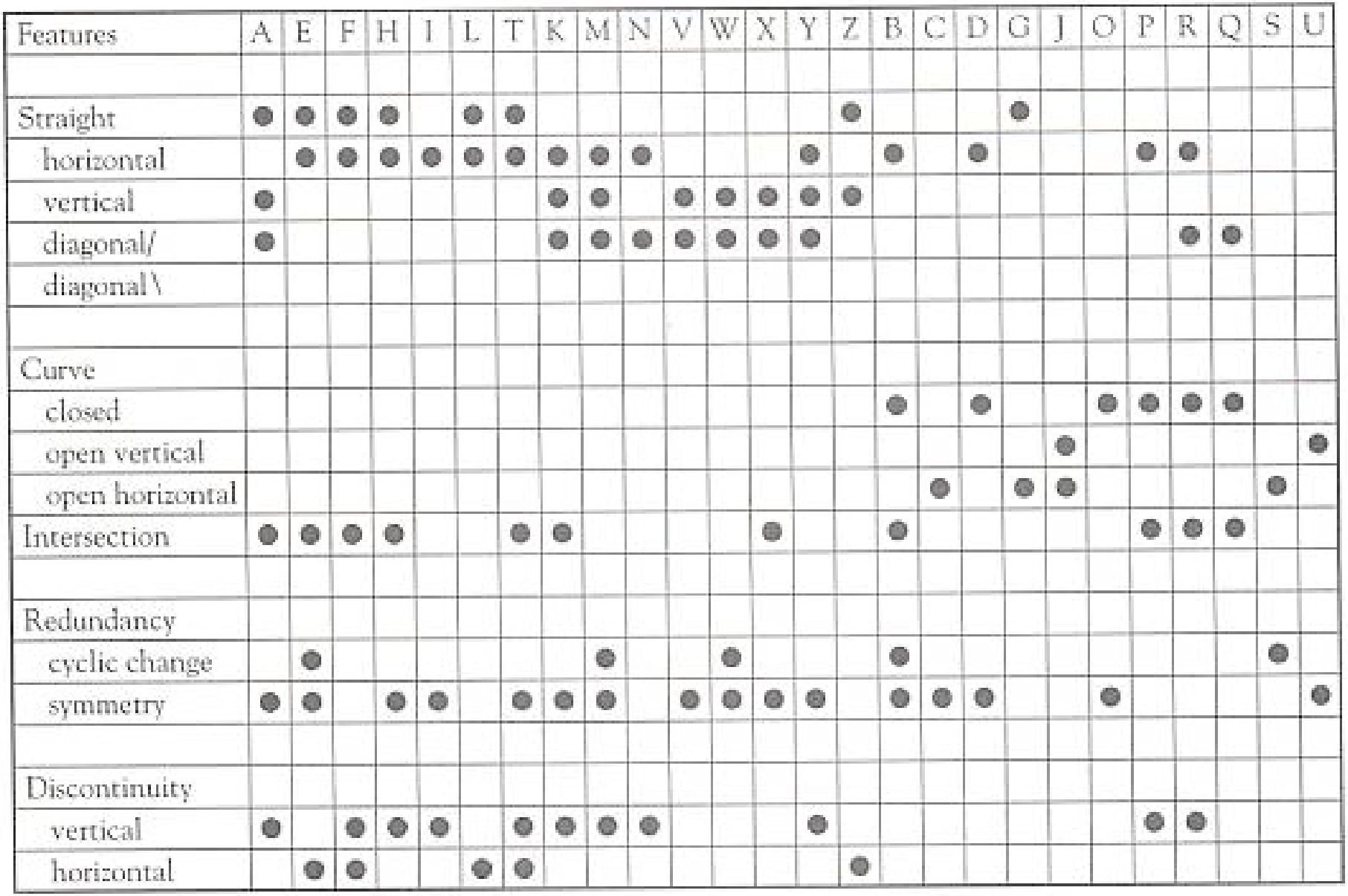

Input
Template
(a)
Input
Template
(b)
Input
Template
(c)
Input
Template
(d)
Input
Template
(e)
Input
Template
(f)
Input
Template
(g)
Input
Template
(h)

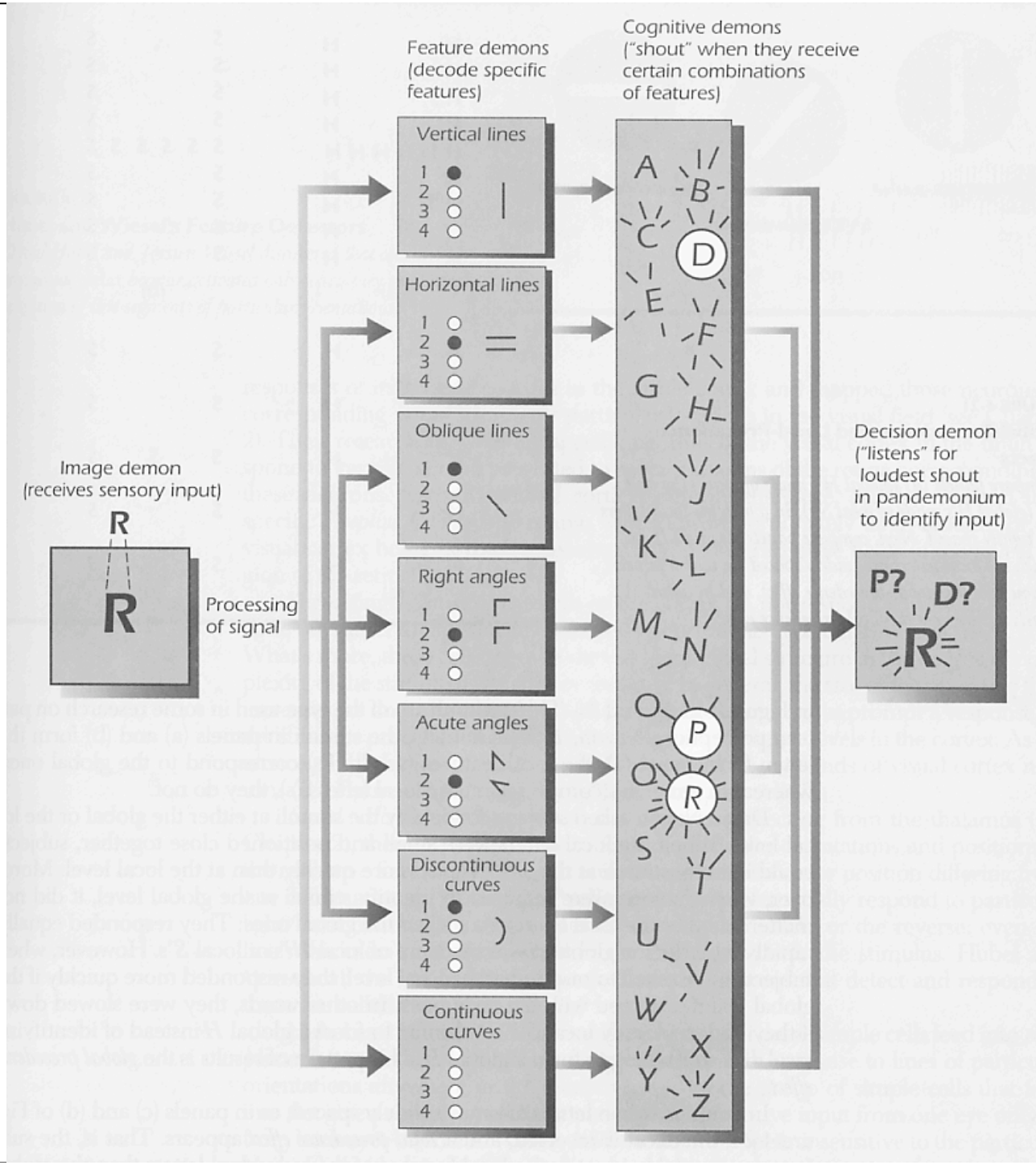
Templates

- Match observed object to stored images
  - "A" is recognized by matching it to stored photograph-like image of previously seen "A"
- Problems
  - Too many templates needed
    - One template for "A" is probably not enough
    - Need a template for "A" with a specific size, orientation, color, etc.
    - Can this problem be solved with pre-processors?
  - Ignores intuition that objects are composed of smaller parts

Feature Analysis

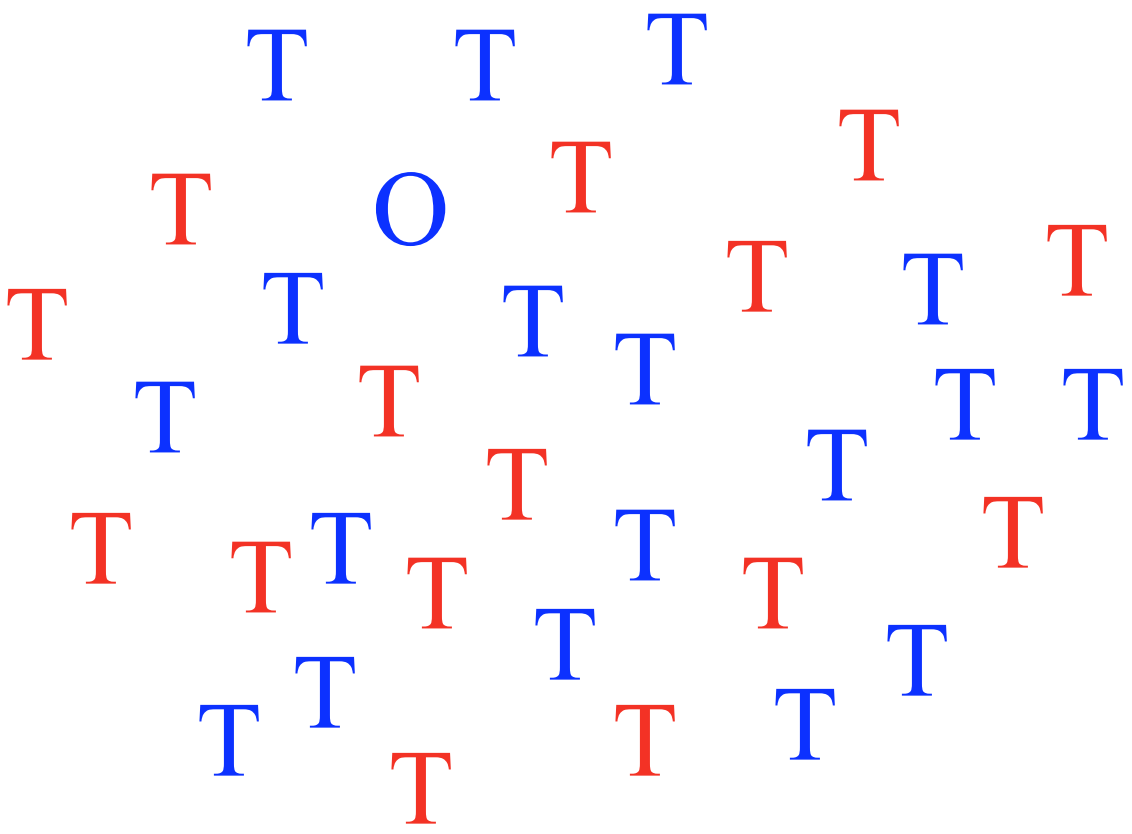
- Recognize an object by breaking it down into features
  - "A" is recognized by combining evidence for \ + - + /
- Evidence for feature analysis
  - Neural feature detectors have been found
  - Simple and conjunctive feature search tasks
    - Simple features are detected in parallel, but combining features requires attention to be moved across an image in a serial manner
    - Asymmetries in feature search
  - Illusory conjunctions


Features
A E F H I L T K M N V W X Y Z B C D G J O P R Q S U
Straight
horizontal
vertical
diagonal/
diagonal\
Curve
closed
open vertical
open horizontal
Intersection
Redundancy
cyclic change
symmetry
Discontinuity
vertical
horizontal


Feature demons
(decode specific
features)
Cognitive demons
("shout" when they receive
certain combinations
of features)
Vertical lines
Horizontal lines
Oblique lines
Right angles
Acute angles
Discontinuous
curves
Continuous
curves
Image demon
(receives sensory input)
R
Processing
of signal
Decision demon
("listens" for
loudest shout
in pandemonium
to identify input)
P?
D?
R

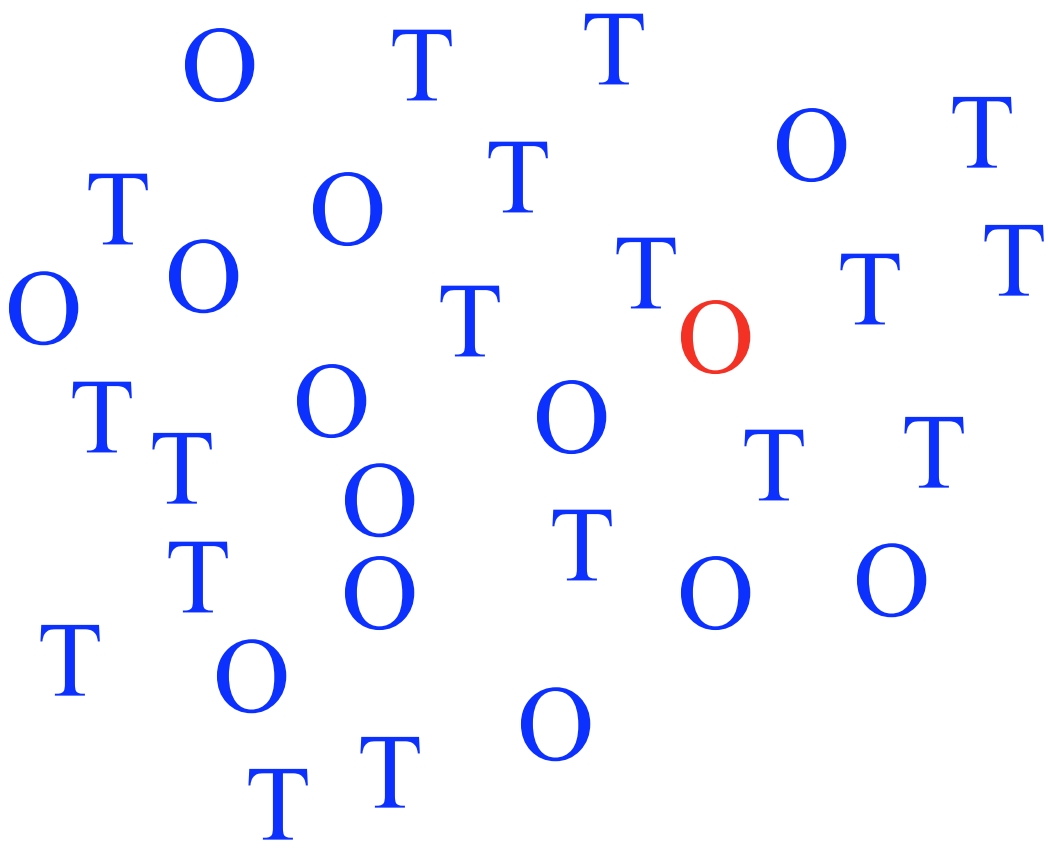
Simple feature search

Look for an “O”



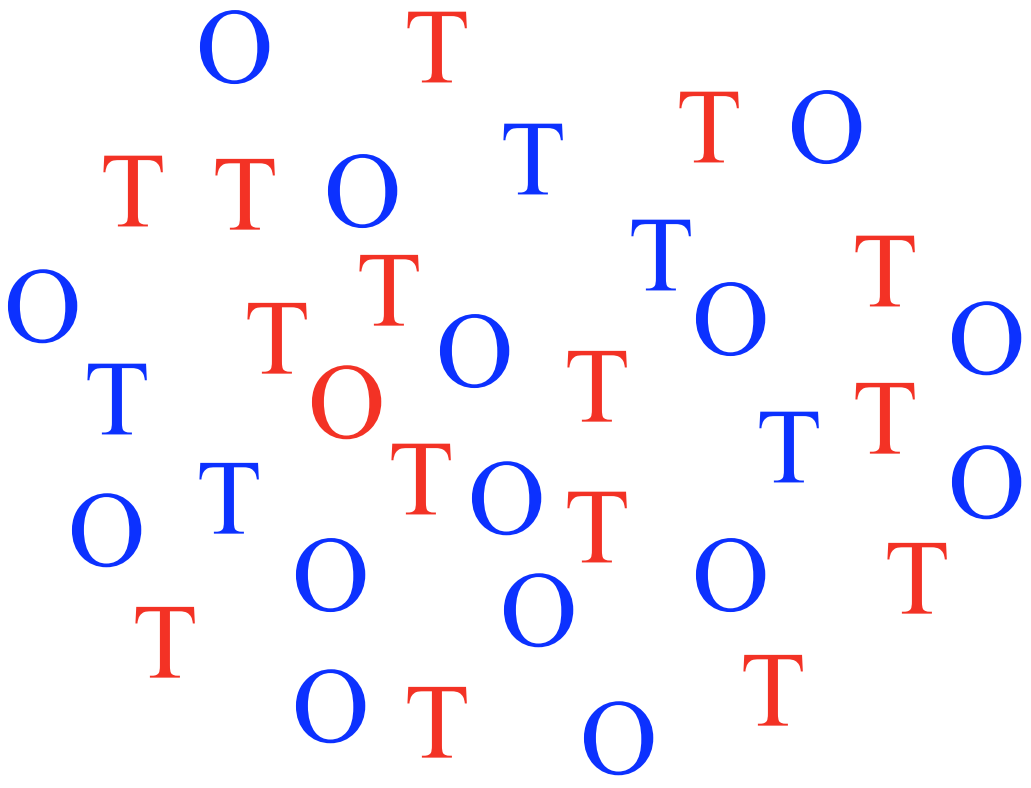
Simple feature search

Look for something red



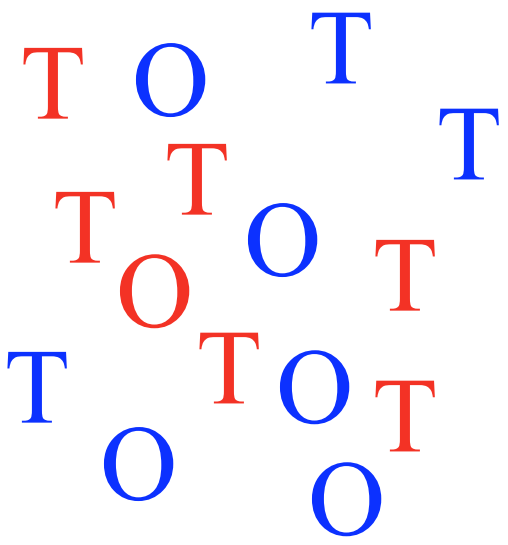
Conjunctive feature search

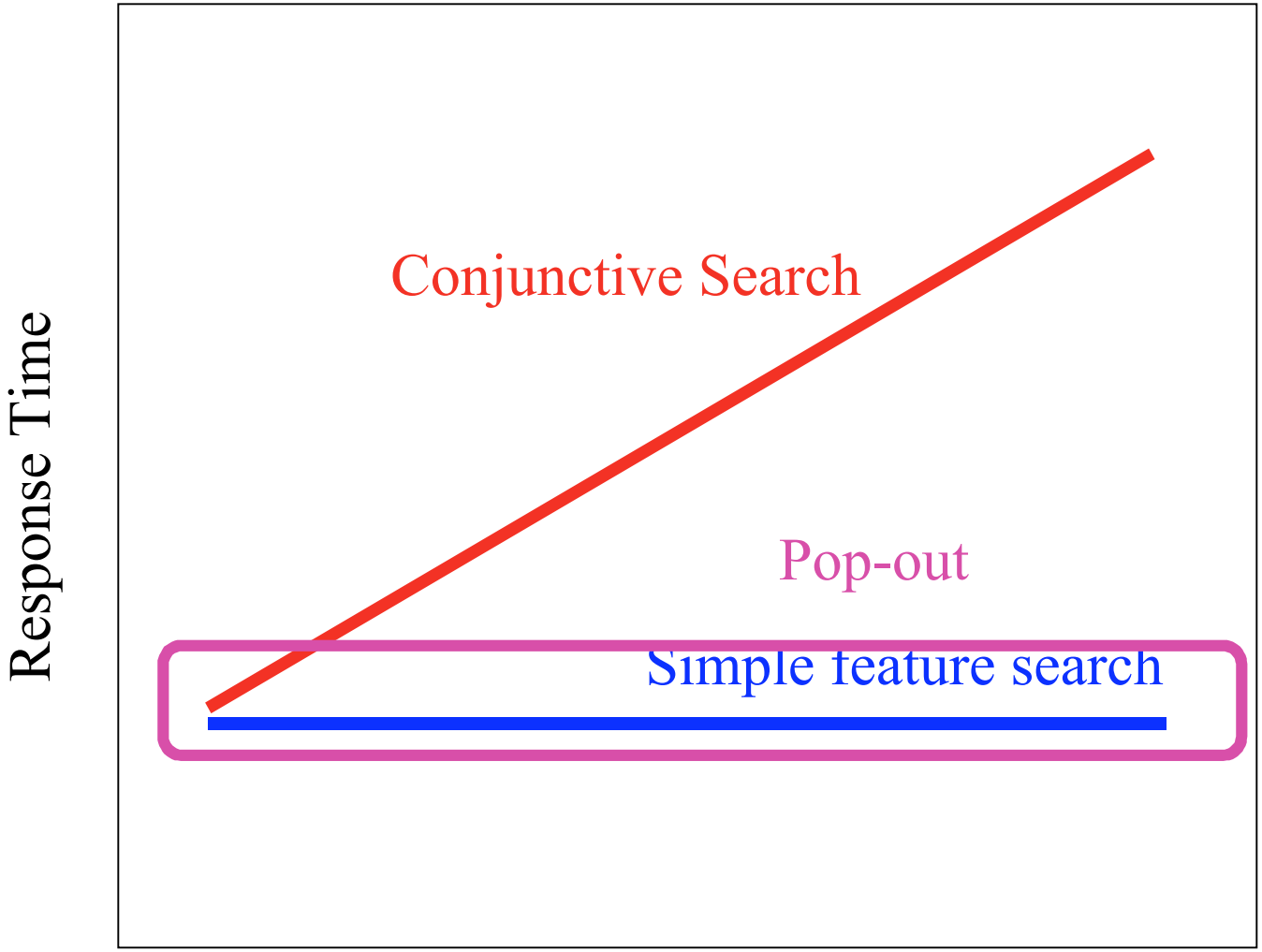
Look for a red “O”



Conjunctive feature search

Look for a red “O”

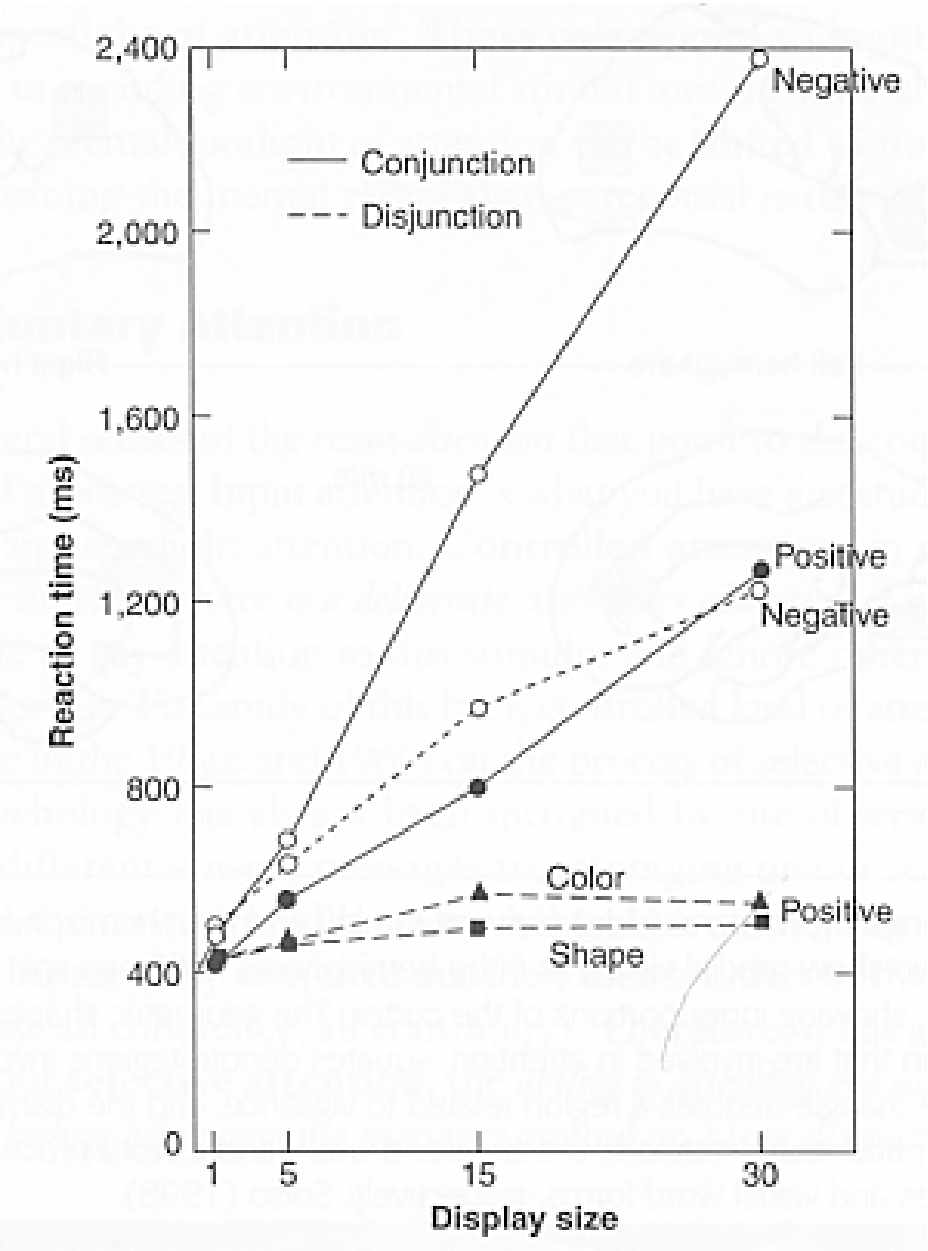



Response Time
Conjunctive Search
Pop-out
Simple feature search
Number of Stimuli in Display

Feature Integration Theory

- Simple features can be detected anywhere in a display, in parallel
  - Pop-out: for a simple feature search task, response time doesn't depend on how many objects are in the display
- Conjunctions of features require attention to bind ("glue") the features together
  - Attention must be deployed serially, to one object at a time
  - Attention as the force that glues otherwise free-floating features together
  - For a conjunctive search task, response increases linearly with number of objects

Disjunctive and conjunctive feature search tasks



Negative = "Absent"

Positive = "Present"

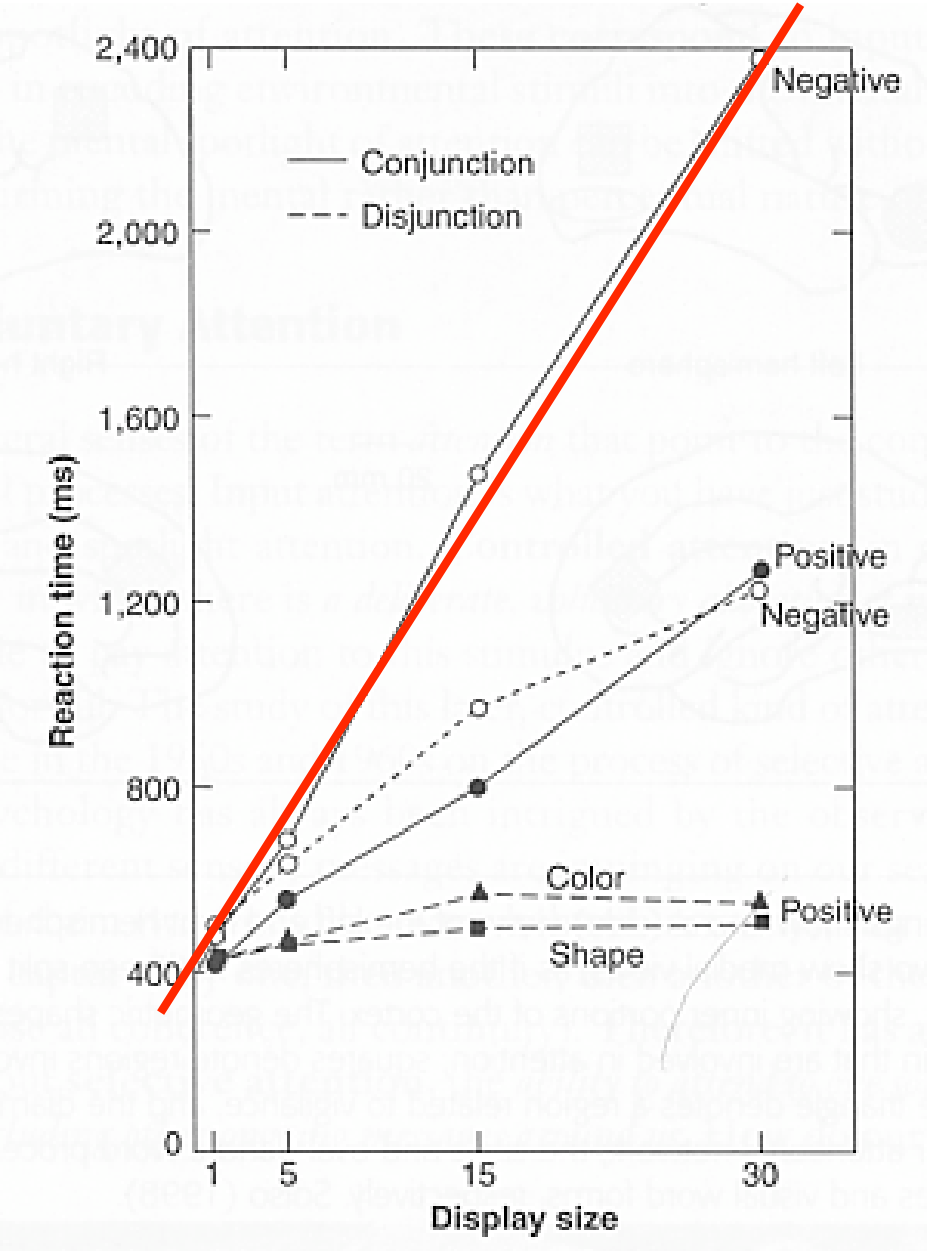
Display size = # of distractors

Conjunctive task: Target is any object that is green AND "S"

Disjunctive task: Target is any object that is green OR "S"

Why is disjunctive task a better control for conjunctive task than a simple feature search like "target is anything green"?

Interpreting Slopes and Intercepts

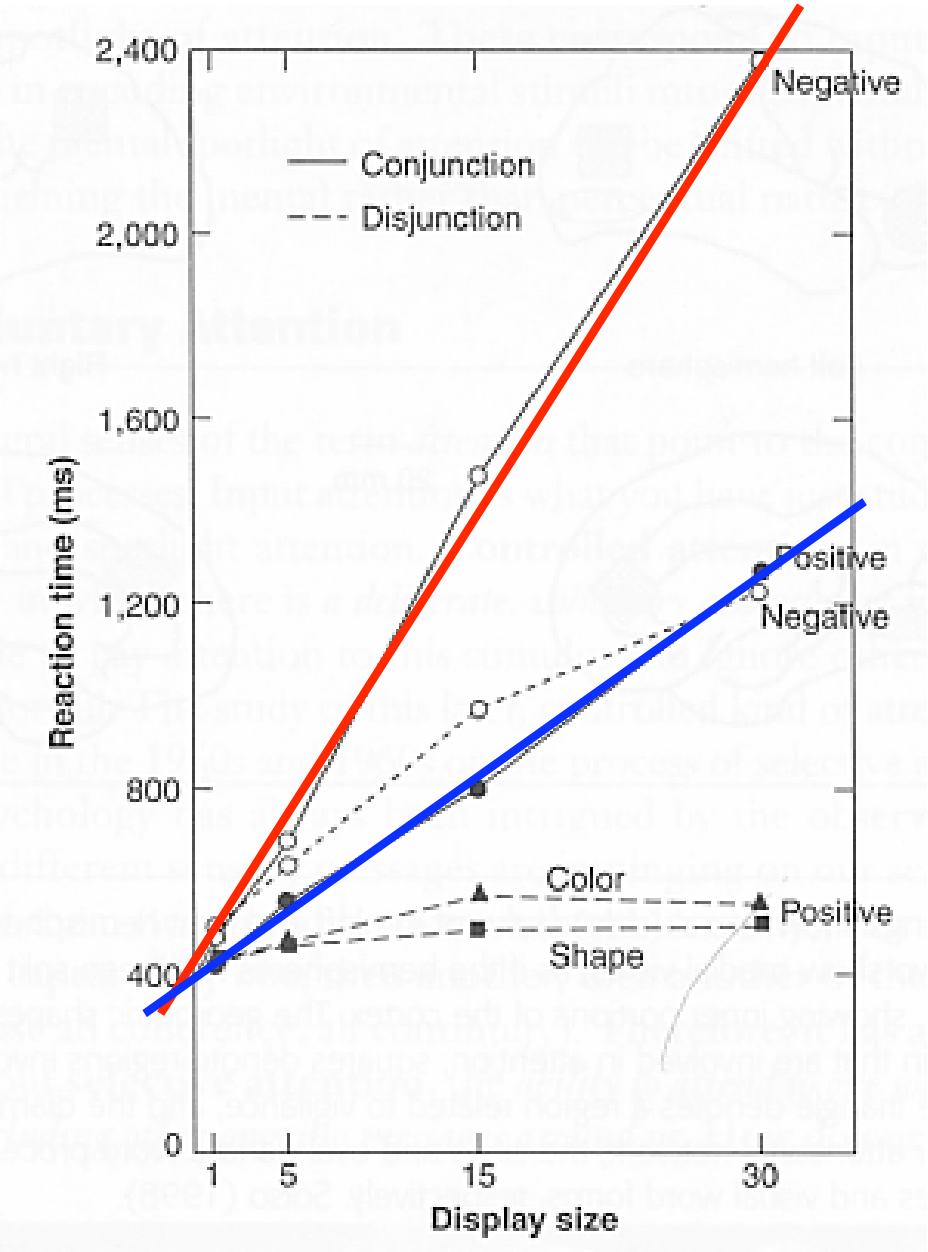


The relation between display size and reaction time is well described by a line.

Slope = number of milliseconds required per item in display

Intercept = amount of time required if there were NO items.

Flat line = "Pop out" = display size does not influence respone time = simultaneous detection of a simple feature everywhere within a display

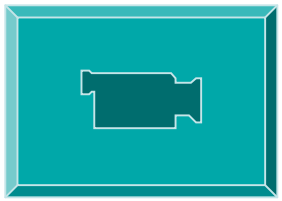


Why is the conjunctive "Absent" slope twice the conjunctive "Present" slope?

Why is there "pop-out" for disjunctive "present" judgments but not "absent" judgments?

Replication of Treisman & Gelade (1980)

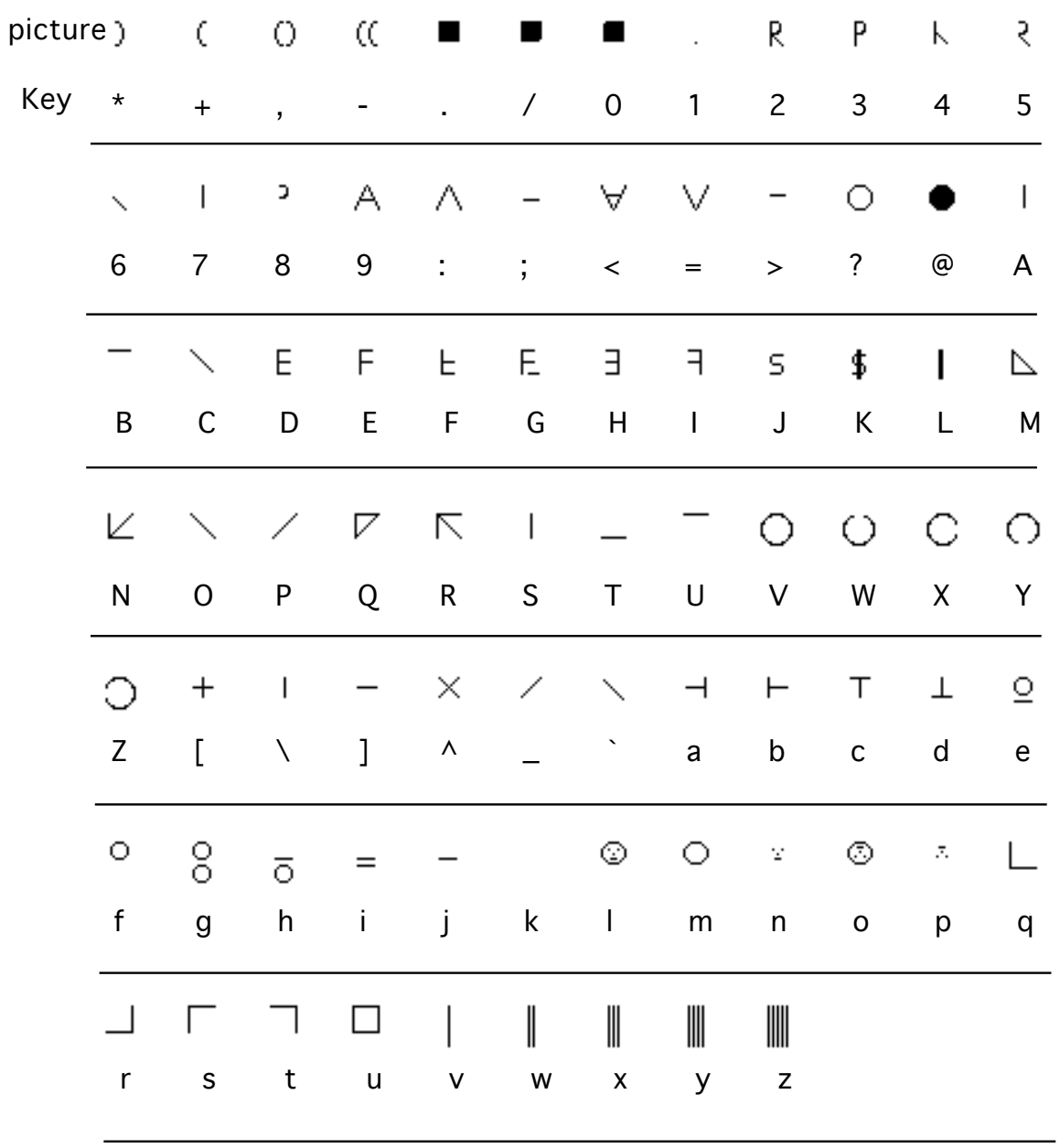
- Variables
  - Task: Conjunctive or Disjunctive
  - Target type: Present or Absent
  - Display size: 1 X 1, 2 X 2, 4 X 4, 6 X 6
  - Dependent variable: Response Time and % Correct
- Predictions
  - Conjunctive response time > Disjunctive response time
  - Interaction between task and display size
    - Display size matters more for conjunctive than disjunctive tasks
  - 2-to-1 slope ratio between conjunctive present and absent tasks
  - Interaction between task, display size, and target type?



Experimental Details

- Tasks
  - Conjunctive
    - Target: Green T
    - Distractors: Red T, Green X
  - Disjunctive
    - Targets: Red or Green S, Blue T or Blue X (anything S OR Blue)
    - Distractors: Red T, Green X
- Display sizes: 1 X 1, 2 X 2, 4 X 4, 6 X 6
- Number of trials
  - 400 total, so 400/(4 array sizes)/(2 tasks) =50 trials per block
  - 50 trials = 25 present and 25 absent trials
- Add jitter: 5 pixels (1 pixel = .034 cm)
- Appearance: Helvetica 24
- Order of 8 conditions: randomized for each participant

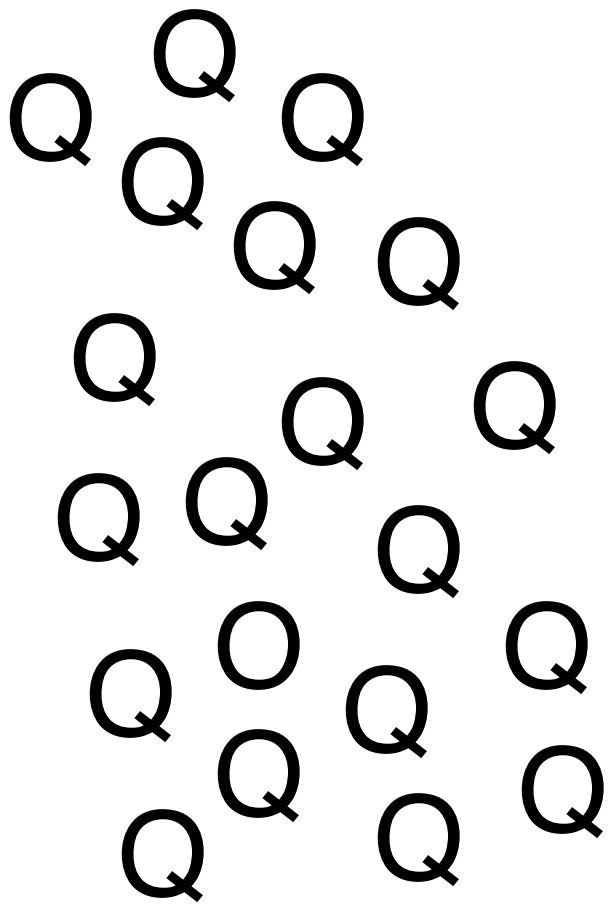
The Cleveland Font



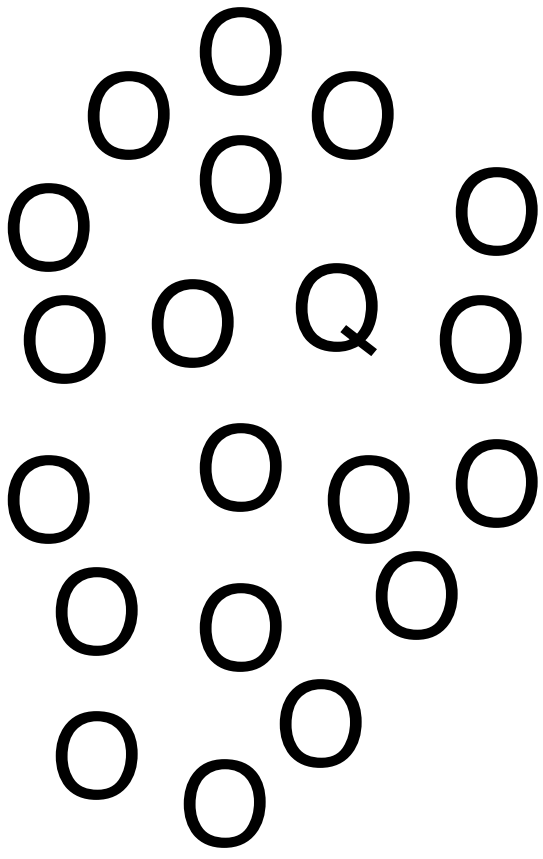
Feature search asymmetries

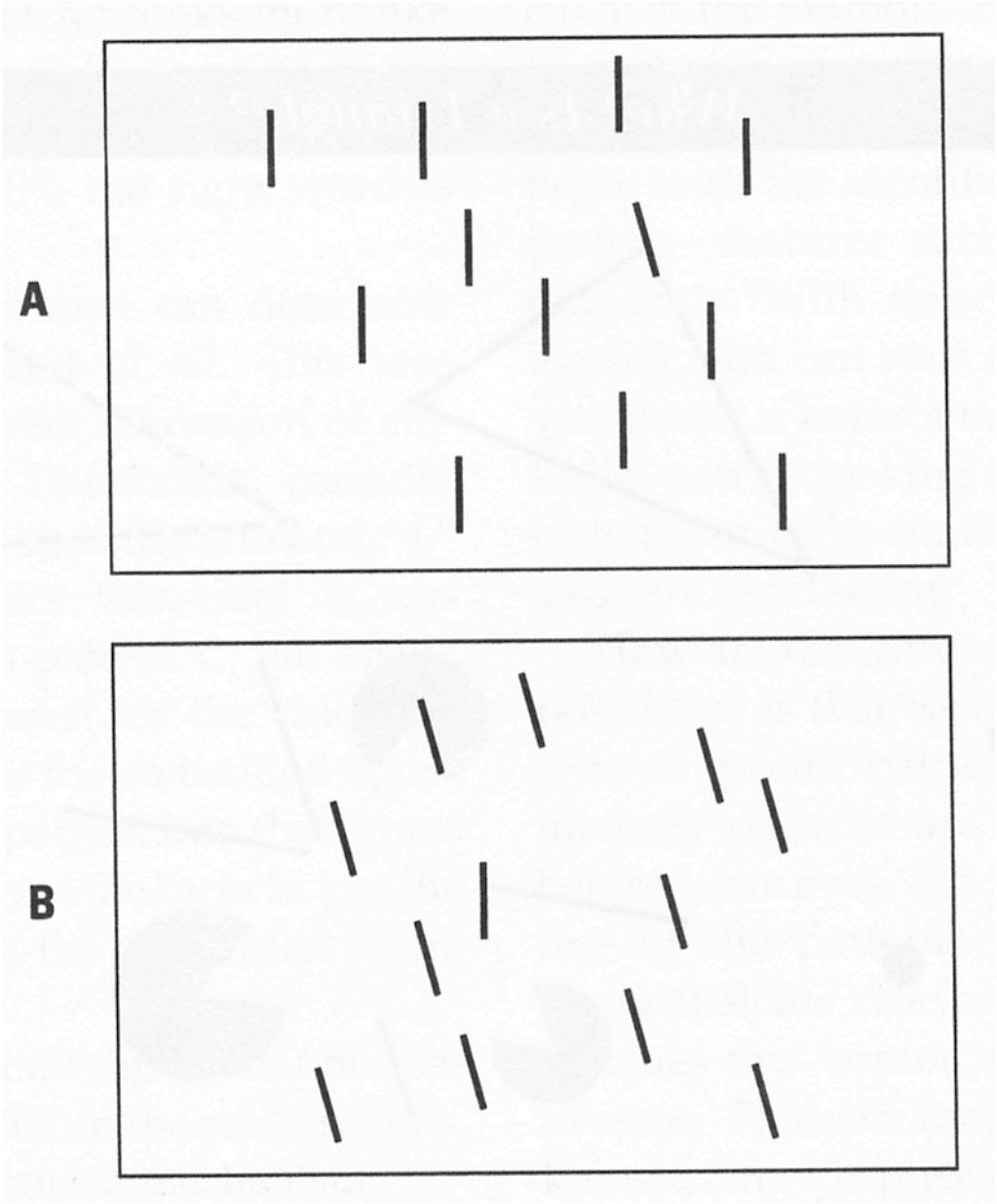
It is easier to find X among Ys than Y among Xs if X has an extra feature compared to Y.

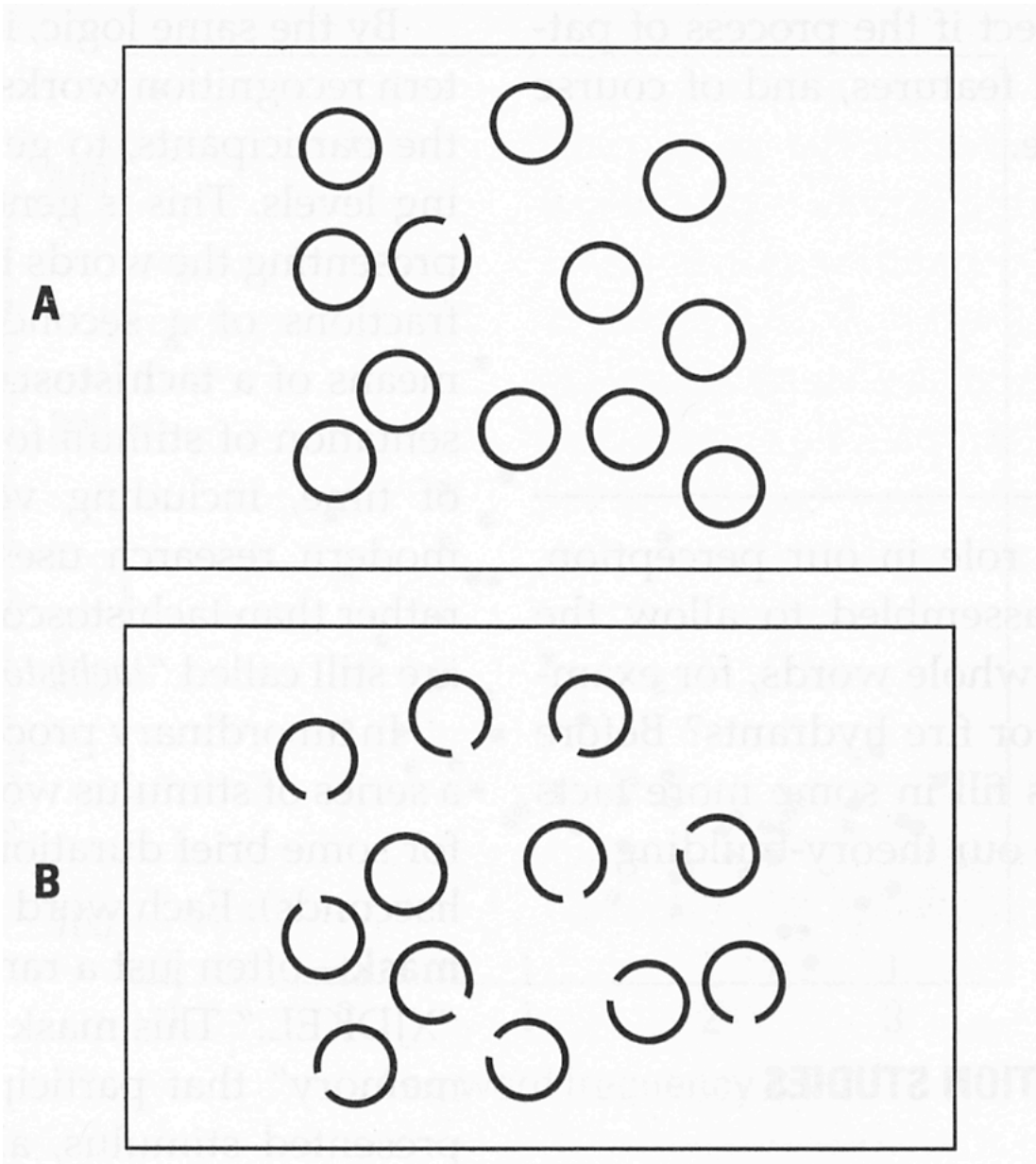
Find the O

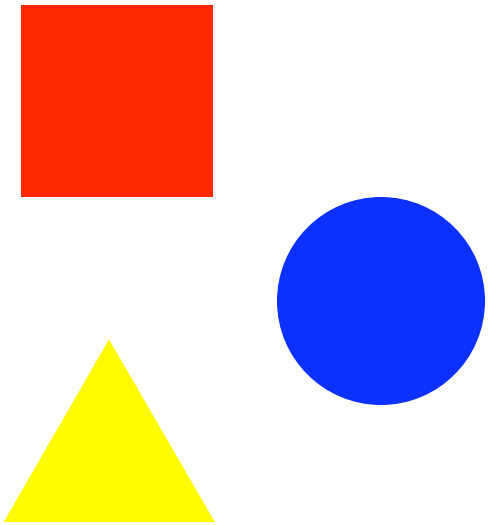


Find the Q




A
B


A
B



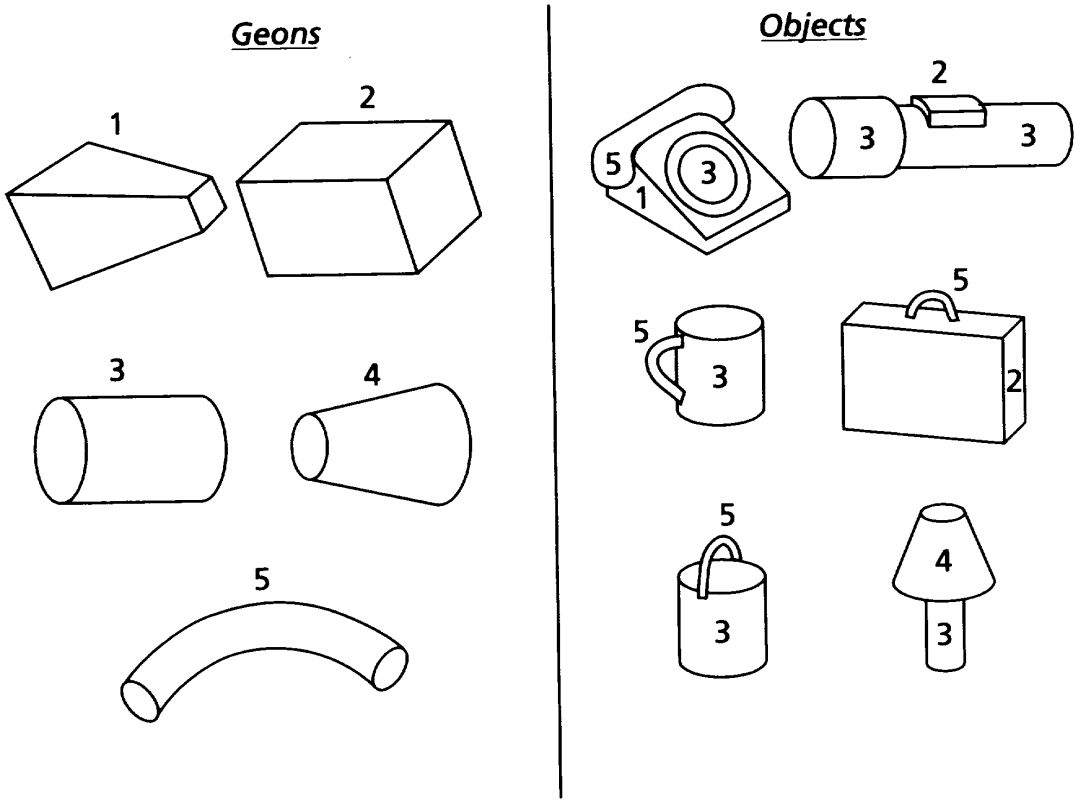
Illusory Conjunctions

When attention cannot be used to bind features together because displays are too fast, then features free-float independently, and may incorrectly recombine with each other.

Structural Analysis

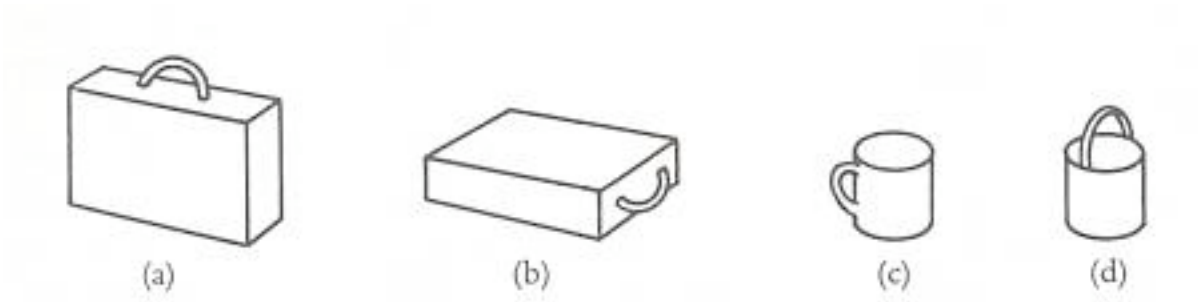
- Represent parts, and relations between parts
- Geon theory
  - A fixed number of primitive geometric components
  - Composed in different arrangements to create all objects
- Evidence for geons
  - Object recognition is hard if object cannot be analyzed into geons

Combining geons to create objects

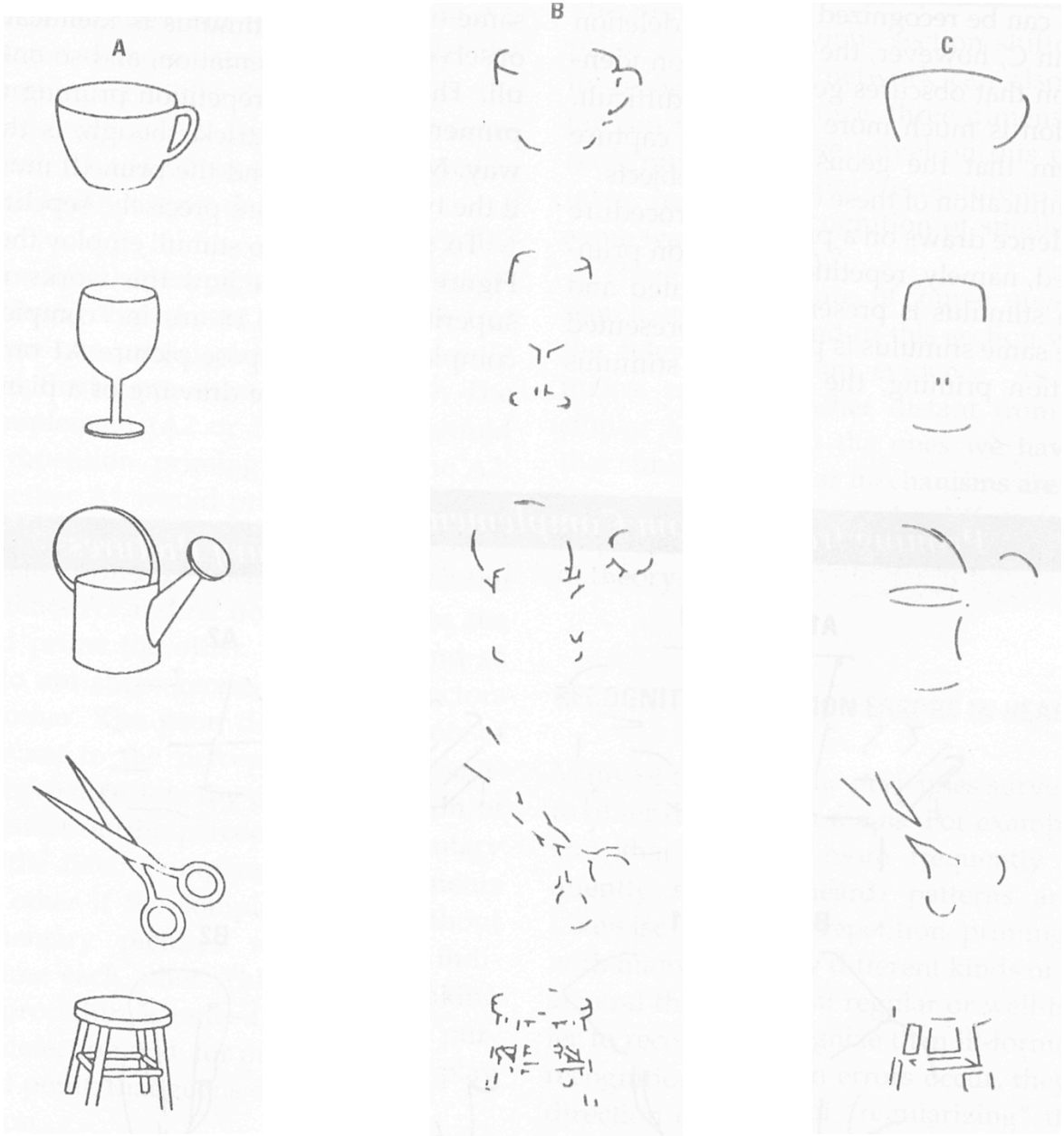


36 geons in all

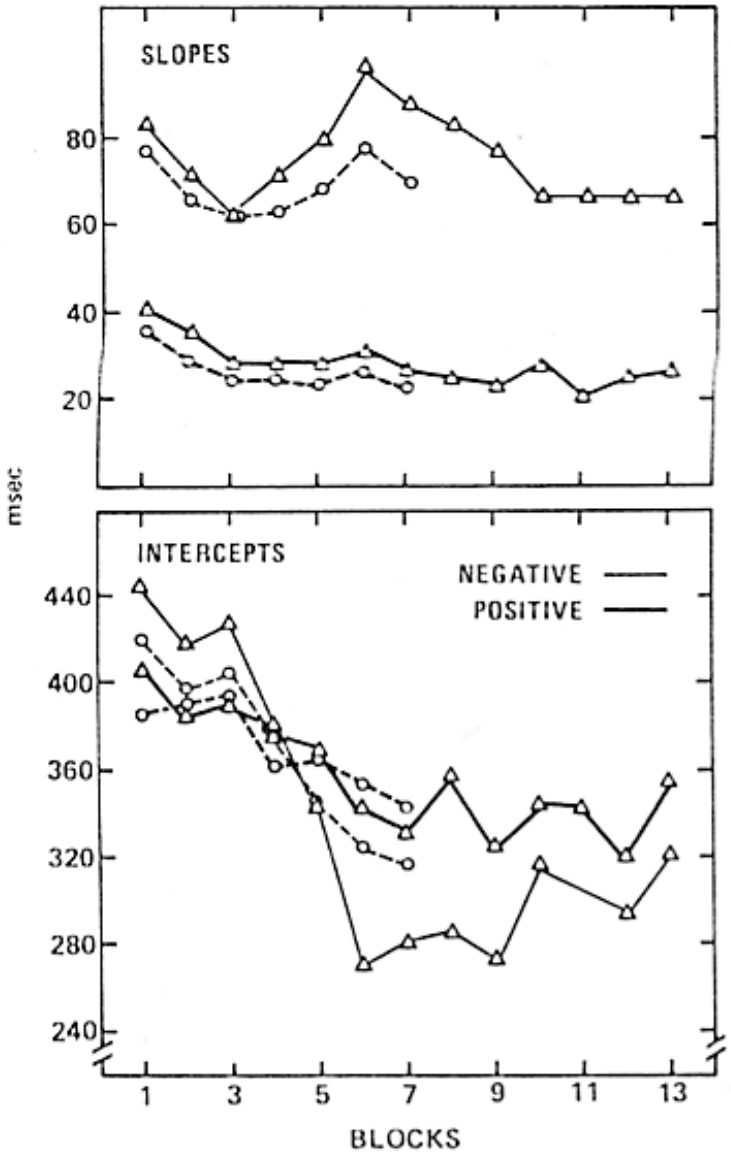
Relations between parts is important



Recognition is easier when geons can be recovered



Practice does not improve Display Size X RT slope for color-form conjunctive task



People do not form color-form units over time

Similarity of Targets to Distractors is not the critical factor for determining search ease

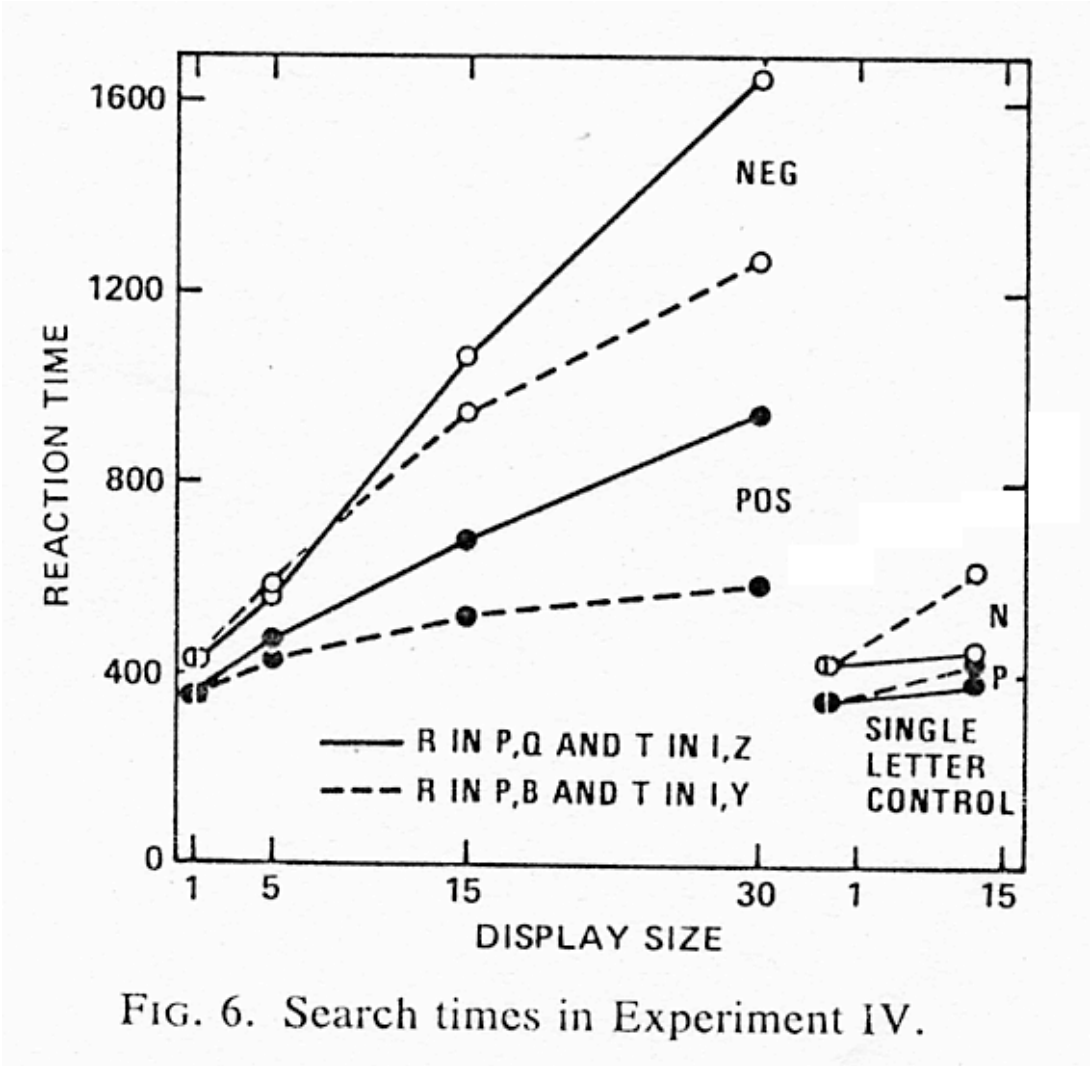


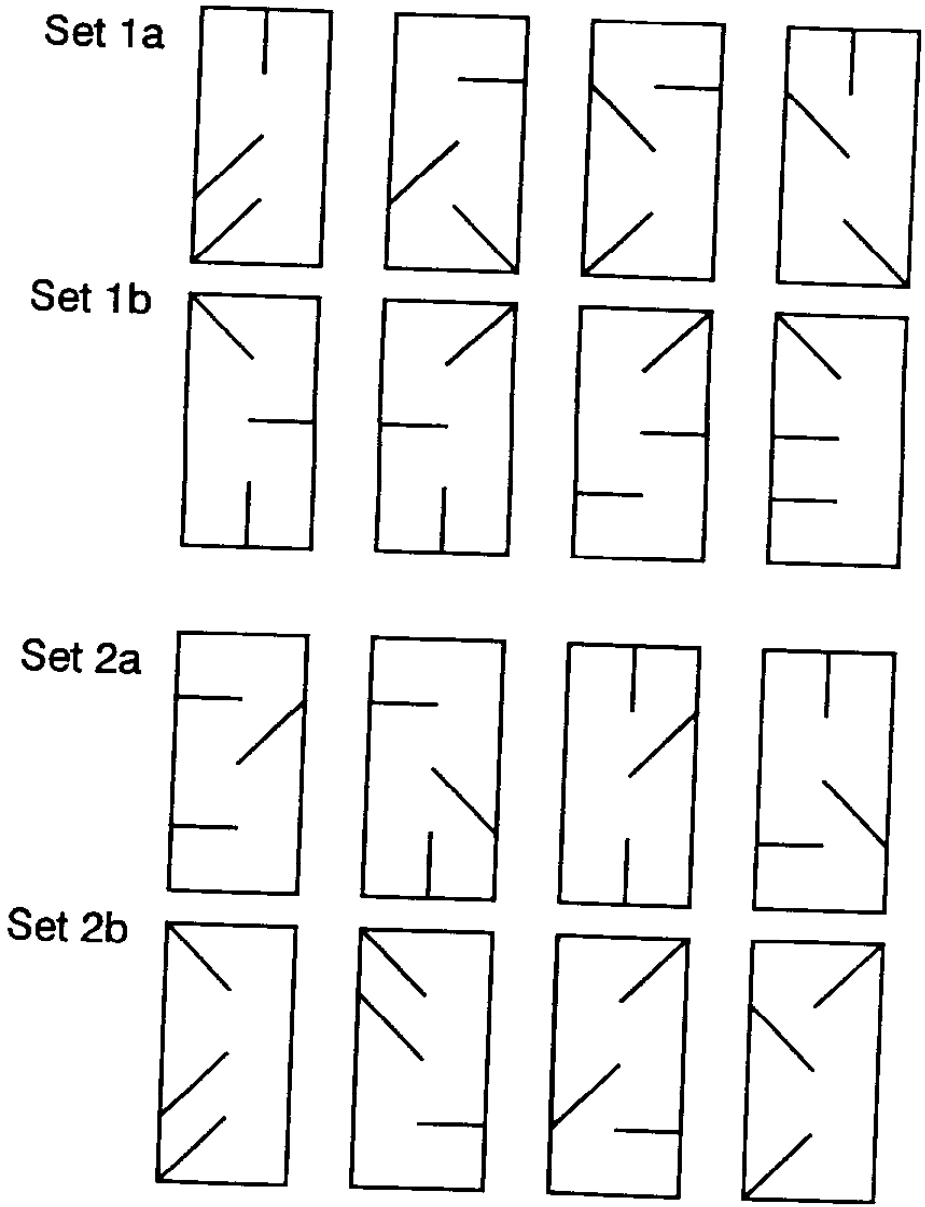
FIG. 6. Search times in Experiment IV.

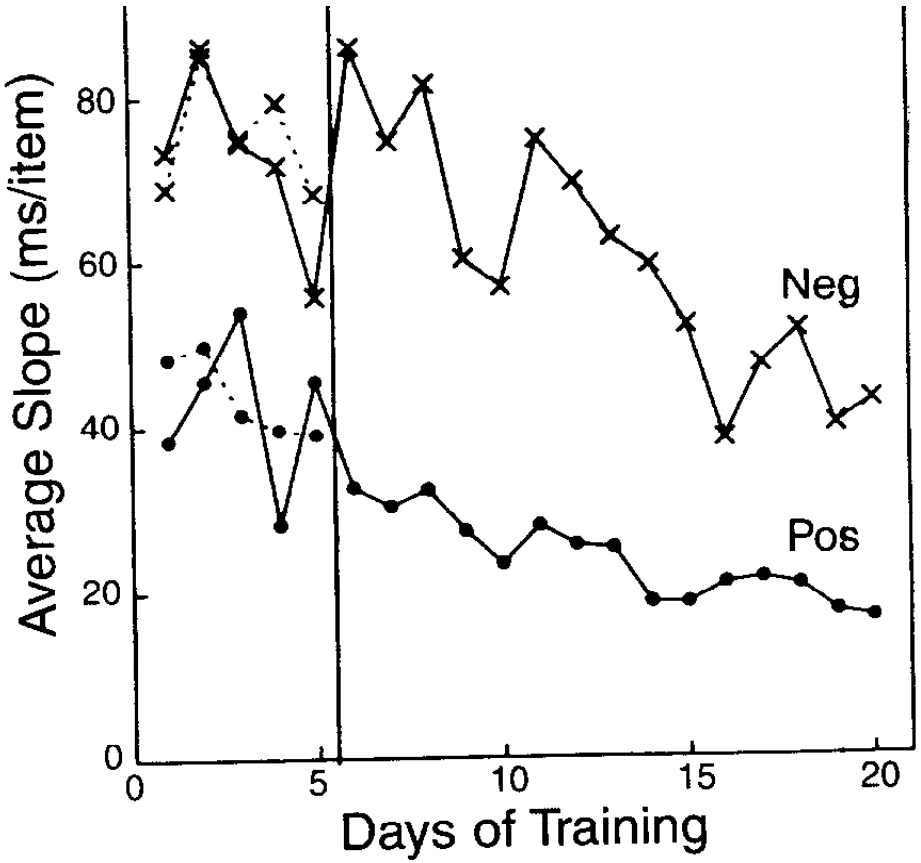
What matters is whether a simple feature can be used to detect a target

Well-learned conjunctions are still recognized by combining their features.

We do not form an "R" unit?

Shiffrin & Lighfoot (1997)





Units made of simple line segments can be created with practice

Unitization of complex forms (Goldstone, 2000)

Category 1

ABCDE

Category 2

ABCDZ

ABCYE

ABXDE

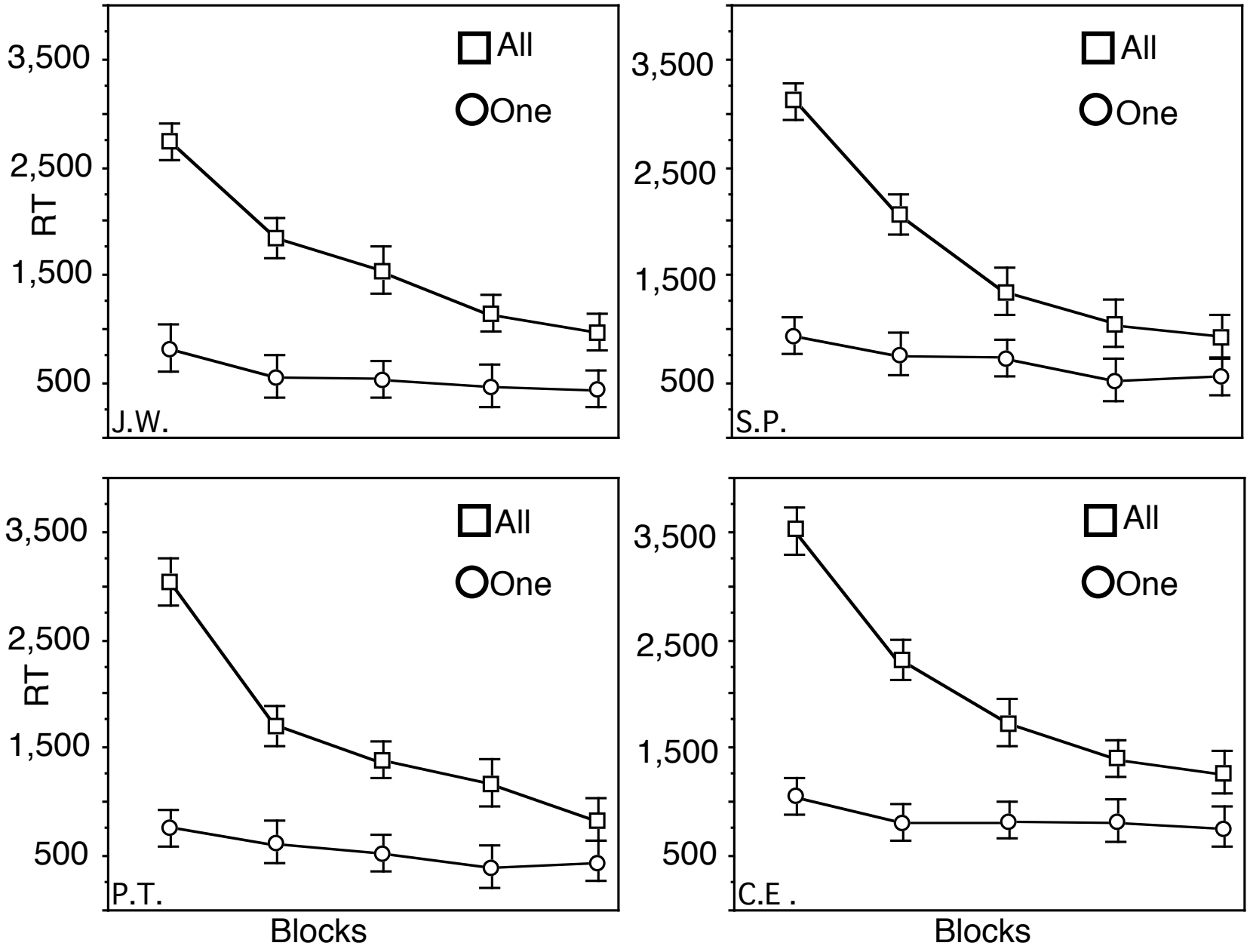
AWCDE

VBCDE

All task: Need to pay attention to A, B, C, D, and E

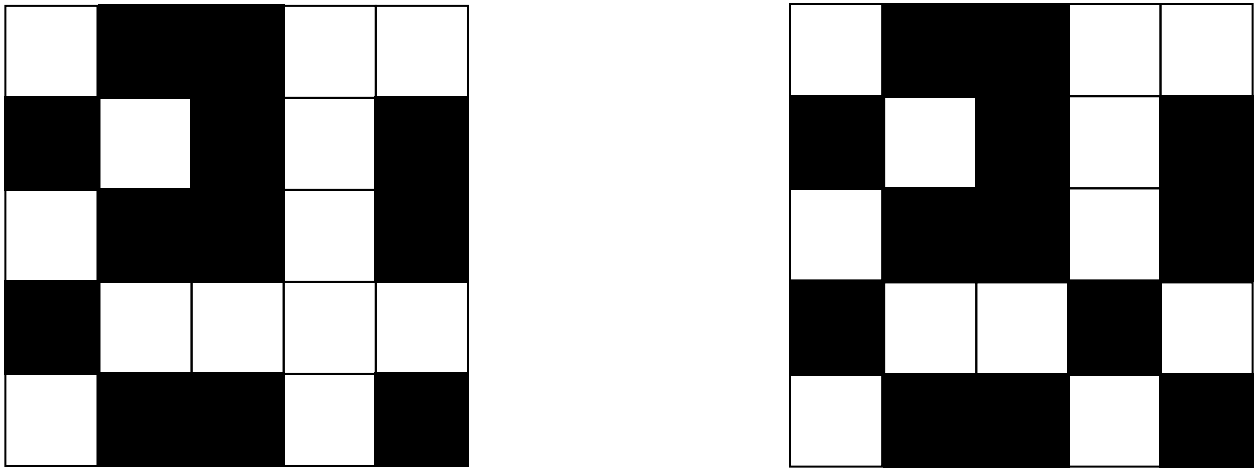
One task: Only need to attend one of the five segments

Responses to a conjunctive target with 20 hours of training

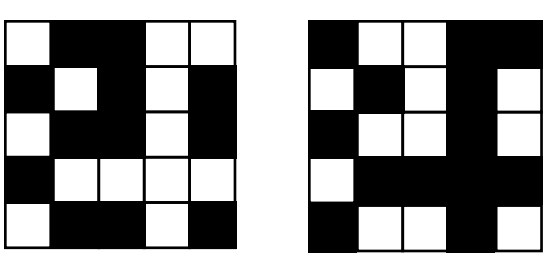


Holism revisited - Holistic "sameness" detector?

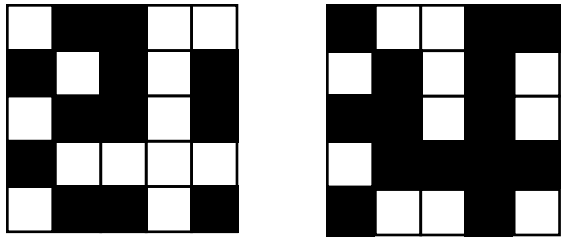
Task 1: Are these two grids cell-for-cell identical?



Task 2: Do these grids have any cells in common?



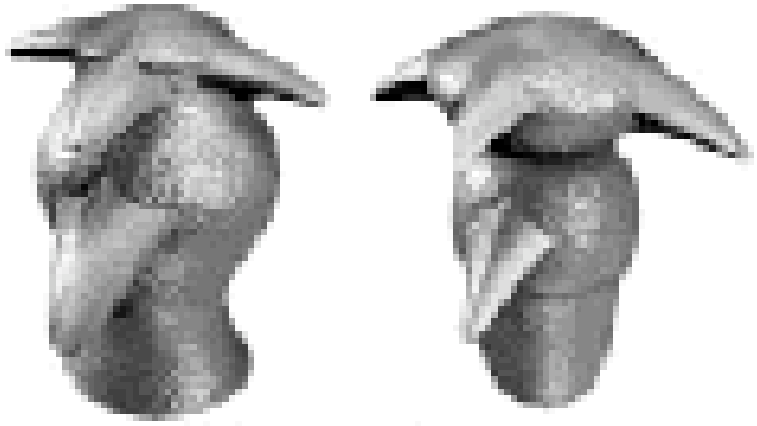
No



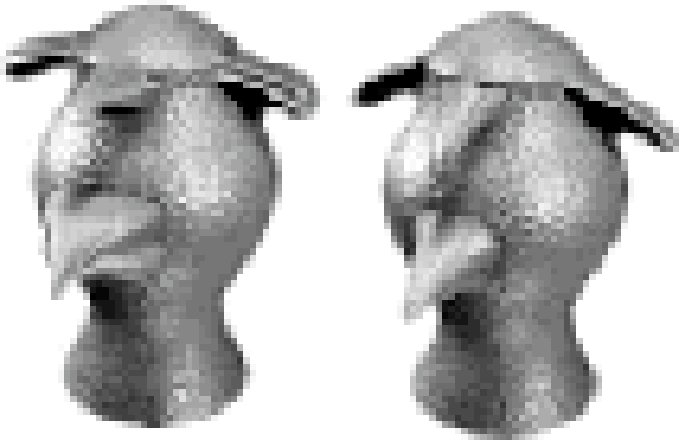
Yes

"Yes" judgments for Task 1 are faster than "Yes" judgments for Task 2, even though, logically (in one sense), Task 1 requires many Task 2-type judgments.

Expertise creates new perceptual units (Gauthier et al, 1999)



Different families



Different individuals

Fusiform gyrus (IT) activity due to novel stimuli (Gauthier et al, 1999)

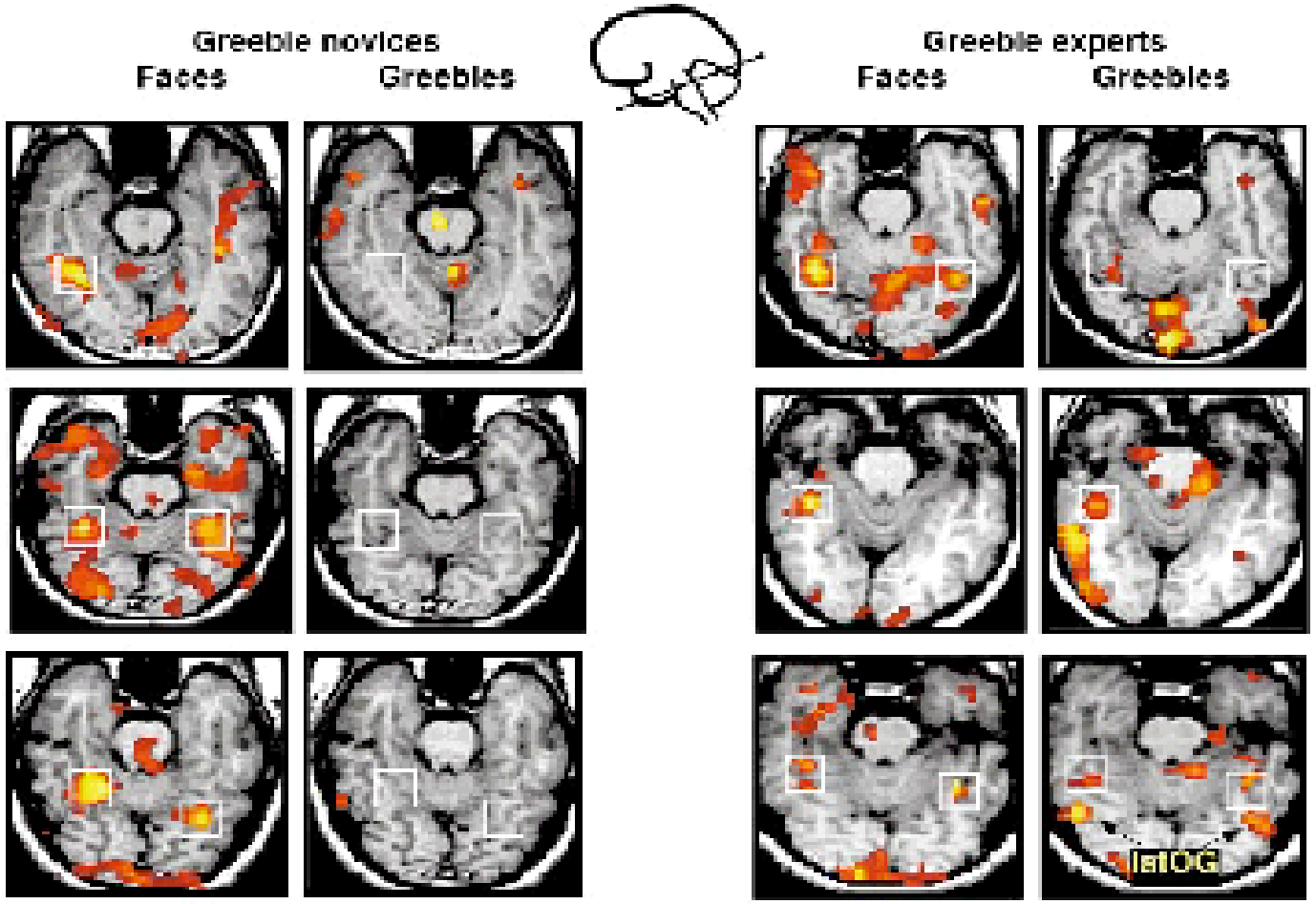
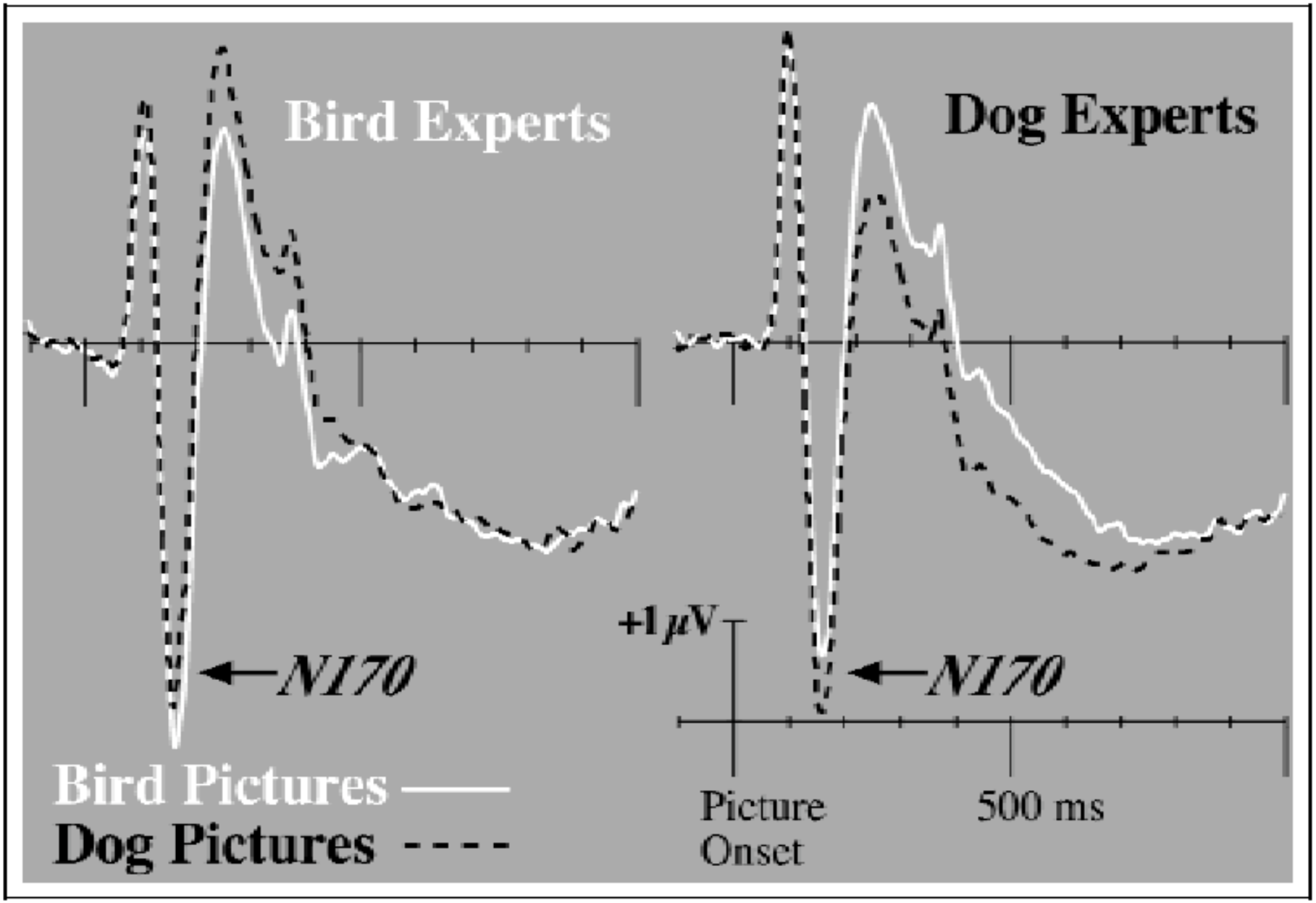


Fig. 4. Activation maps for three novices and three experts in the passive-viewing tasks with faces and greebles. A baseline of passive viewing of objects is used in both conditions, and only the voxels showing more activation for faces or greebles than objects are shown. Images are thresholded at an arbitrary value of $t = 0.75$. Note that we do not attribute any statistical meaning to individual subjects' t-values. The statistical significance of the effects is determined by their representation in the group sample. White squares, middle fusiform gyri ROI; arrows, lateral occipital gyrus foci for one expert (bottom right).

Differences in object recognition due to expertise occur within 170 msec of stimulus onset (Curran & Tanaka, 2001)



Holism as interactions between parts (Farah, 1992)

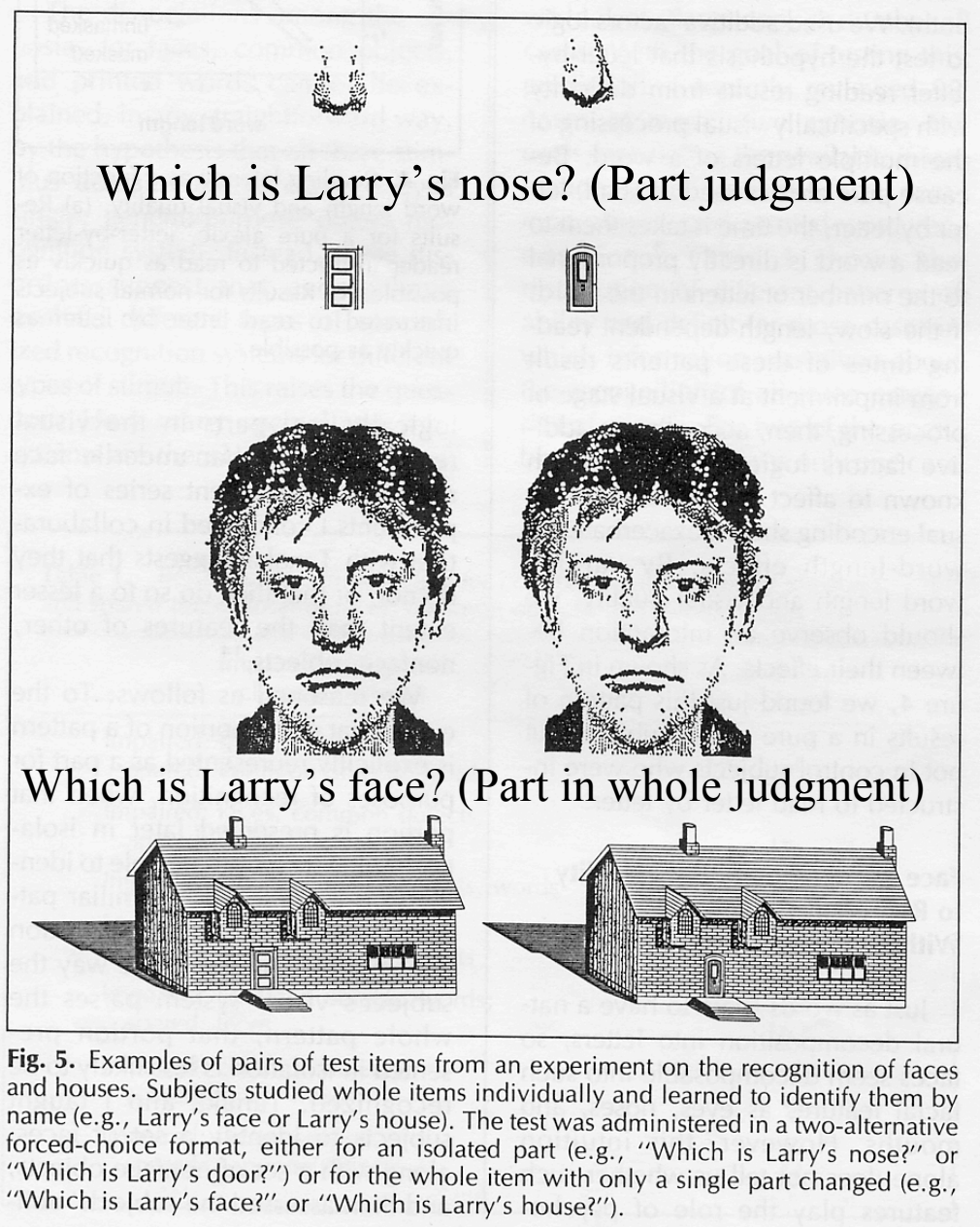
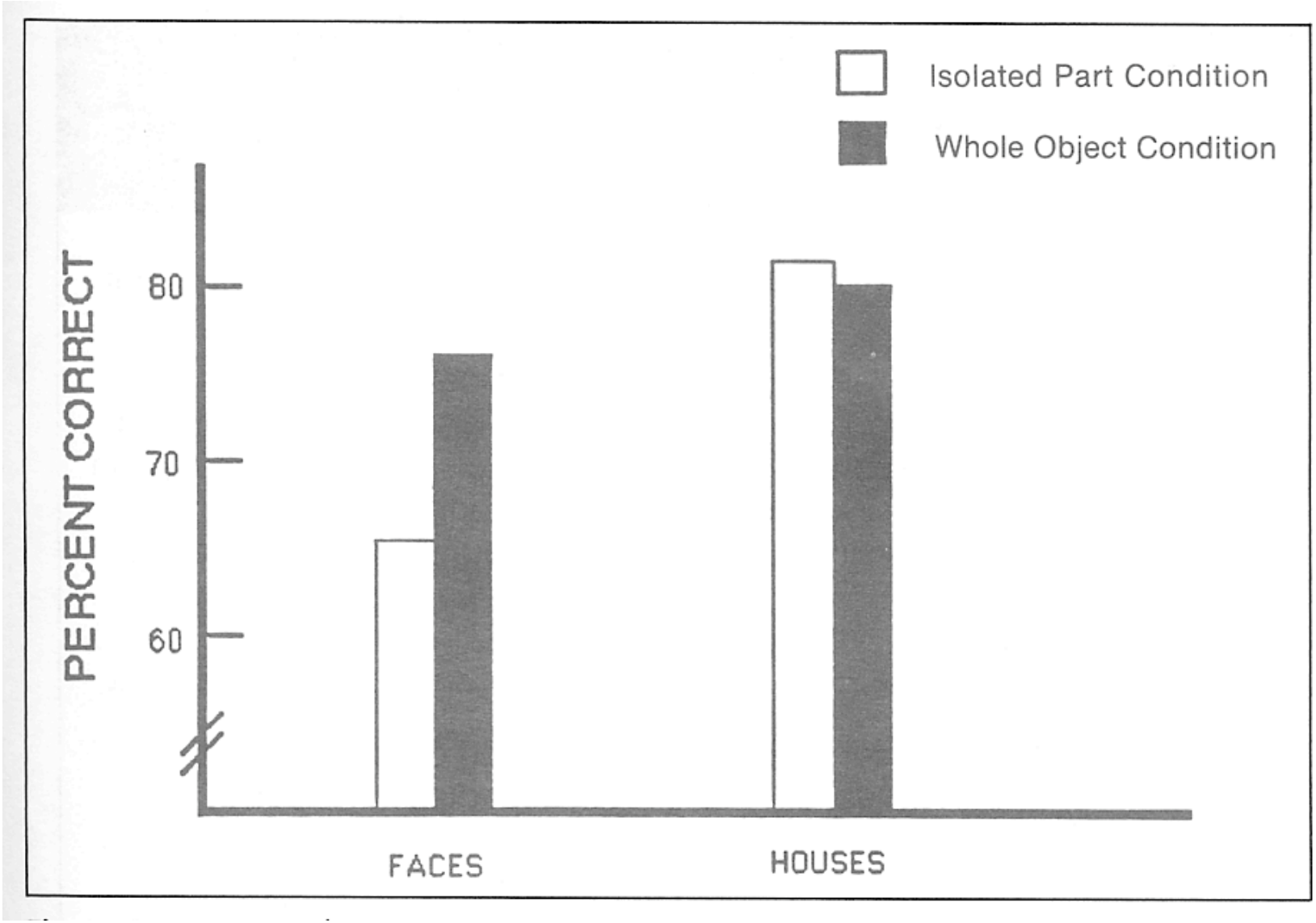


Fig. 5. Examples of pairs of test items from an experiment on the recognition of faces and houses. Subjects studied whole items individually and learned to identify them by name (e.g., Larry's face or Larry's house). The test was administered in a two-alternative forced-choice format, either for an isolated part (e.g., "Which is Larry's nose?" or "Which is Larry's door?") or for the whole item with only a single part changed (e.g., "Which is Larry's face?" or "Which is Larry's house?").

Part in whole judgment is much easier than part judgment for faces

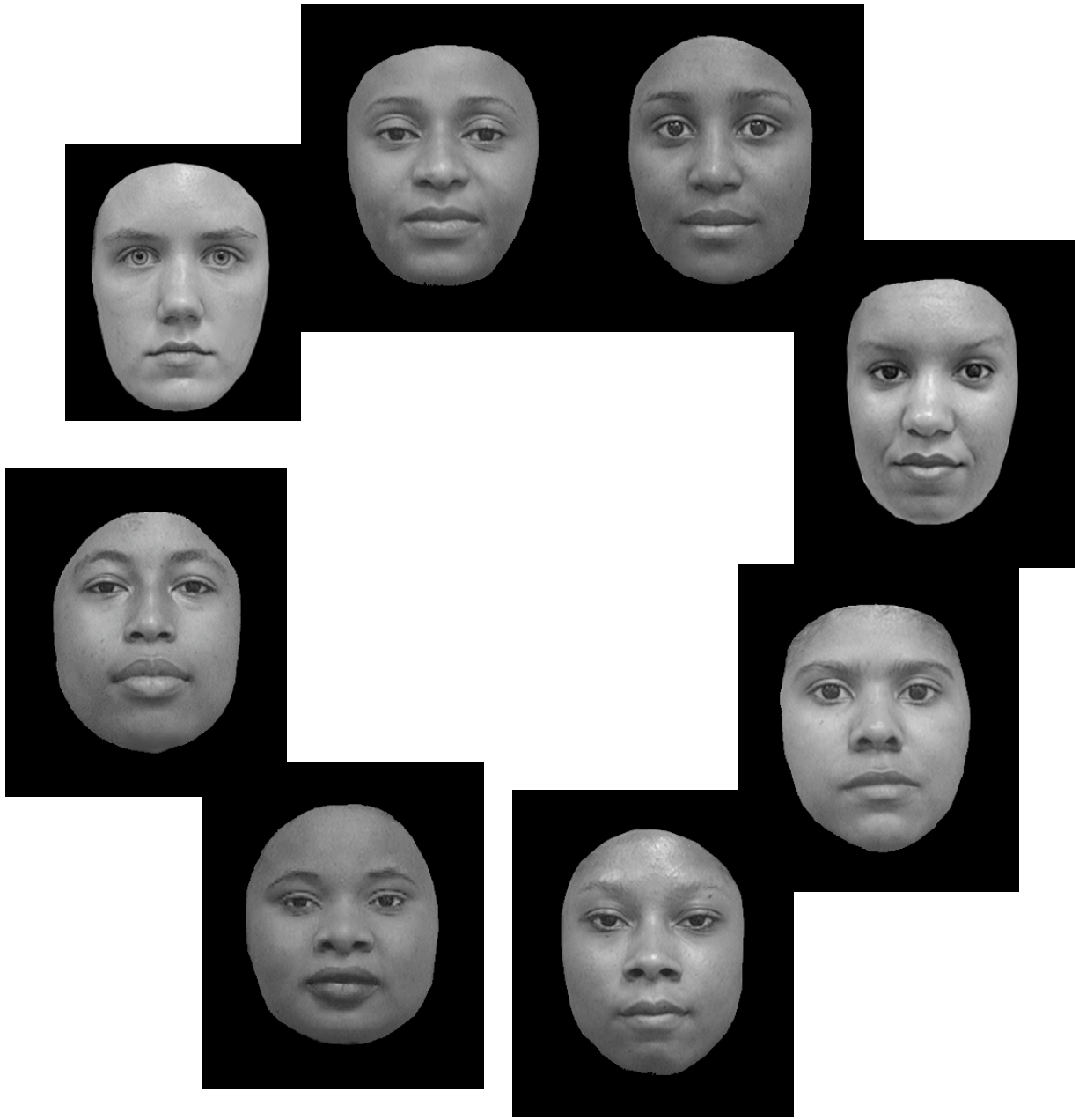
Faces are holistically perceived

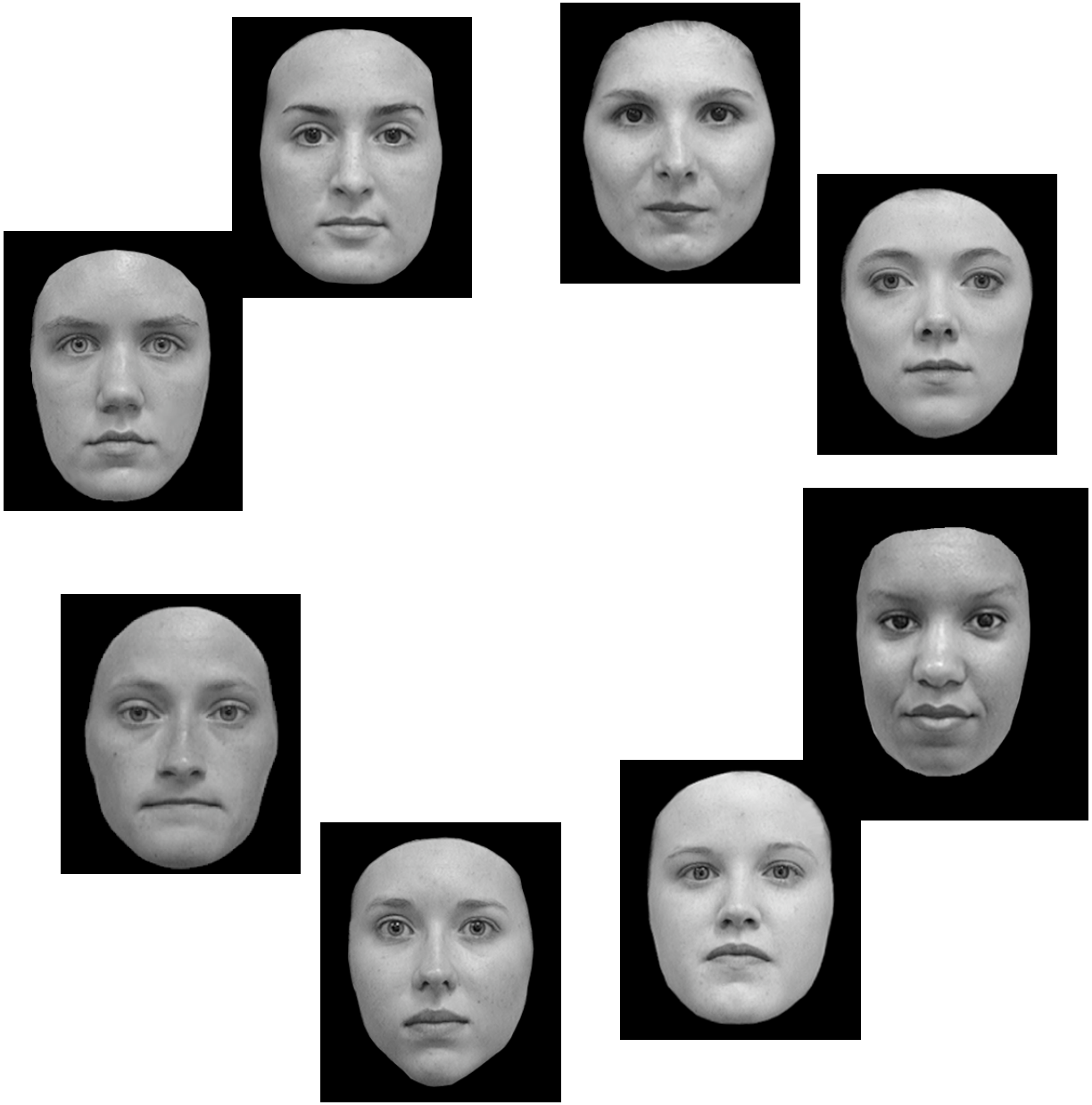
Tanaka & Farah (1993)



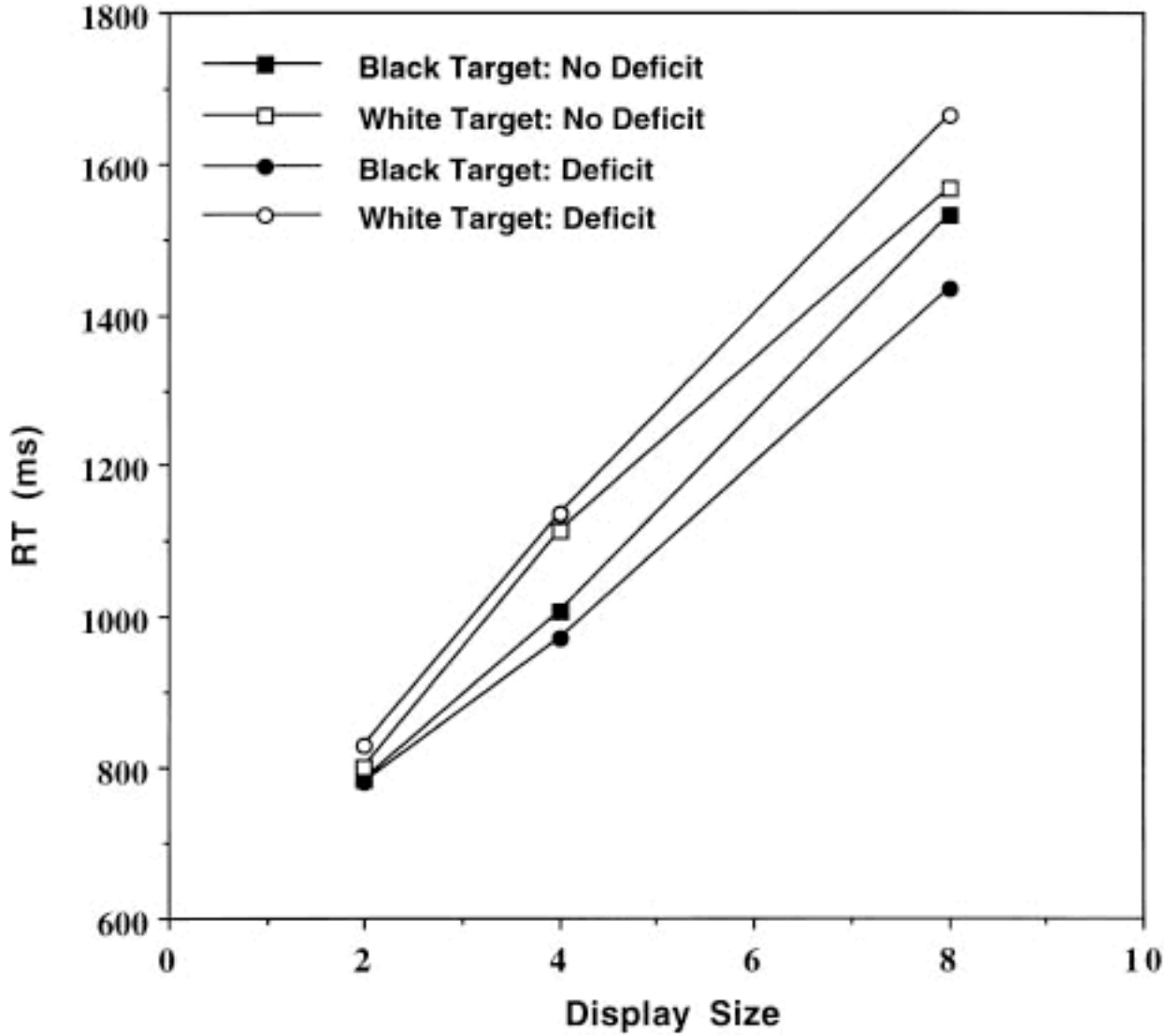
Feature Search for Race

The influence of life-long experience on face perception





Feature search task performance is more efficient with “other race” targets than “same race” targets (Levin, 2000)



Conclusions

- There is strong evidence for an analytic account of pattern recognition
  - Parallel detection of features
  - Combining features together takes time and attention
  - Early perception of an object is as a “bag of features”
- Holistic perception also occurs
  - Perception of entire forms without decomposition
  - Unitization of well-learned forms
  - Context effects on perception of features
- Open Question
  - When do analytic and holistic perceptual processes occur?
  - What does it mean to be a psychological feature?