

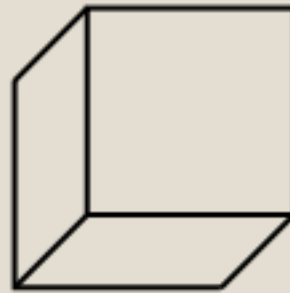
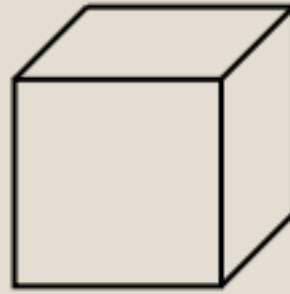
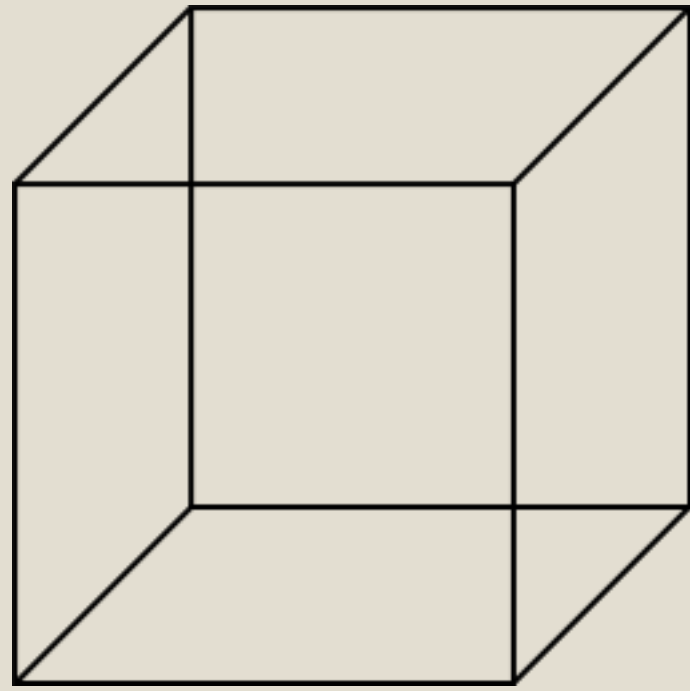


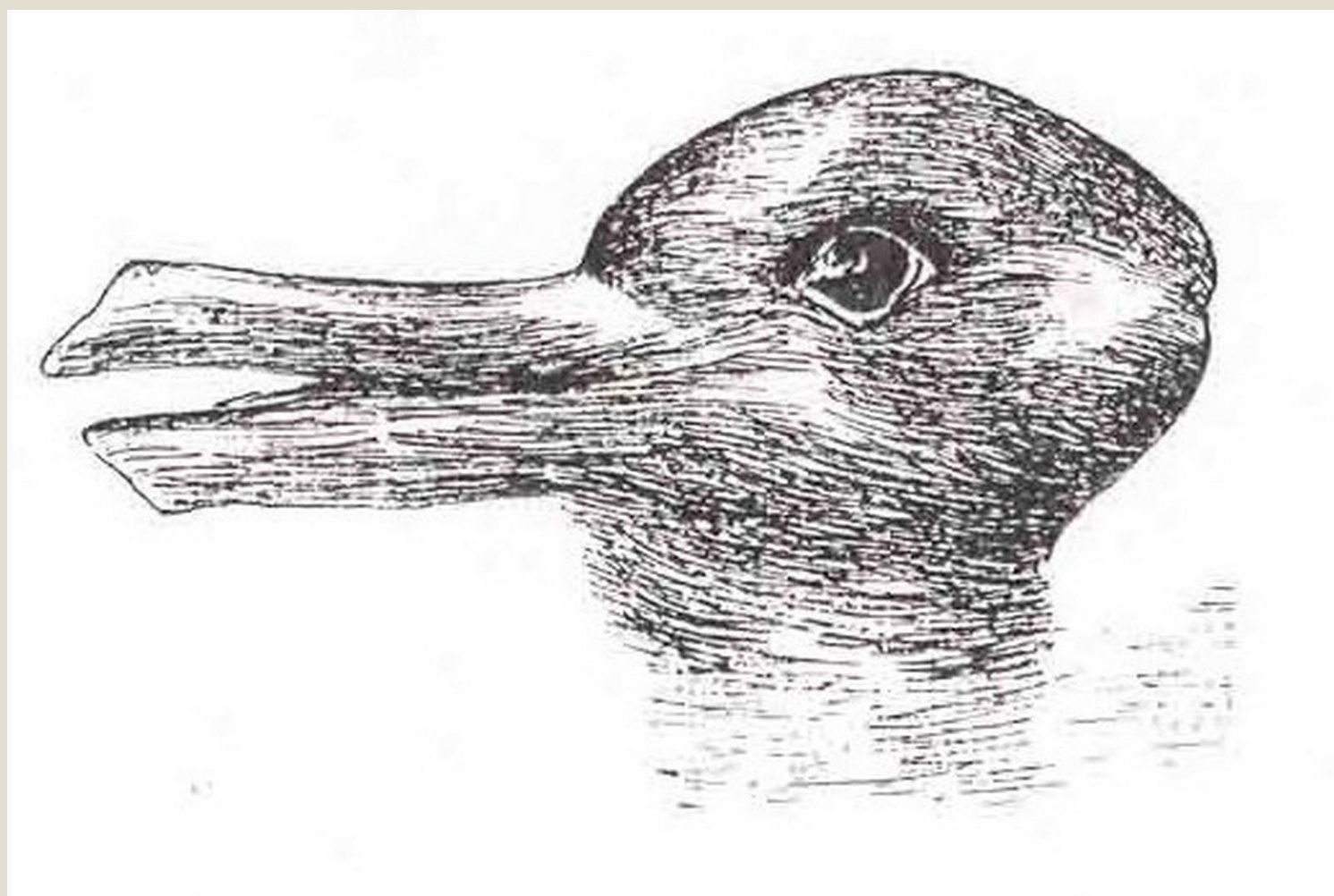
OBJECT RECOGNITION

PSY435

Roadmap

- How is recognition different from perception?
- What are some theories of object recognition?
- How does Gestalt Psychology explain object perception?
- How do computers recognize objects?







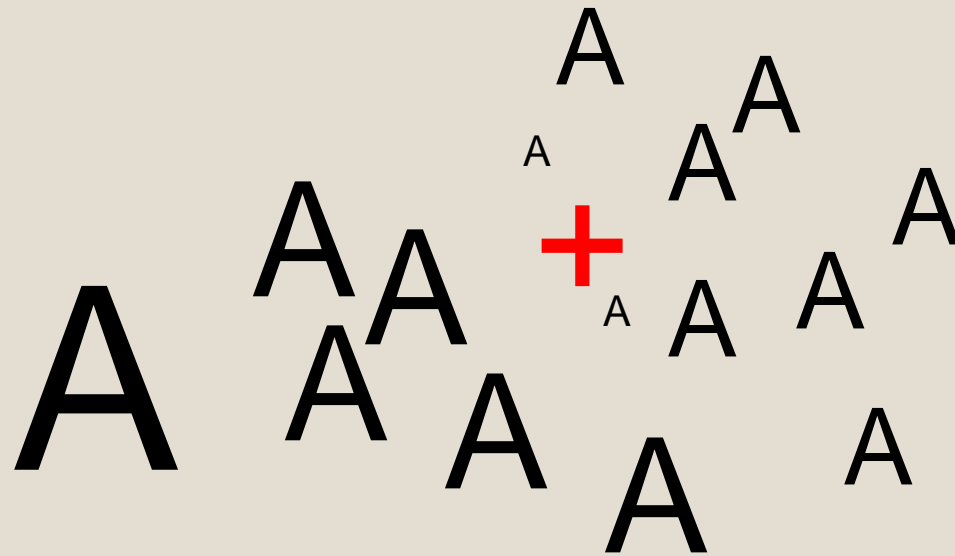
Why Object Recognition is Hard

1. Object Boundaries not easily determined



The visual system must carry out the process of “image segmentation”.

2. Objects may appear anywhere on retina
in any size

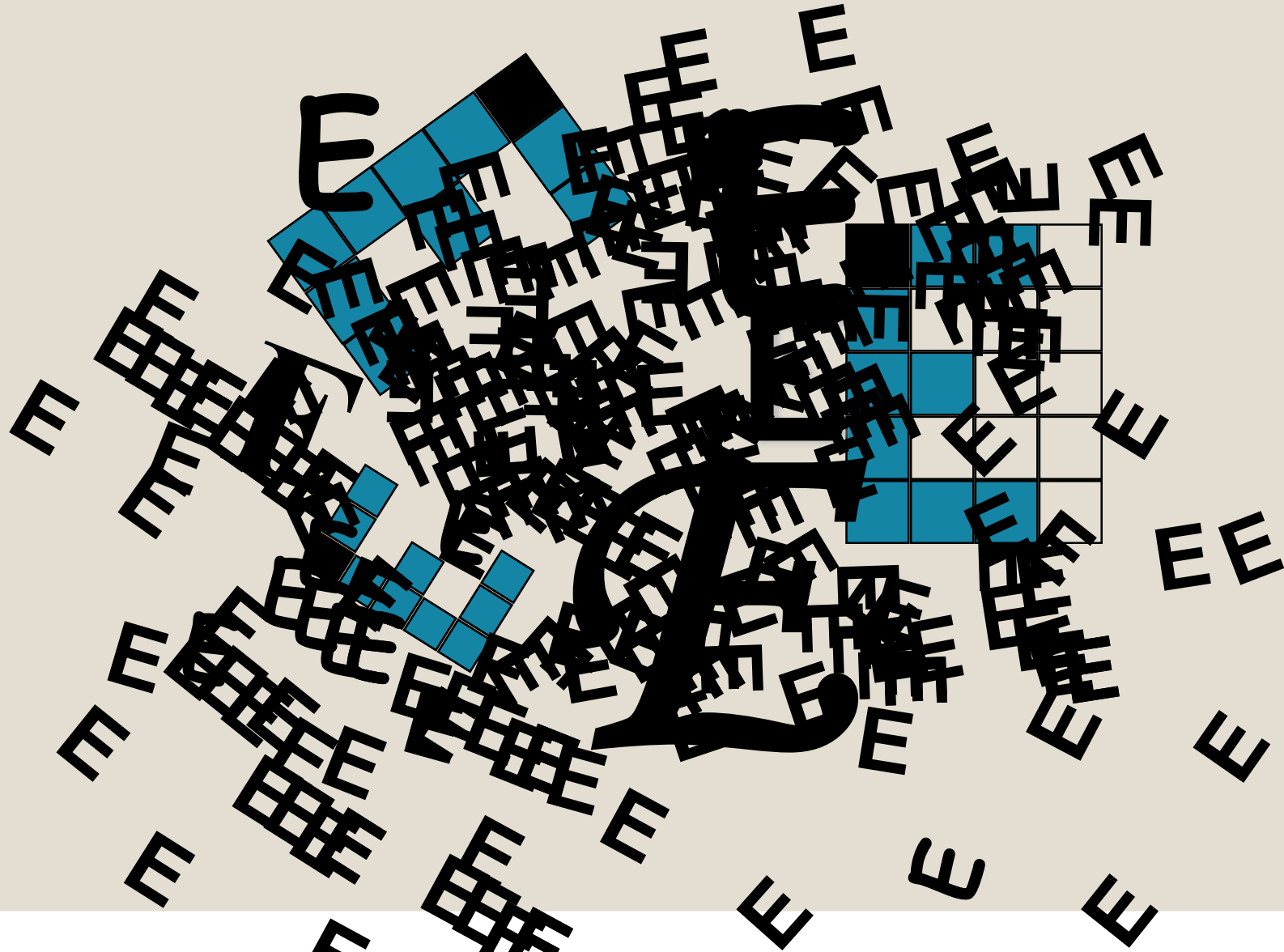


Perhaps the brain is able to represent these objects in a way that is “translationally invariant” and “size invariant”.

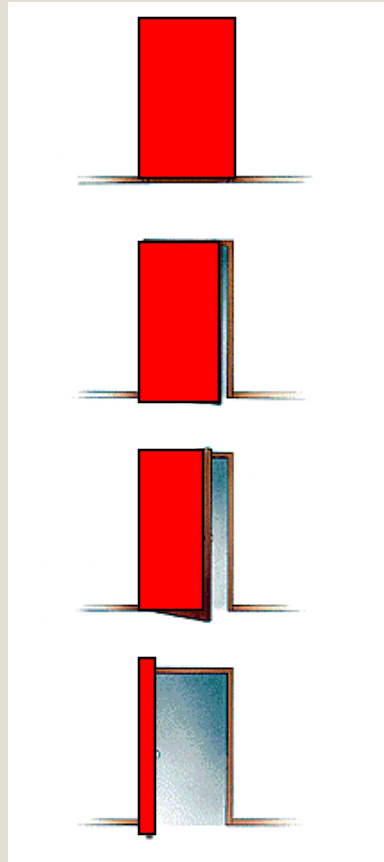
3. Objects may occlude each other or they may be superimposed on each other.



4. Same object may have different shapes



Perceptual constancy: Shape



Visual System Infers:

1. Object Identities
(man, woman, child, oar, lake)
2. Spatial Layout of
Scene
3. Properties of Surfaces
(color, texture, etc.)

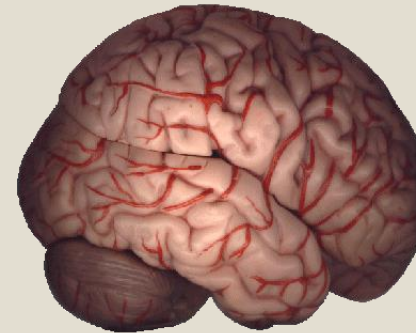
Specific Disorders suggest functionally independent types of visual analysis

1. Visual Agnosia:

- loss of ability to recognize objects
- preserved ability to navigate, reach
- Bilateral damage to occipital/temporal
 - Prosopagnosia (specific to face recognition)

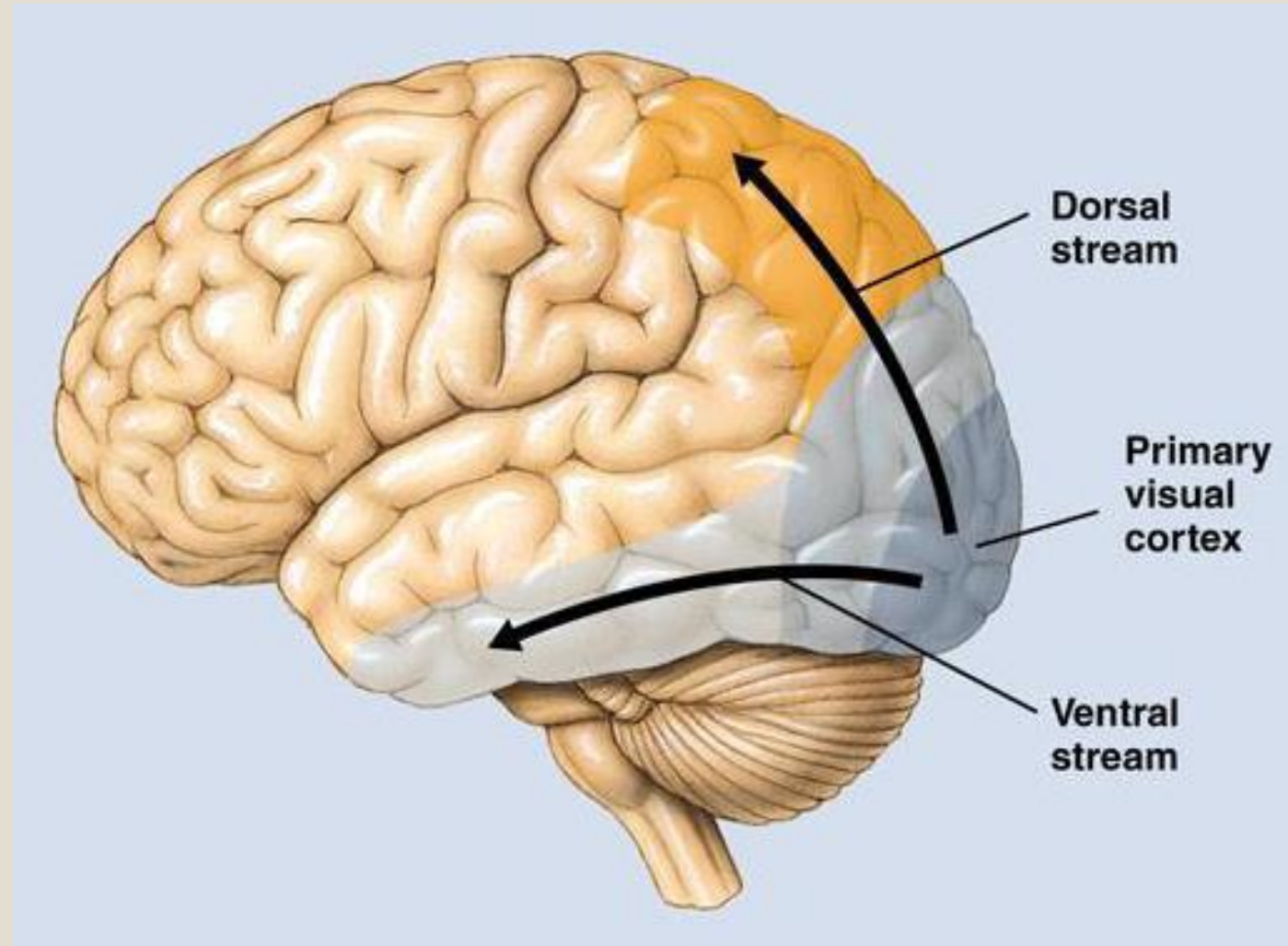
2. Balint's Syndrome:

- inability to navigate, reach
- preserved ability to recognize objects
- damage to parietal areas



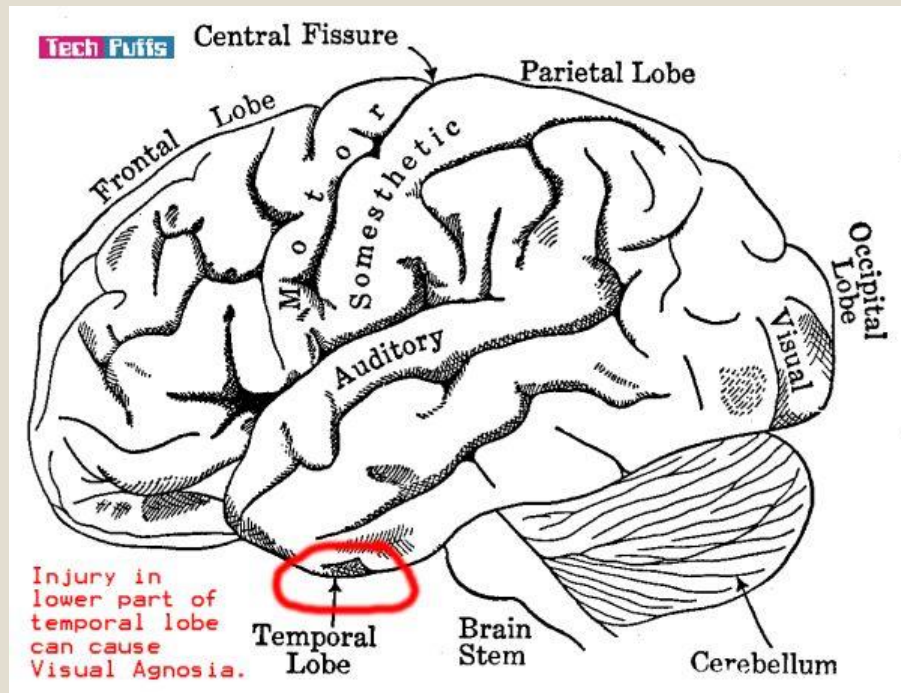
3. Cerebral Achromatopsia

- inability to discriminate colors
- **distinct from color blindness**--Color blindness results from abnormalities in the photoreceptors of the eye. But cerebral achromatopsia results from damage to posterior visual areas in the brain



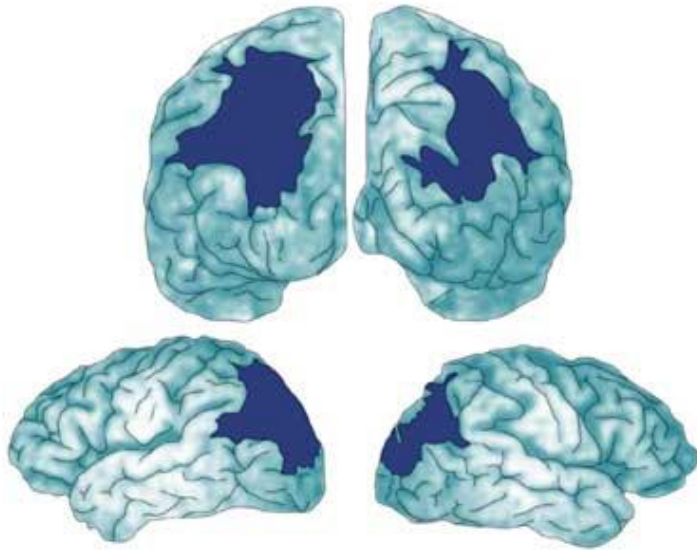
Visual Agnosia

<https://www.youtube.com/watch?v=T1qnPxwalhw>



Balint's Syndrome

- <https://www.youtube.com/watch?v=4odhSq46vtU>



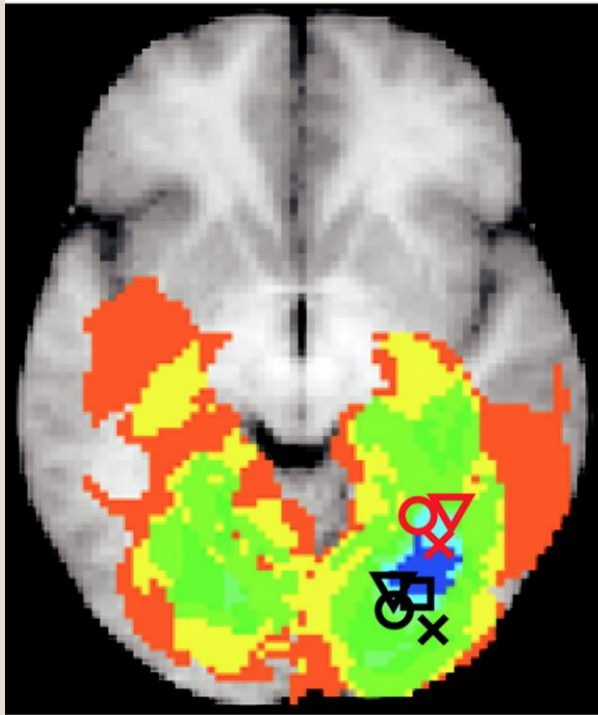
Nature Reviews | Neuroscience



Cerebral Achromatopsia

<https://www.youtube.com/watch?v=ygRNoieAnzI>

Achromatopsia



Some theories of Object Recognition

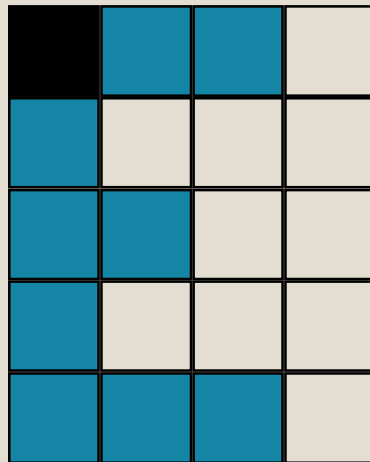
- Older Theories
 - Template
 - Feature
 - Recognition by Components
- Current Theories
 - Interactive activation
 - Neural network models
 - Top-down bottom up at all times
 - Context/Relative
 - Meaning stored relatively

Template Theory

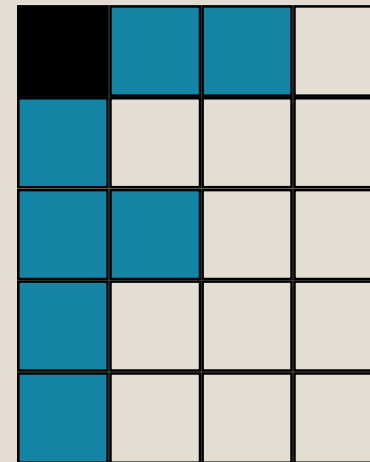
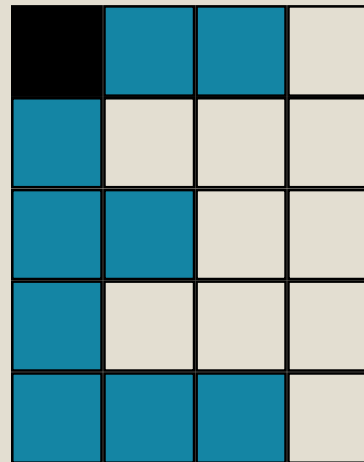
- Store in brain a copy of what every possible input will look like.
- Match observed object to the proper image in memory

Template Theory

Perceptual Representation



Memory Representations



Can we think of issues?

- (1) **Massive numbers** of templates are required
- (2) Predicts no **transfer** to novel views of the same object
- (3) Objects are often obstructed
- (4) Doesn't explain how we recognize novel objects

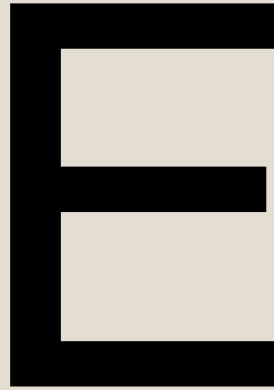
Feature Analysis

- Object is built up from features
- Fixed set of elementary properties analyzed independently in parallel across visual field.

A Simple Version of Feature Theory

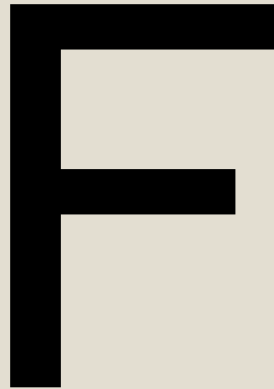
Perceptual Representation

3 Horizontal lines
1 Vertical line
4 Right angles



Memory Representation

3 Horizontal lines
1 Vertical line
4 Right angles



2 Horizontal lines
1 Vertical line
3 Right angles

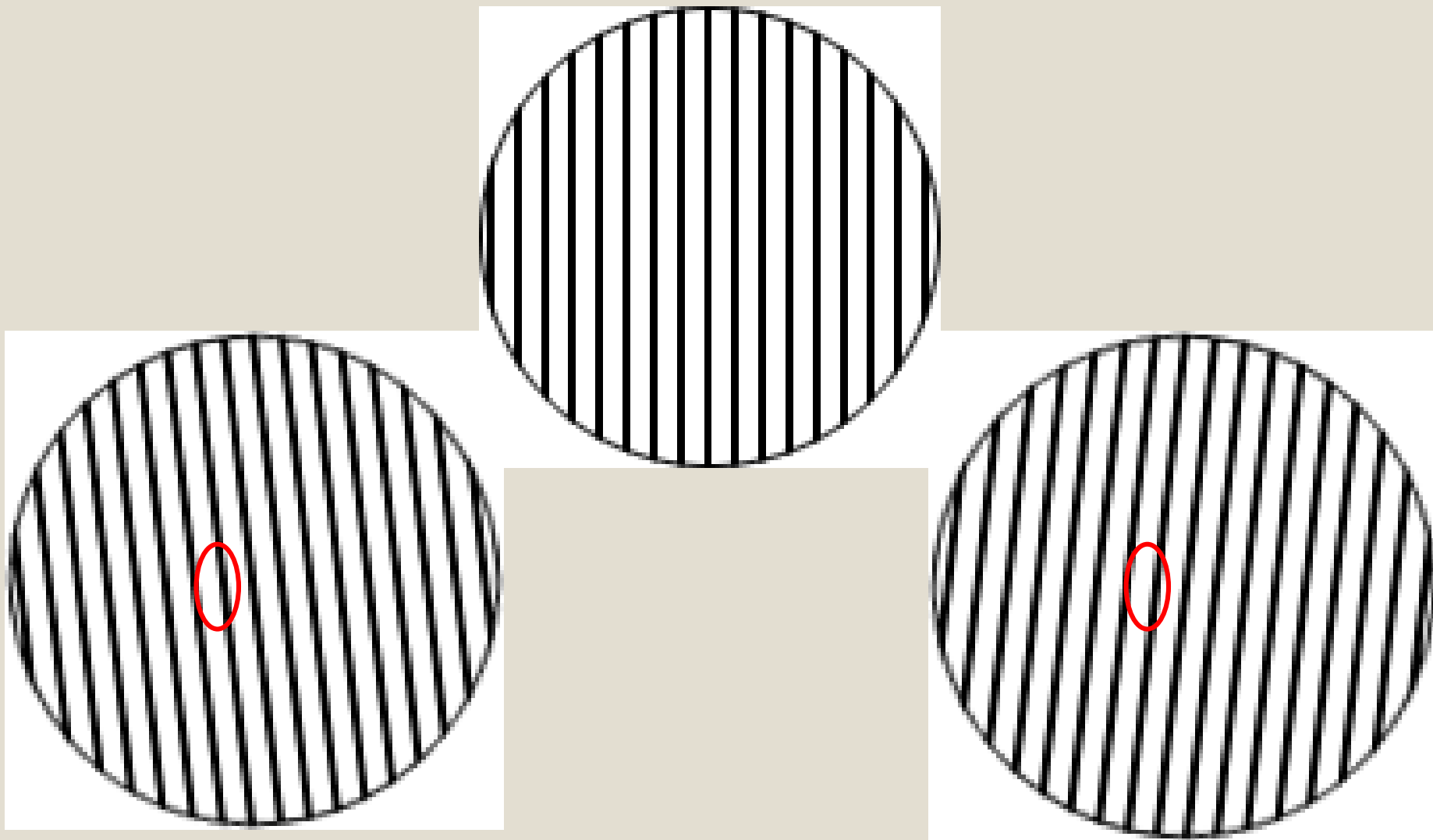
Evidence that Features are really basic elements of visual processing

- Physiological Evidence
 - Individual neurons can be shown to respond preferentially to different kinds of simple visual features
 - **simple cells**--respond best to lines or angles of a specific orientation and retinal position
 - **complex cells**--fire maximally to lines or angles without respect to location. They often also have preference for stimuli moving in certain directions...
 - **hypercomplex cells**--fire maximally have even more complicated sets of requirements for maximal firing. (e.g. corners, notches....etc.)

Adaptation Effects are indicators of elementary visual features...

- Your eyes are always moving, even when “fixated” (microsaccades).
- But if an image is kept completely still on the retina, it will slowly fade from view!
- We believe this happens because individual feature detectors become habituated, or fatigued with prolonged, sustained stimulation.
- Retinal stabilization procedure allows direct demonstration of this phenomenon.

Adaptation to Orientation



More Evidence for Features...

Visual Search:

(1) Subject scan through groups of stimuli searching for a specific “target stimulus.”

(2) Targets are *easier* to detect when they can be *defined by a single feature*.

(3) Targets are *harder* to detect when they are defined on the basis of *combination of basic visual features*.

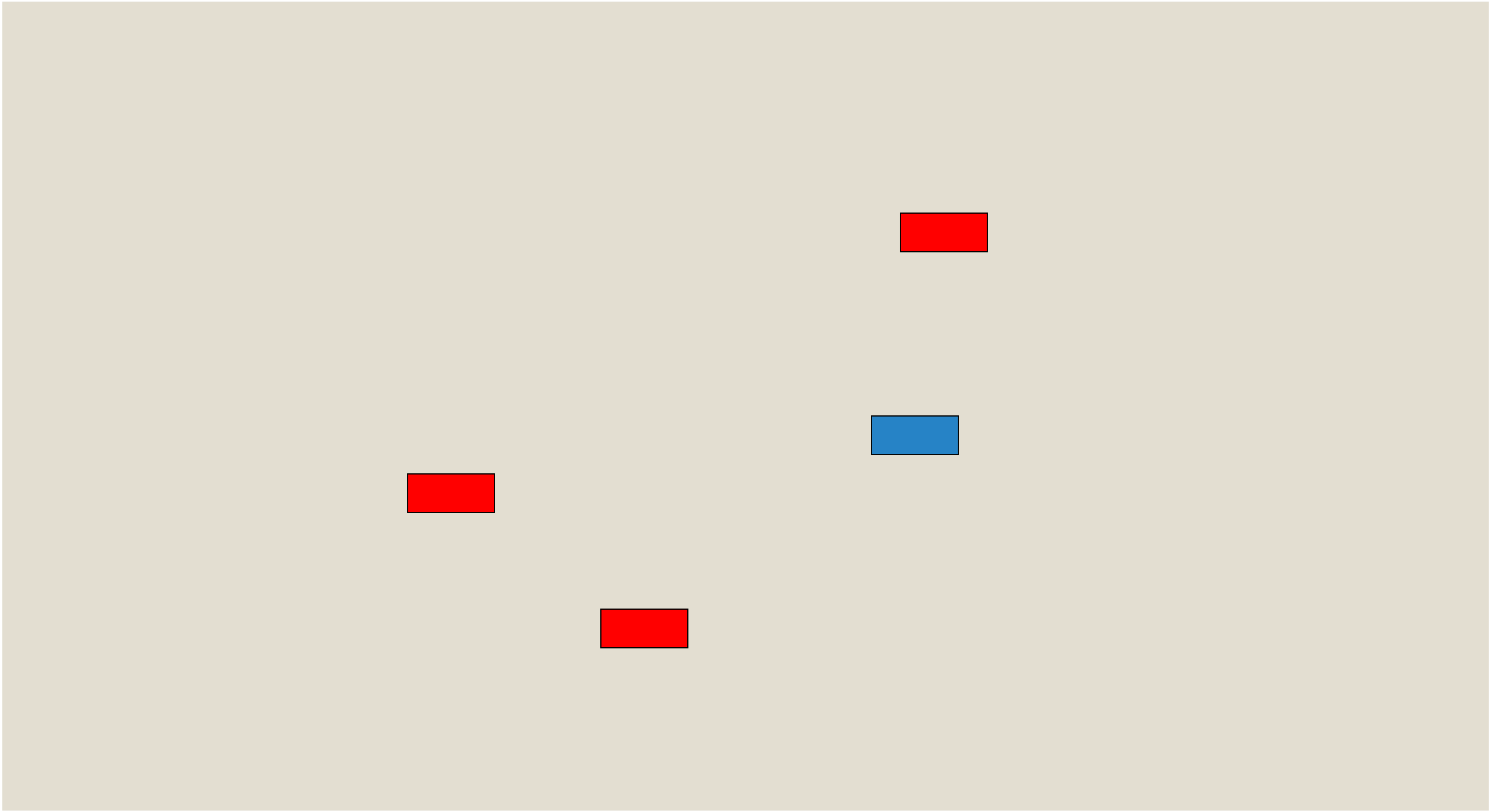
(4) This result suggests that the easily detected features are *elementary units of visual perception*.

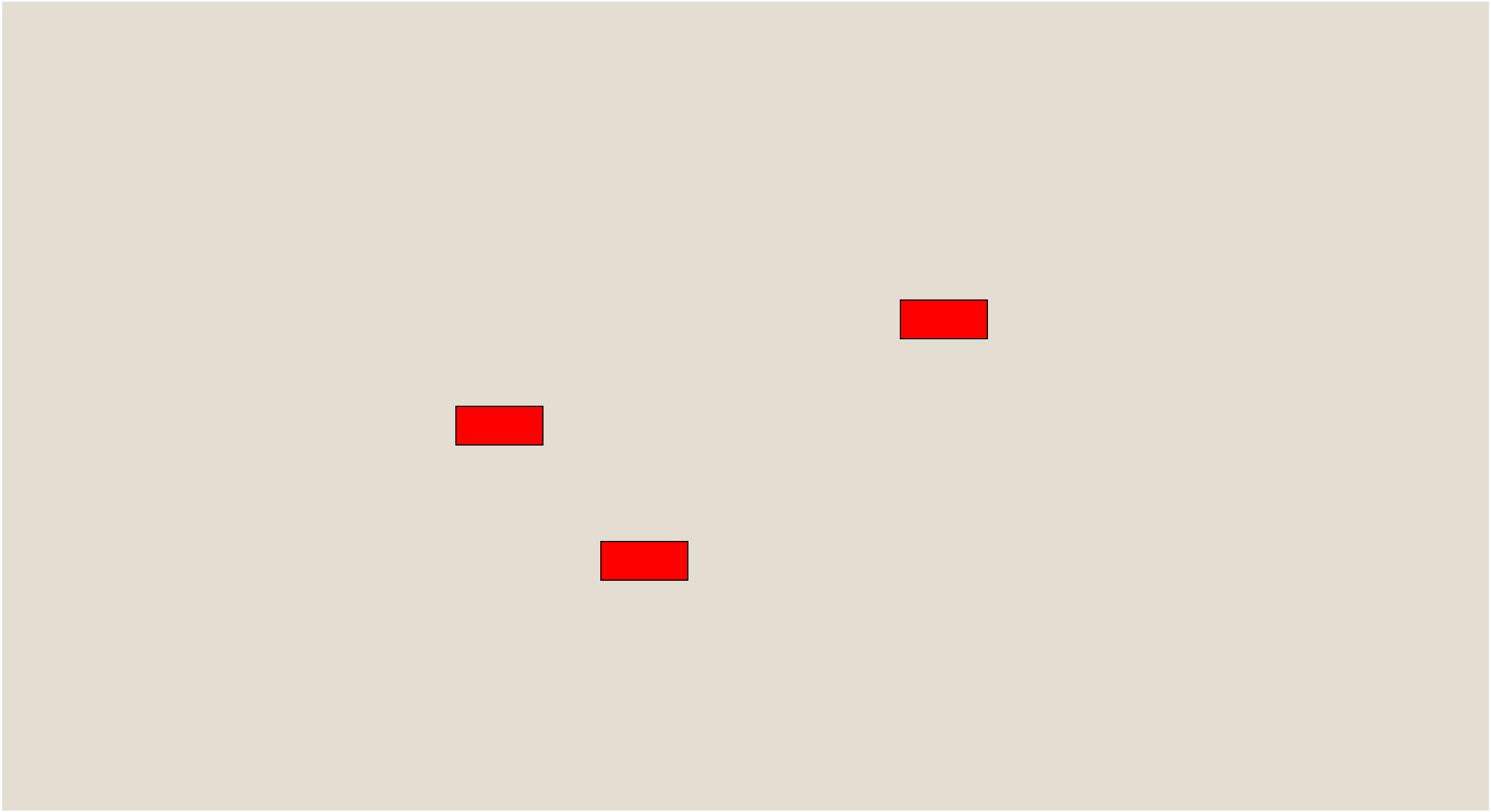
Let's try it out ourselves...

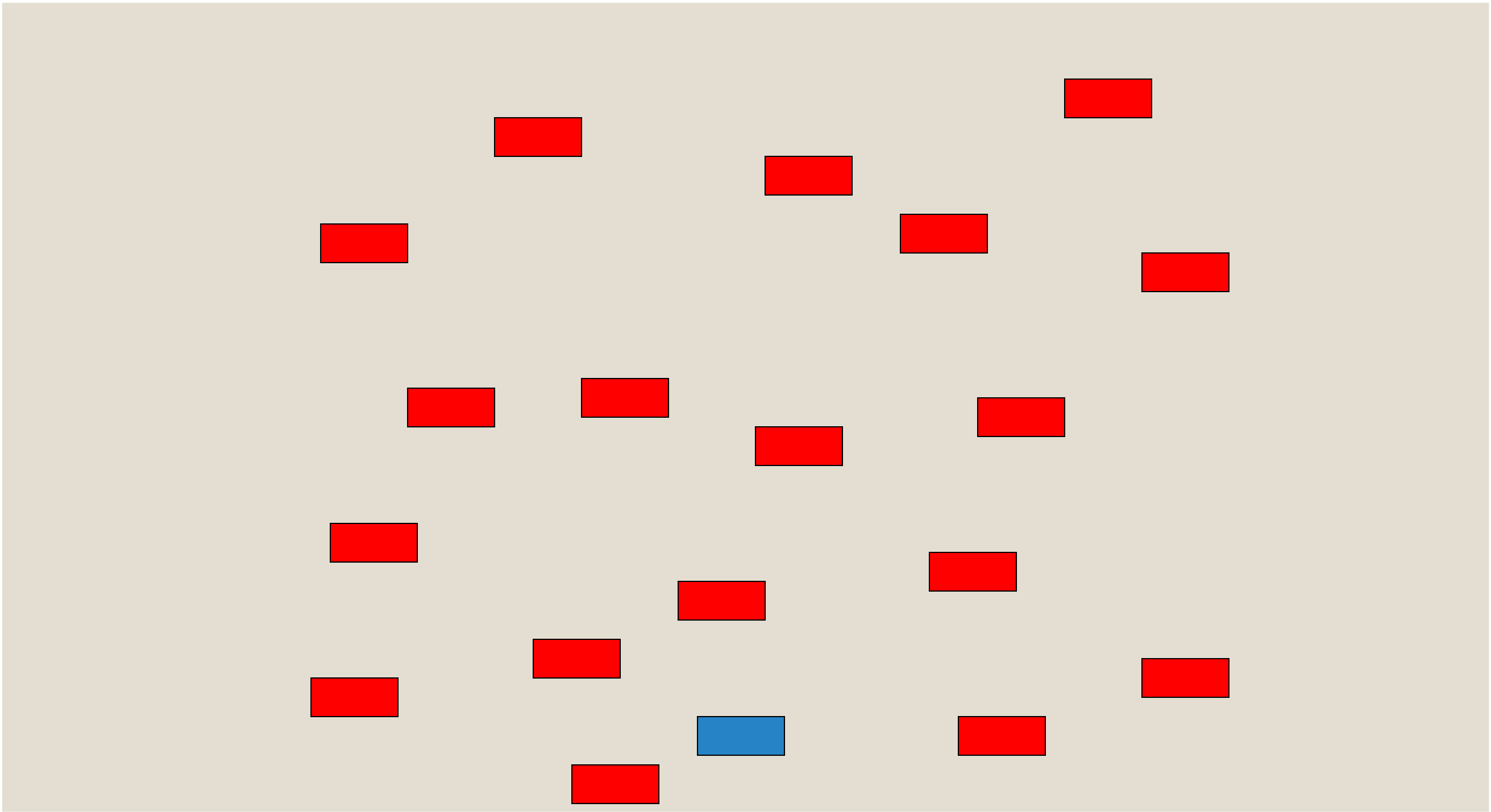
“FEATURE SEARCH”

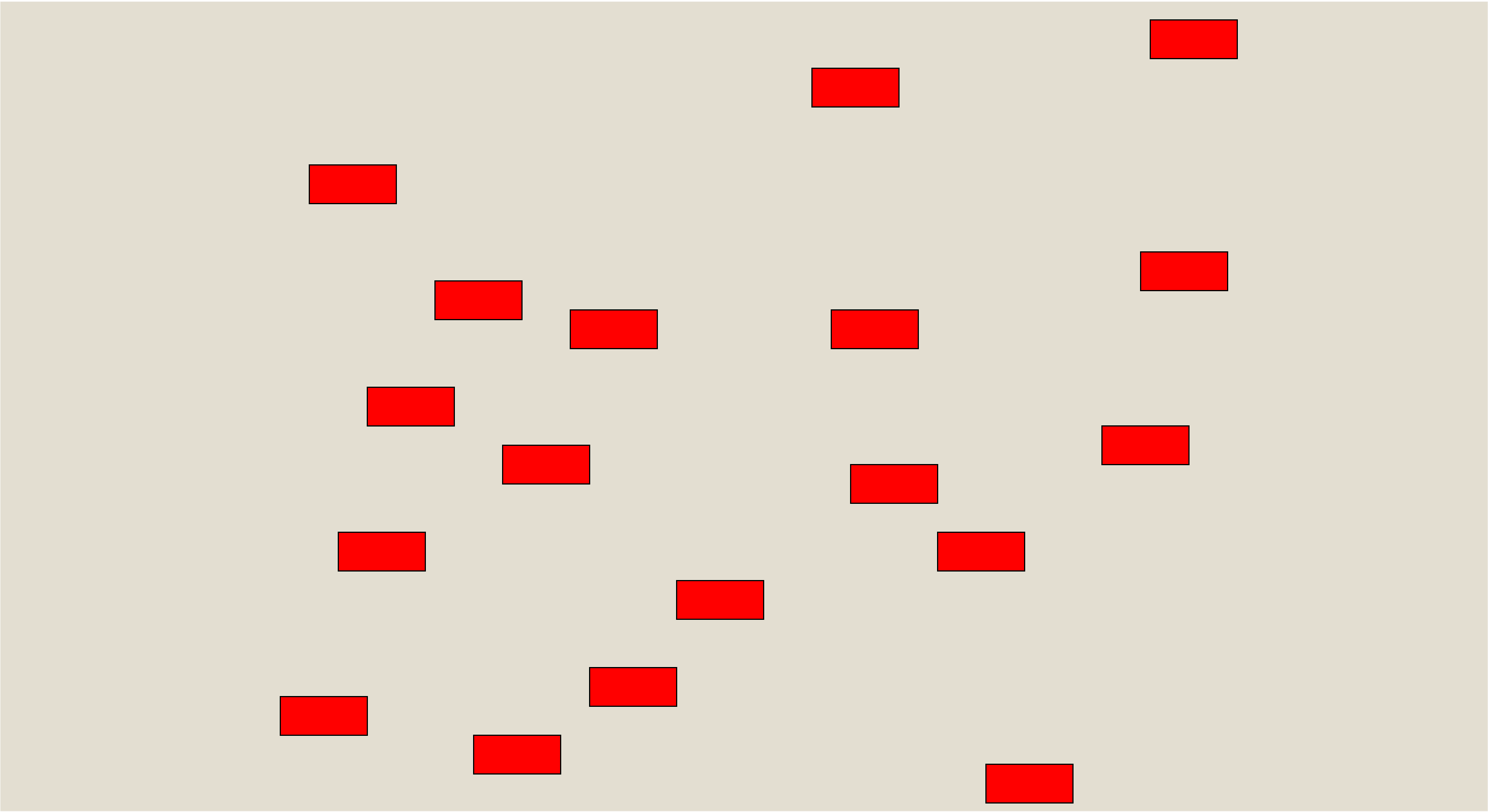
Clap when you see the bar that is blue

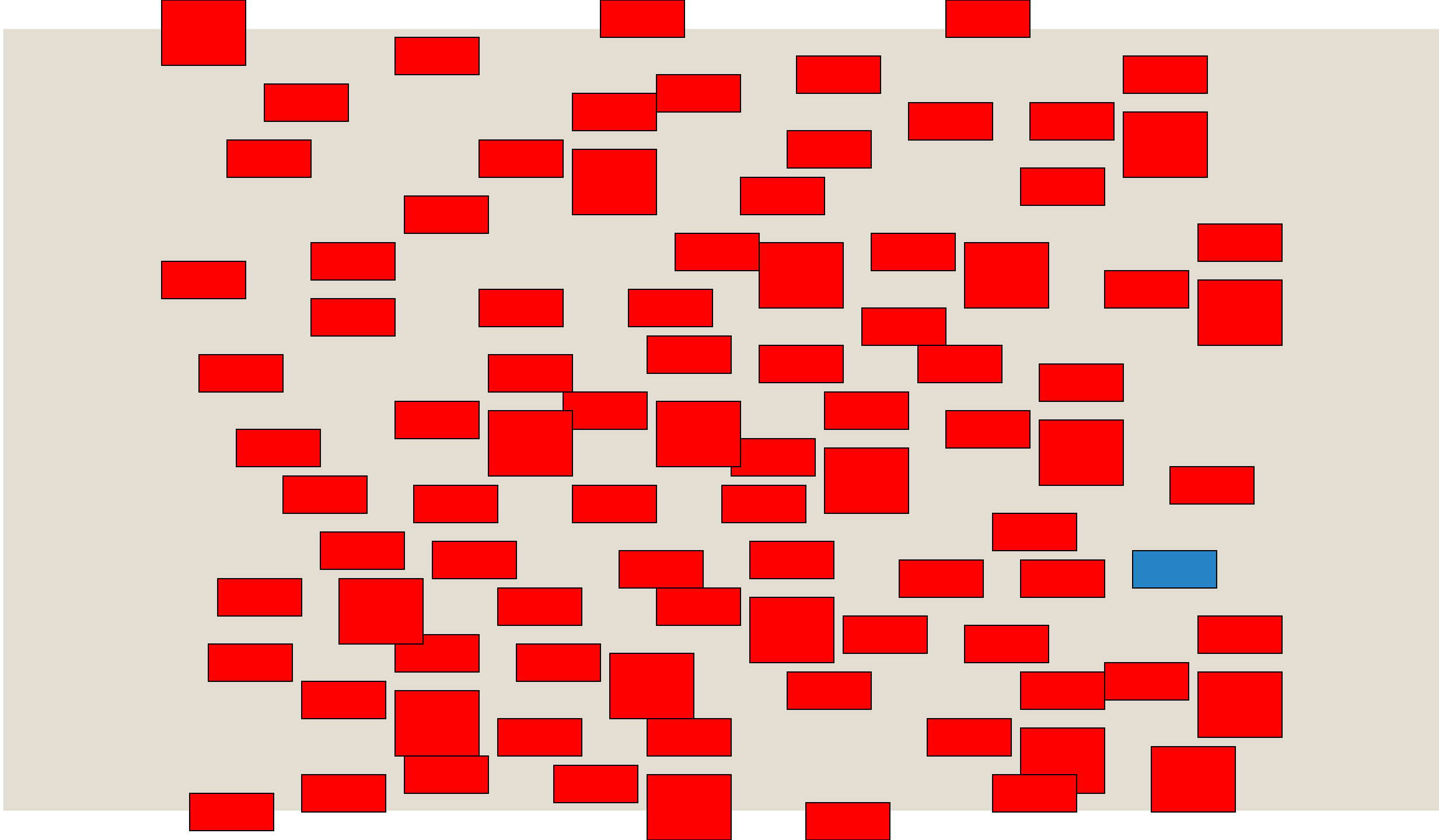




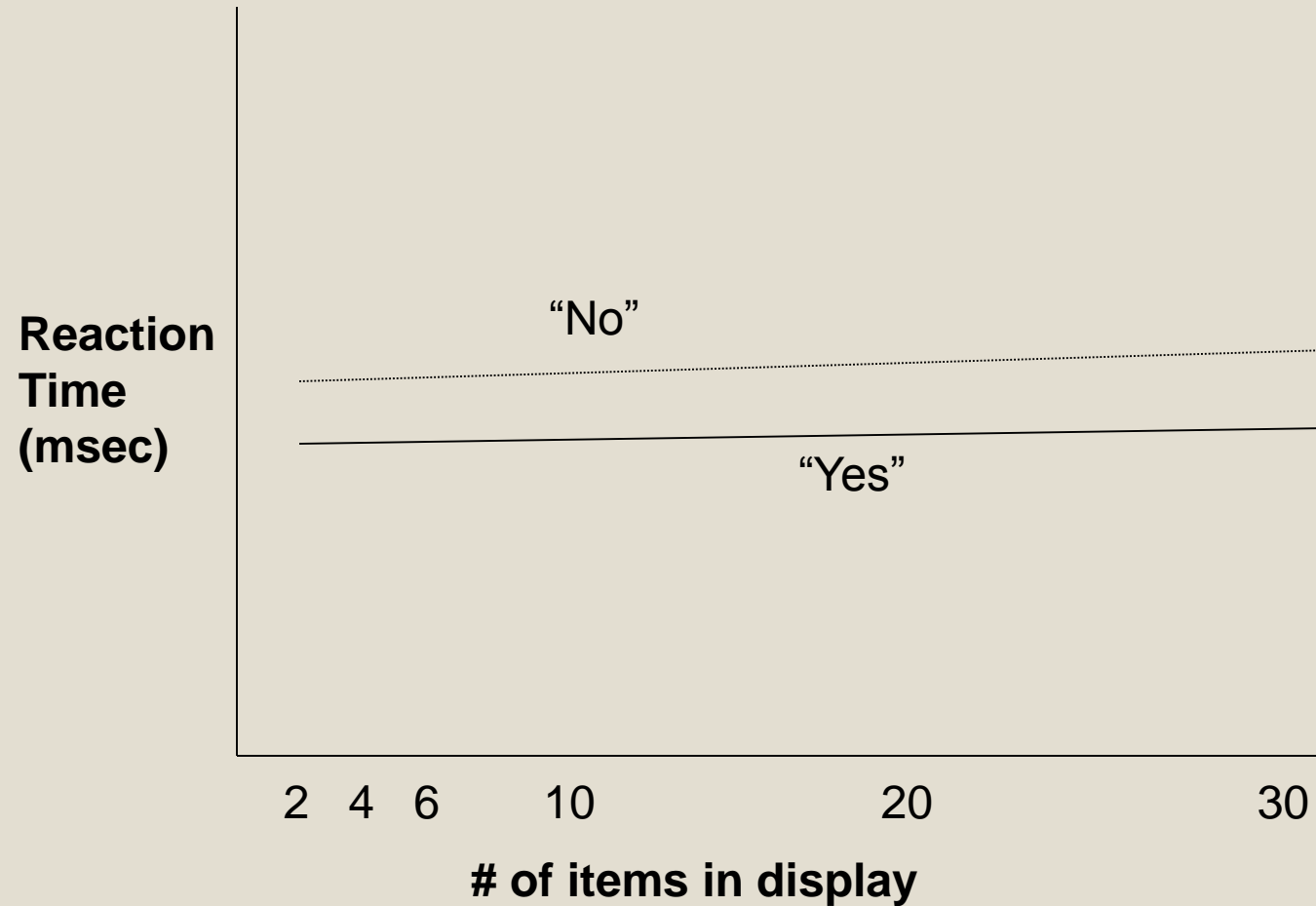








Typical Results for “Feature Search”



These results suggest parallel analysis and detection of simple visual features.

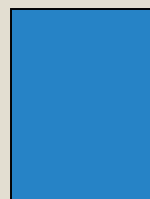
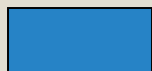
Conjunction Search

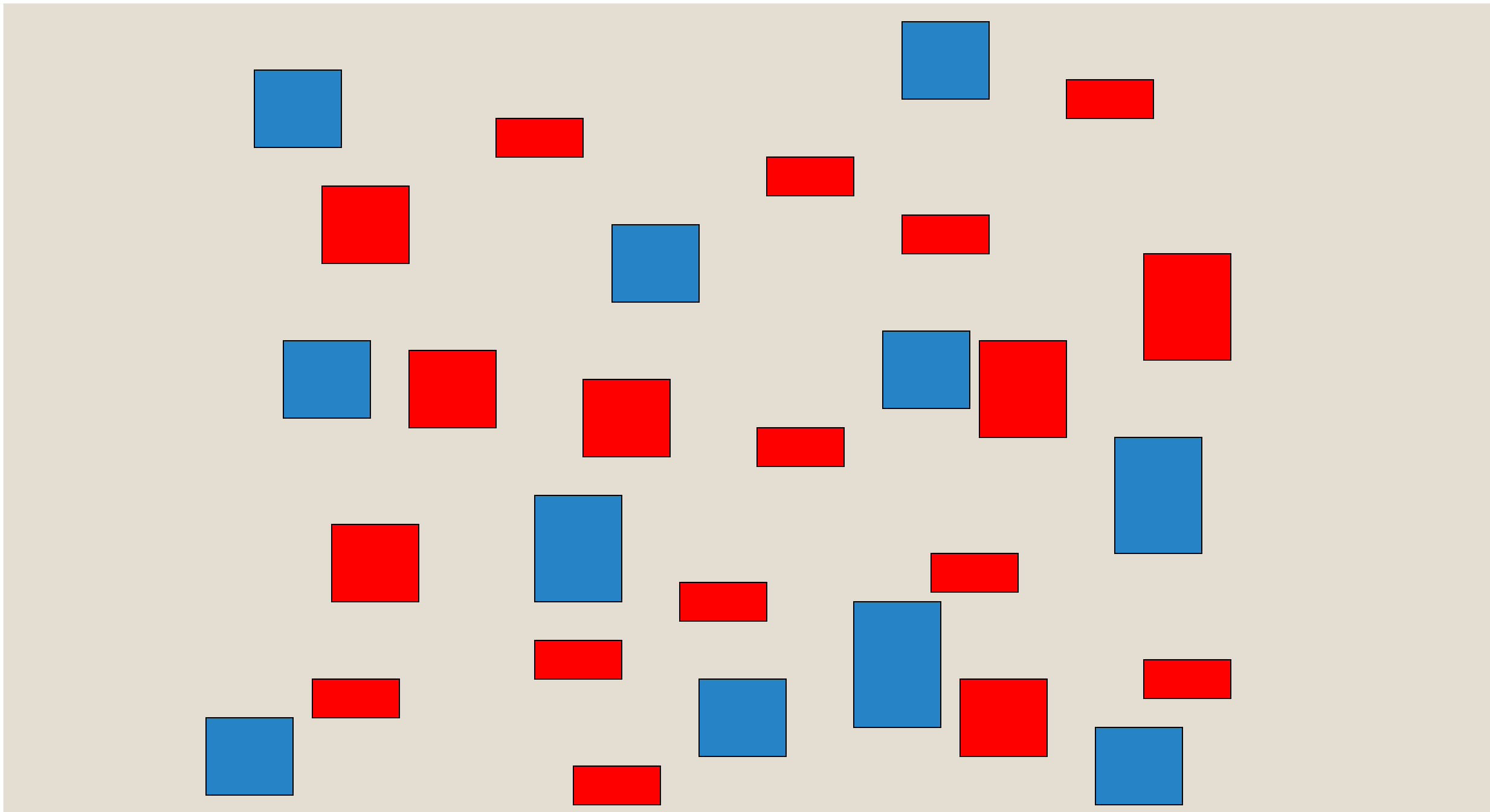
When targets are defined by:

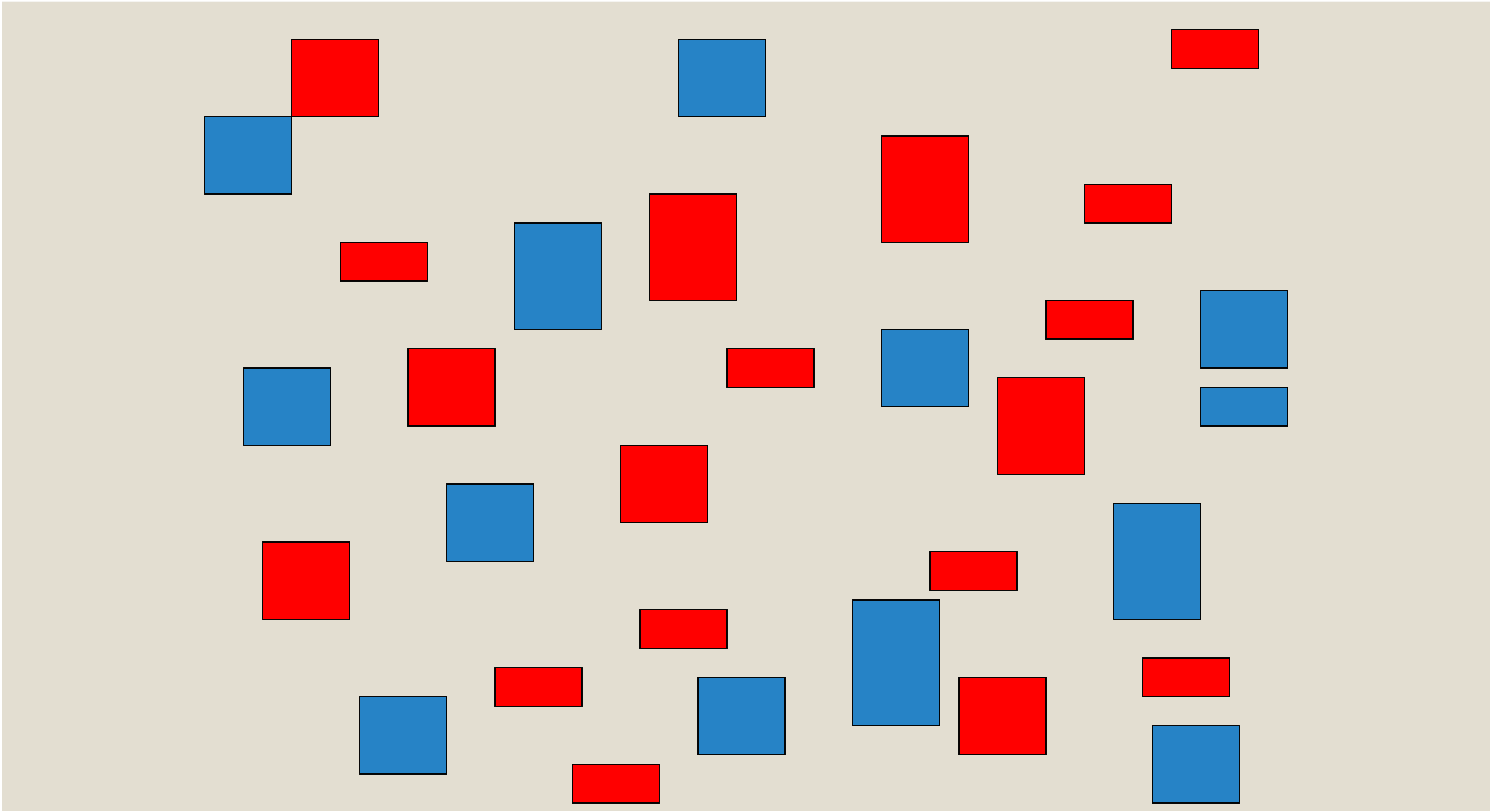
- Combination of features (e.g., red AND horizontal)
- Spatial arrangements of features (e.g. black above white)

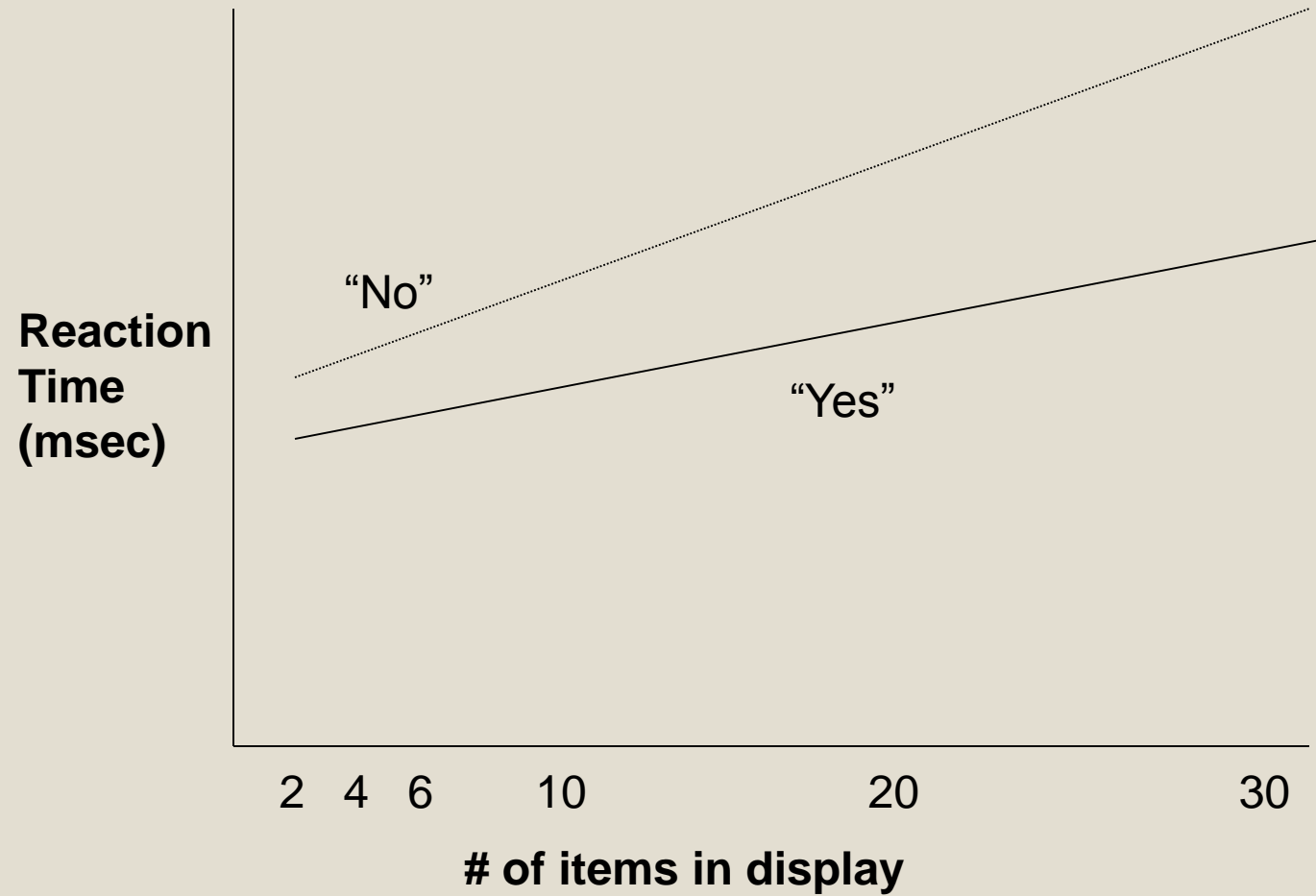
“CONJUNCTION SEARCH”

Clap when you see the bar that
is both horizontal AND blue









When higher order analysis or integration of multiple features is required, search is much harder, and reaction time rises with number of distractors.

Feature Theory: a possible problem

Absence of a feature:

O vs. \emptyset

“ABSENCE OF FEATURE Search”

Clap when you see an “O” without a slash through it

O and not Ø

O

Ø

Ø

Ø

Ø

O

Ø

Ø

Ø

Ø

O

Ø

Ø

Ø

Ø

Ø

Ø

Feature Theory: a ~~possible~~ problem

Absence of a feature:

O vs. \emptyset

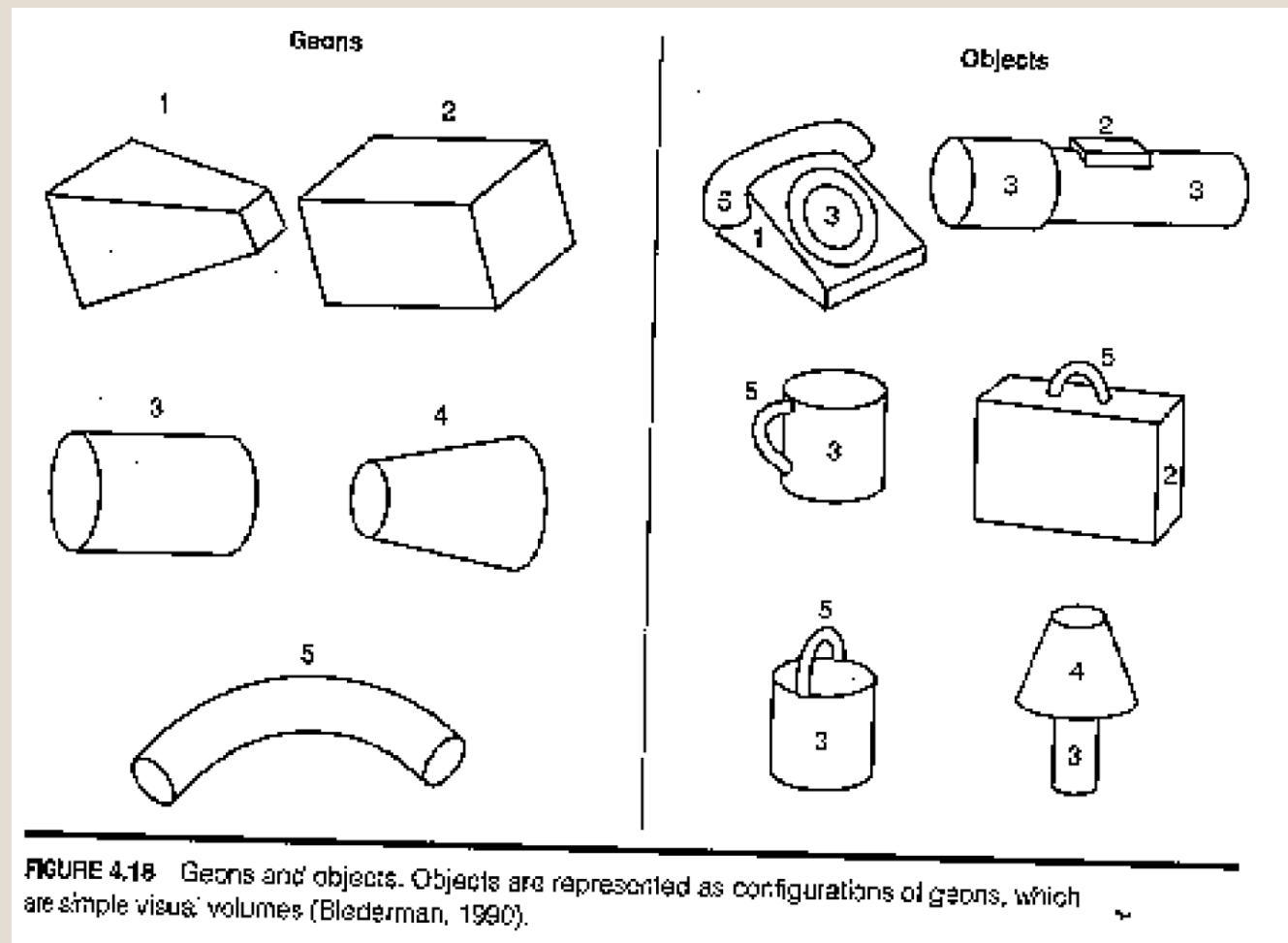
A feature detector, by definition cannot detect the absence of a feature

Issues with Feature Theory?

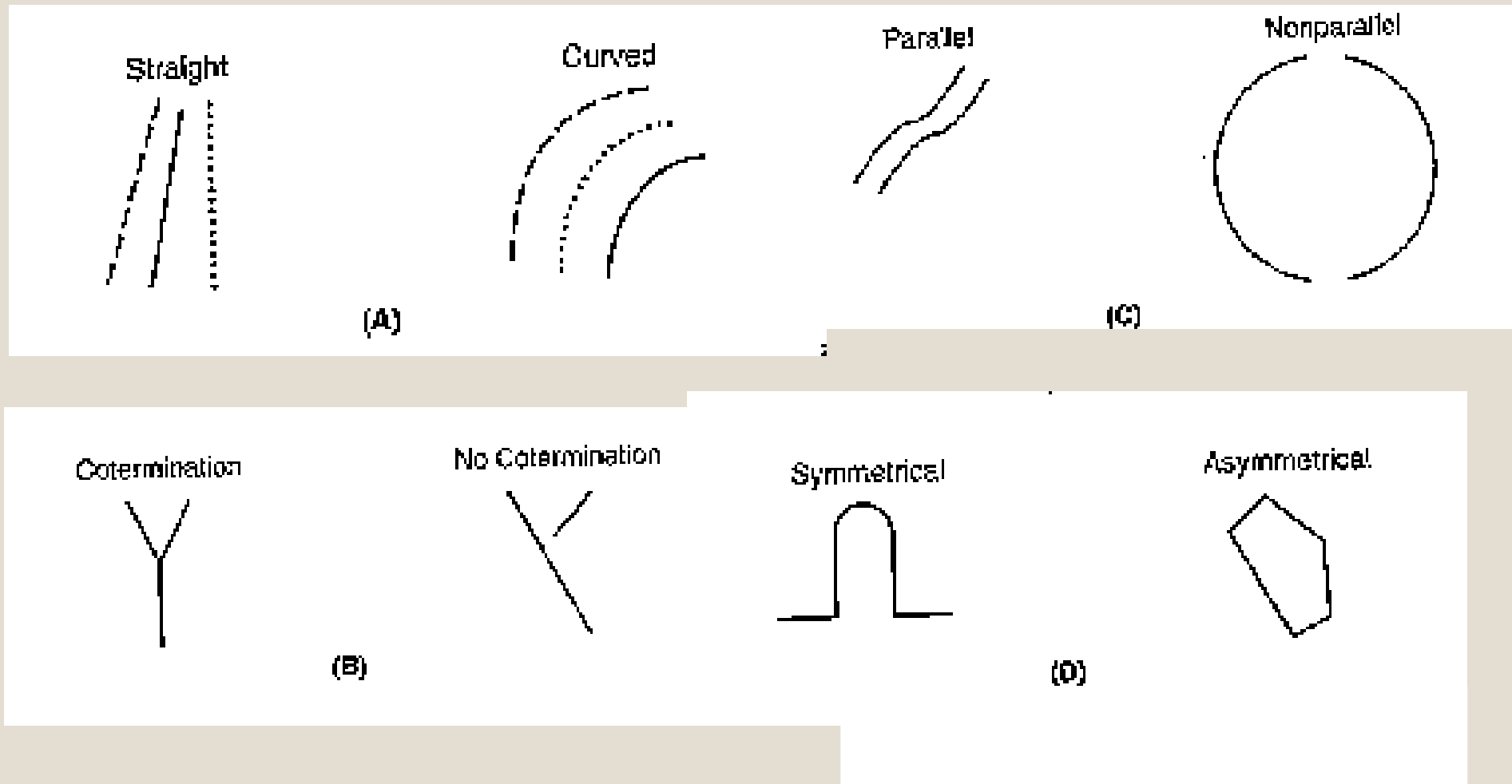
- Pushes the issue down a stage
- We recognize absences as much as presences in features
- Objects appear as unities in the world, not piecemeal constructions

Recognition by Components

- Spatial relationship between feature is taken into account
- Set of “Geons” are proposed that all objects are built from
 - 36 basic shapes
- Describes 3D geometry



Non-Accidental Features



Visual features that are evident regardless of the observer's viewpoint:
Viewpoint invariance

Some Nonaccidental Differences Between a Brick and a Cylinder

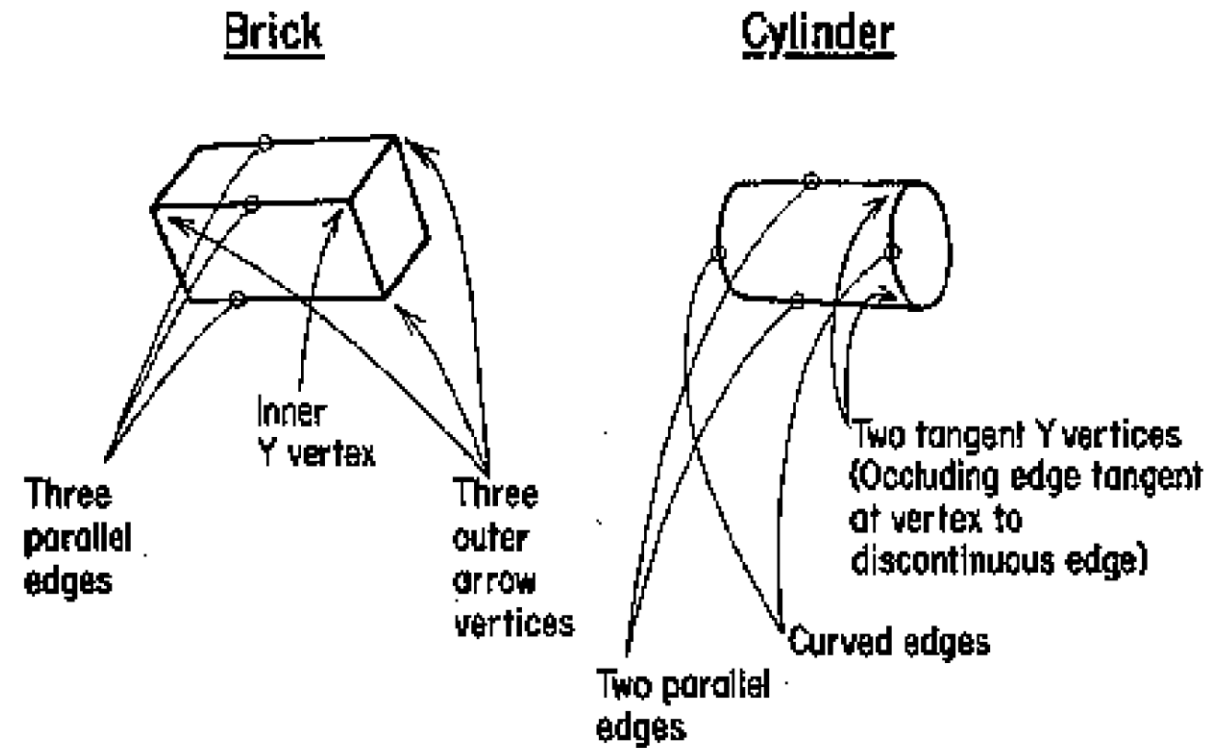


Figure 5. Some differences in nonaccidental properties between a cylinder and a brick.

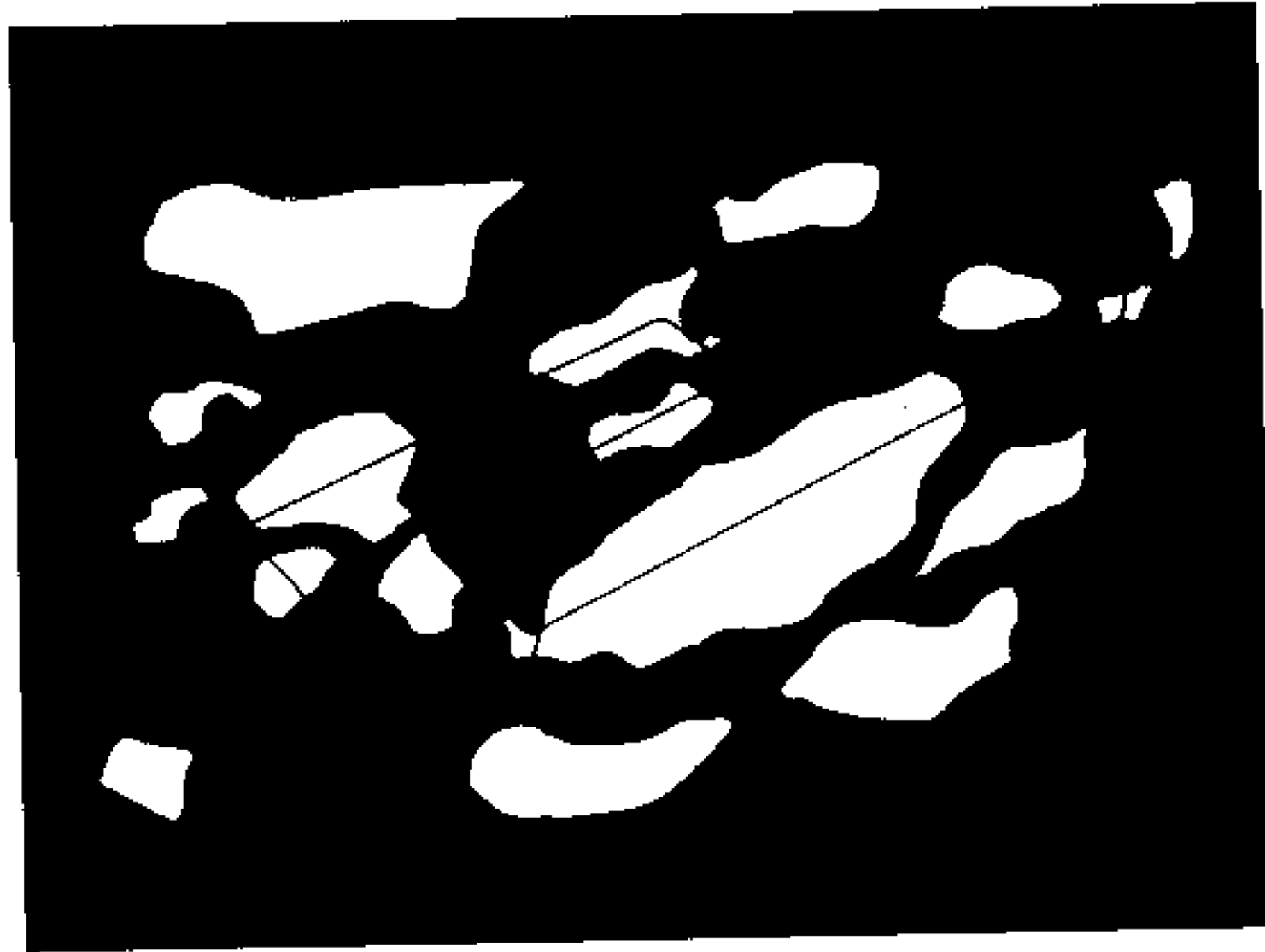


Figure 25. Nonrecoverable version of an object where the contour deletion is produced by an occluding surface.





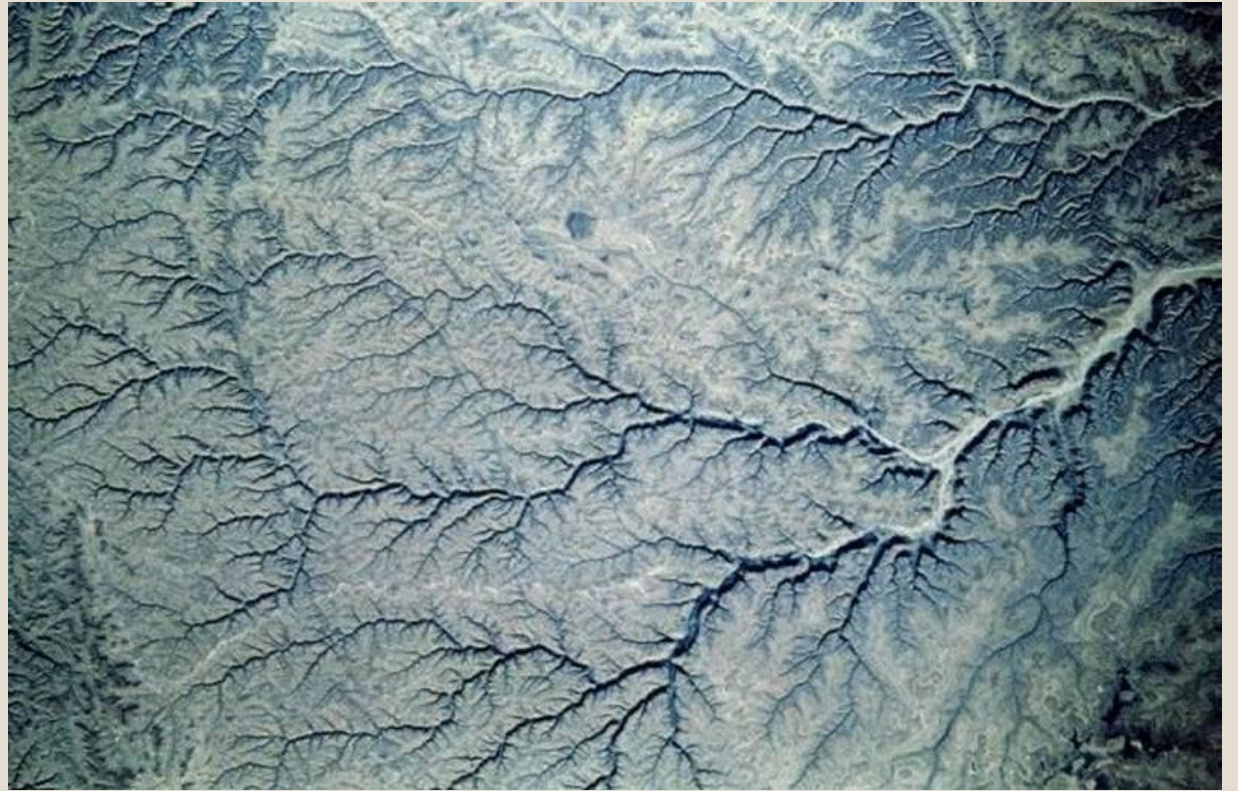


Issues?

- Specifies manmade objects, but not natural objects
- Most things in the world are natural
- Like Feature theory, they cannot be easily broken down into parts

Fractals

- Unlike manmade objects, most things in nature cannot be described using simple geometric shapes.
- Fractals are self-repeating patterns that describe nature much better.



Better Models

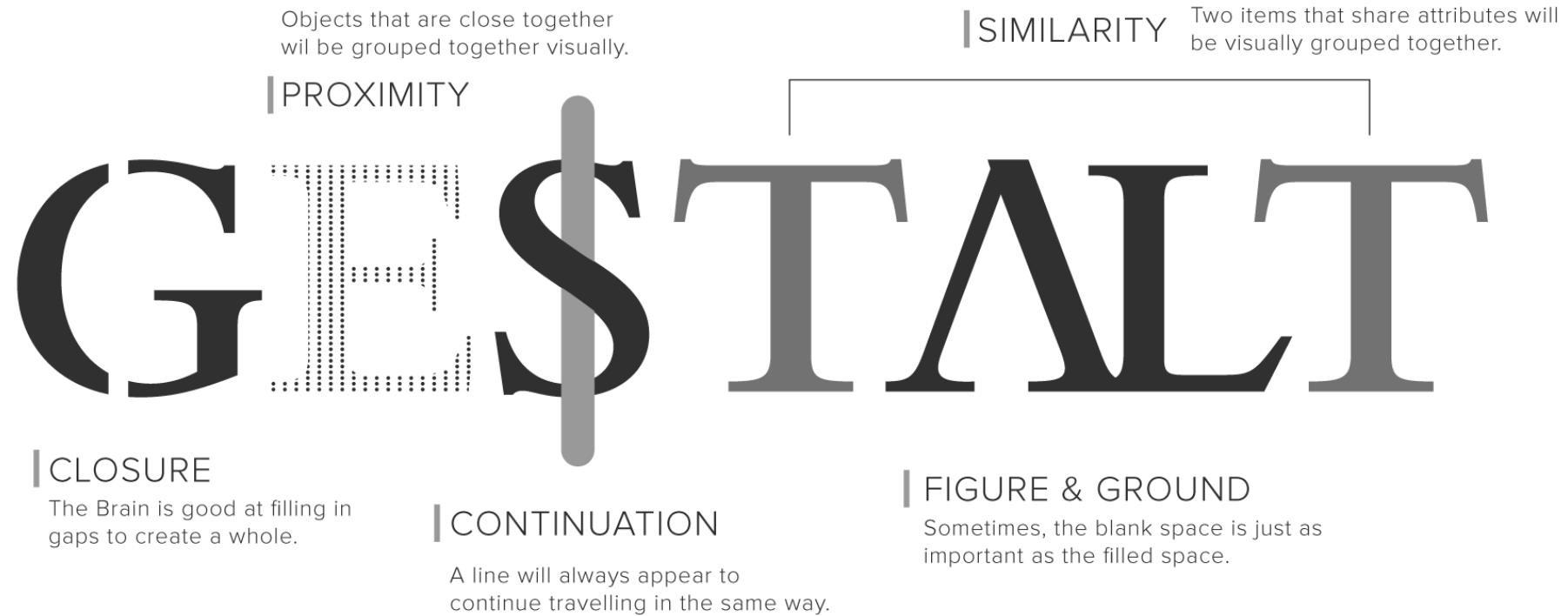
- Gestalt
- Interactive Network Models



Gestalt

- “The whole is more than the sum of its parts”
- Top-down organizational principles as important as individual features of scene
- Relationship between what is perceived is paramount.





PRINCIPALS OF VISUAL PERCEPTION

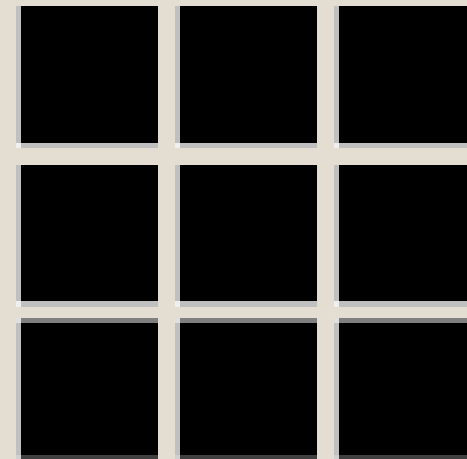
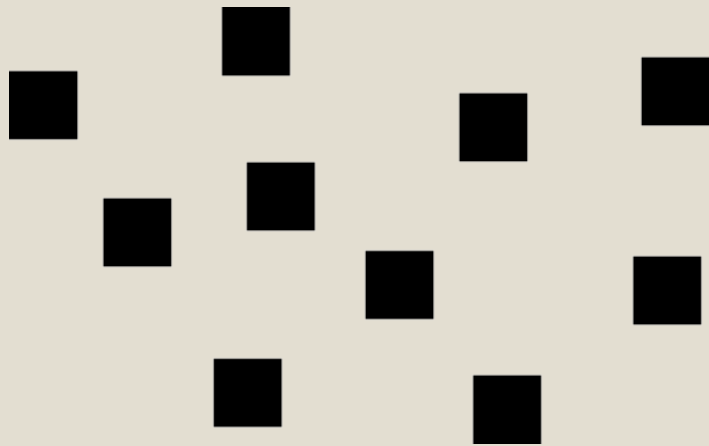
Closure

- Closure occurs when an object is *incomplete* or a space is not *completely enclosed*. If enough of the shape is indicated, people perceive the whole by filling in the missing information.



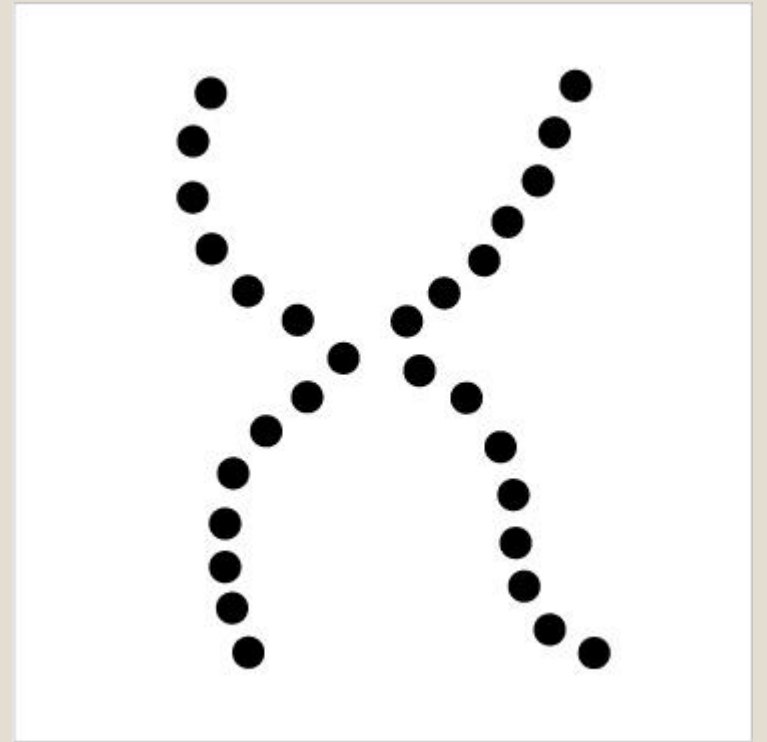
Proximity

- When elements appear together, they tend to be perceived as a group or unit.



Continuation

- When objects seems to continue into or through one another



Similarity

- When similar objects are near one another, they are perceived as a pattern.

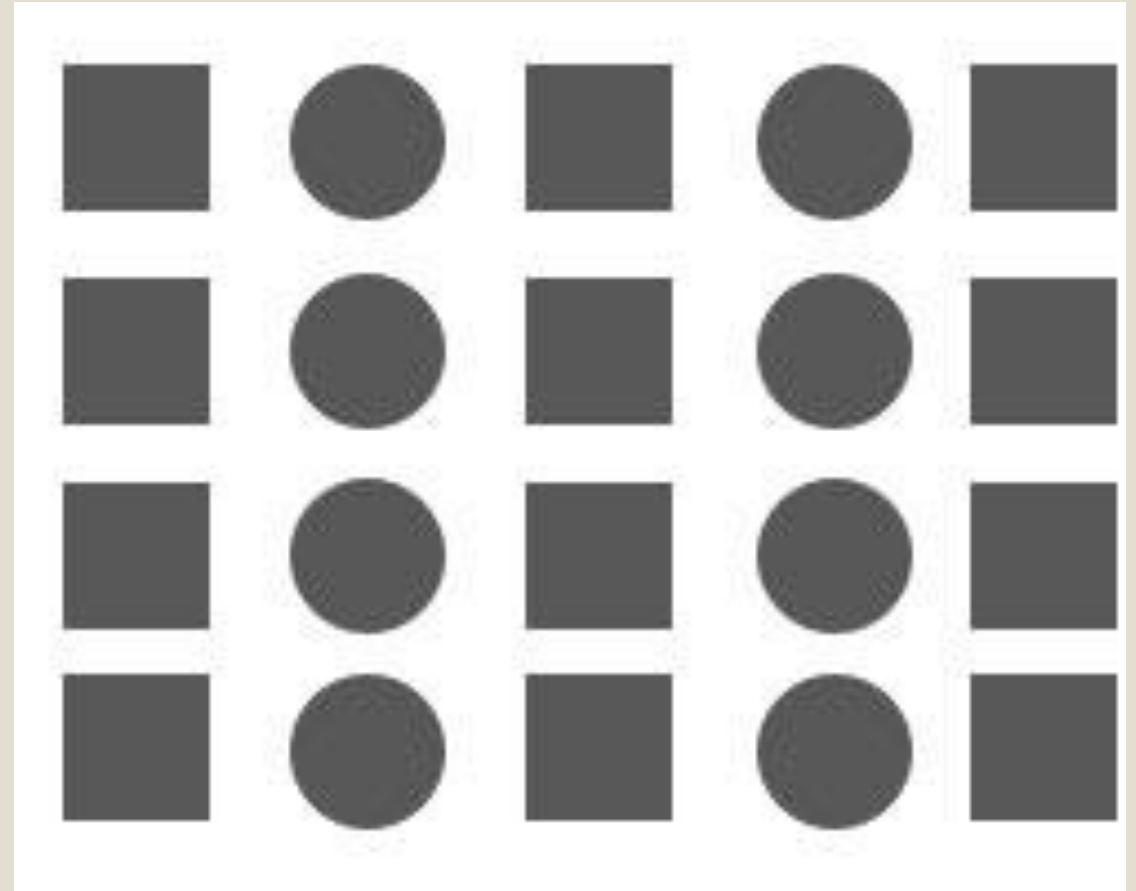
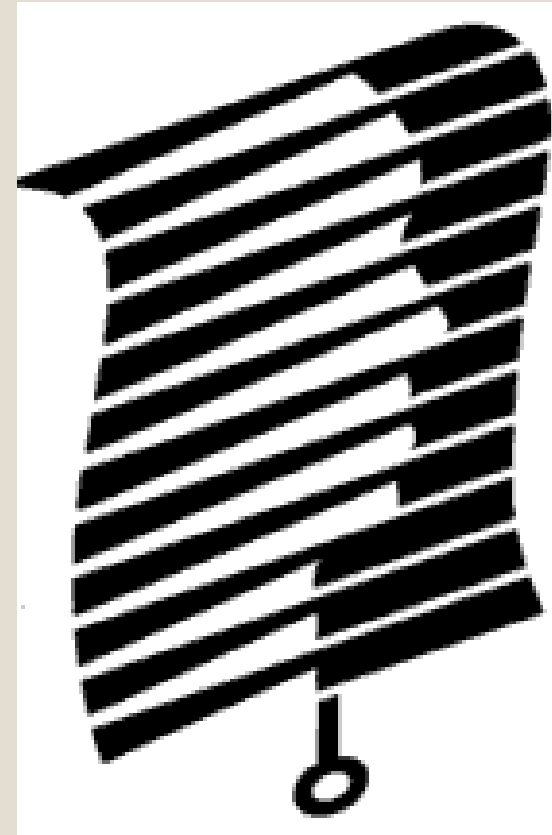
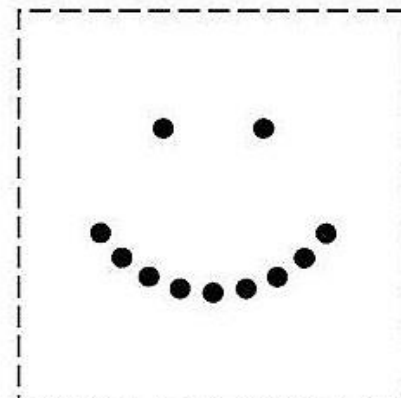
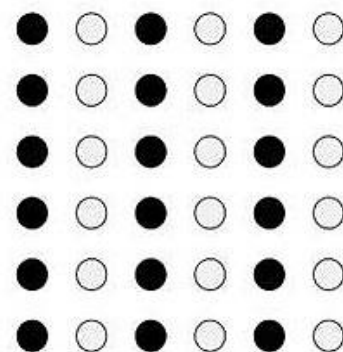
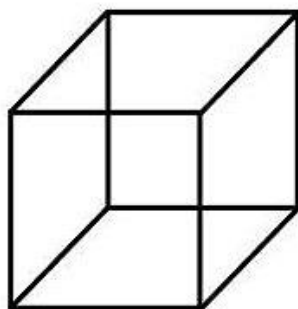
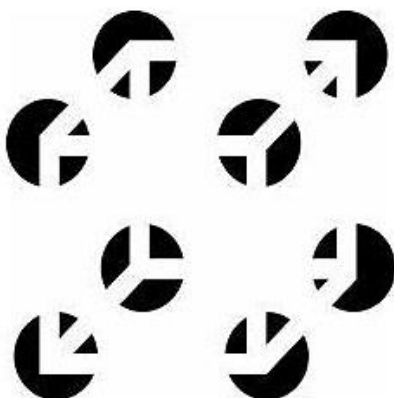
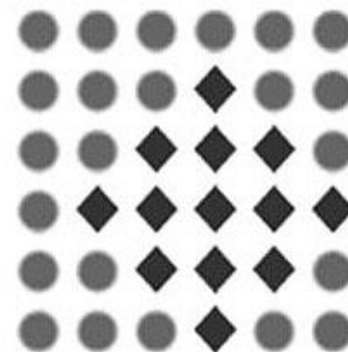
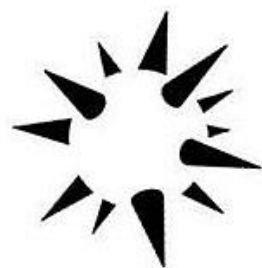
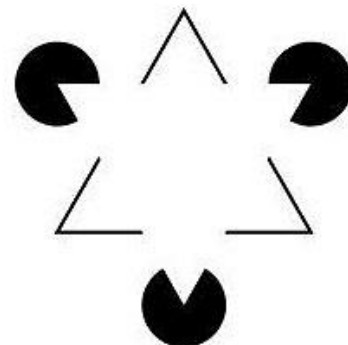
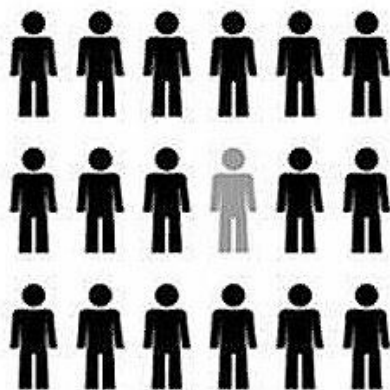
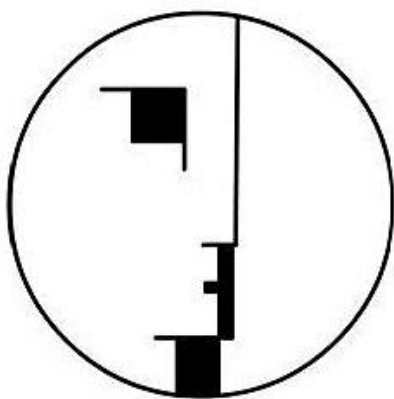


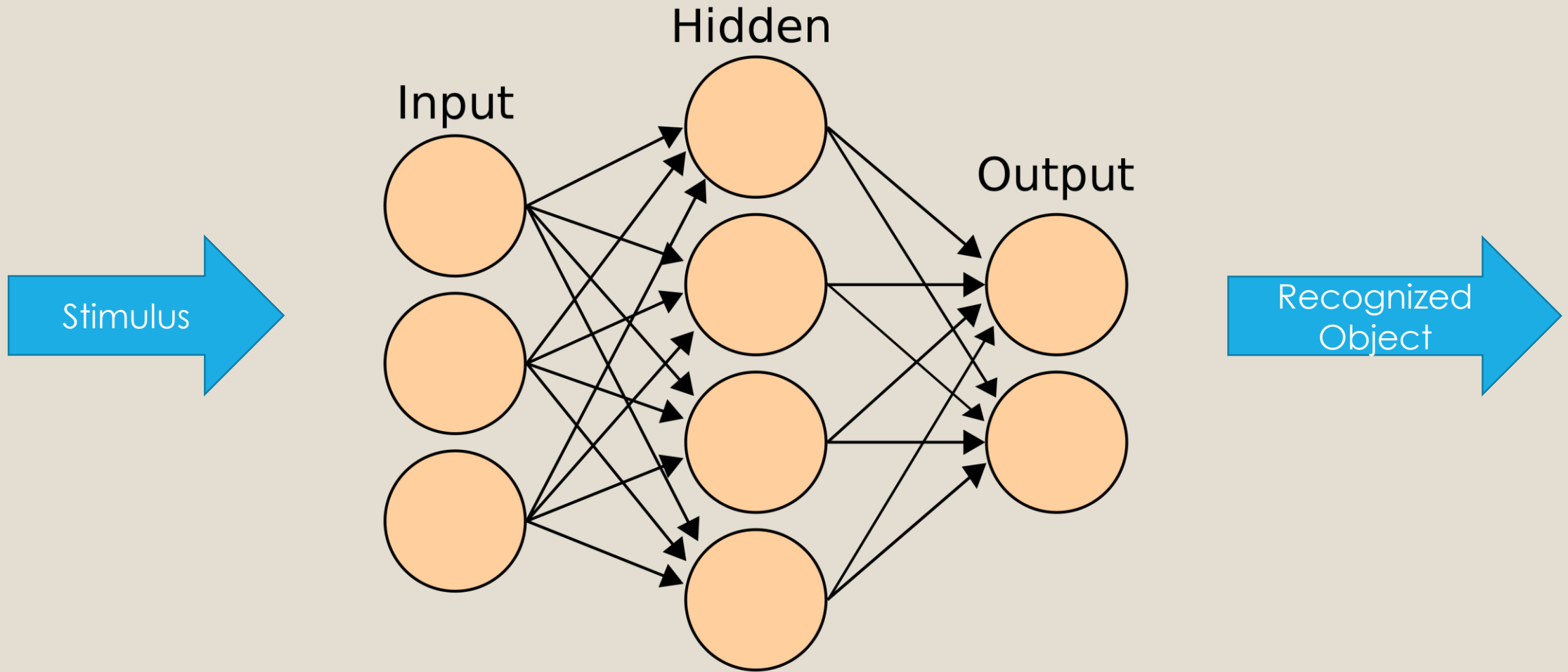
Figure & Ground

- The relationship between what is the ground of the scene and the figure determine the nature of the perceived object.



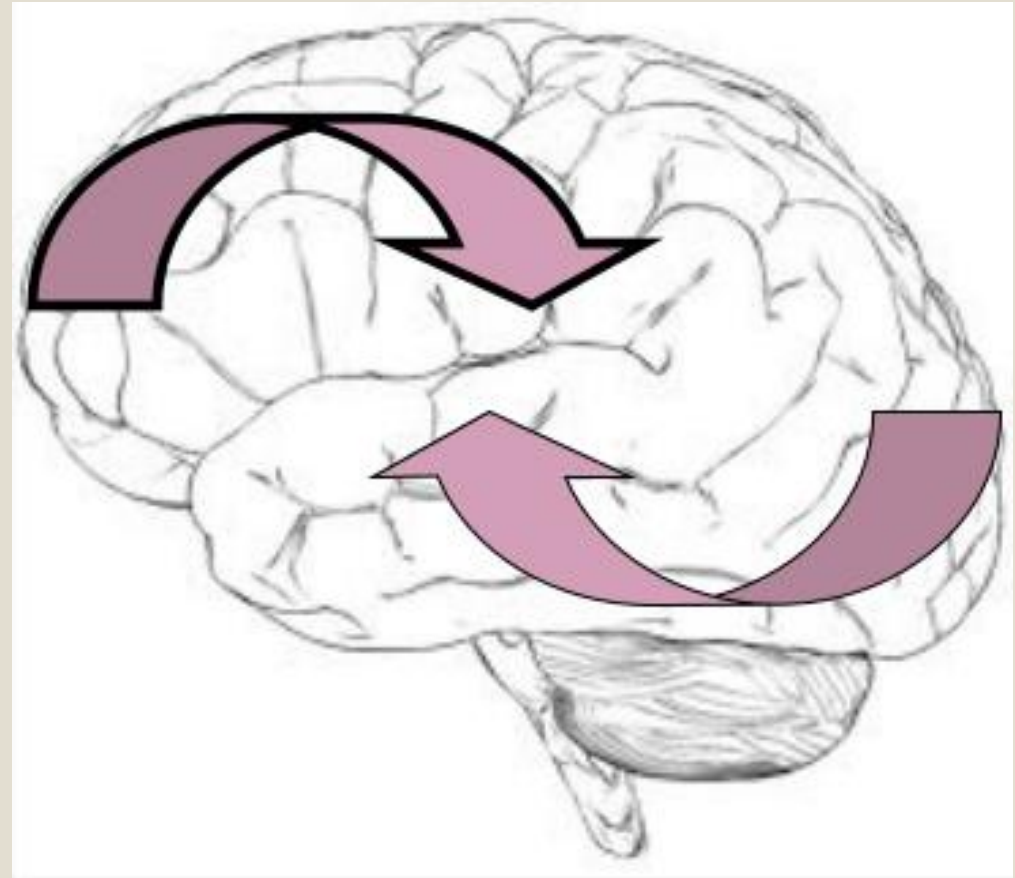


Interactive Network Model

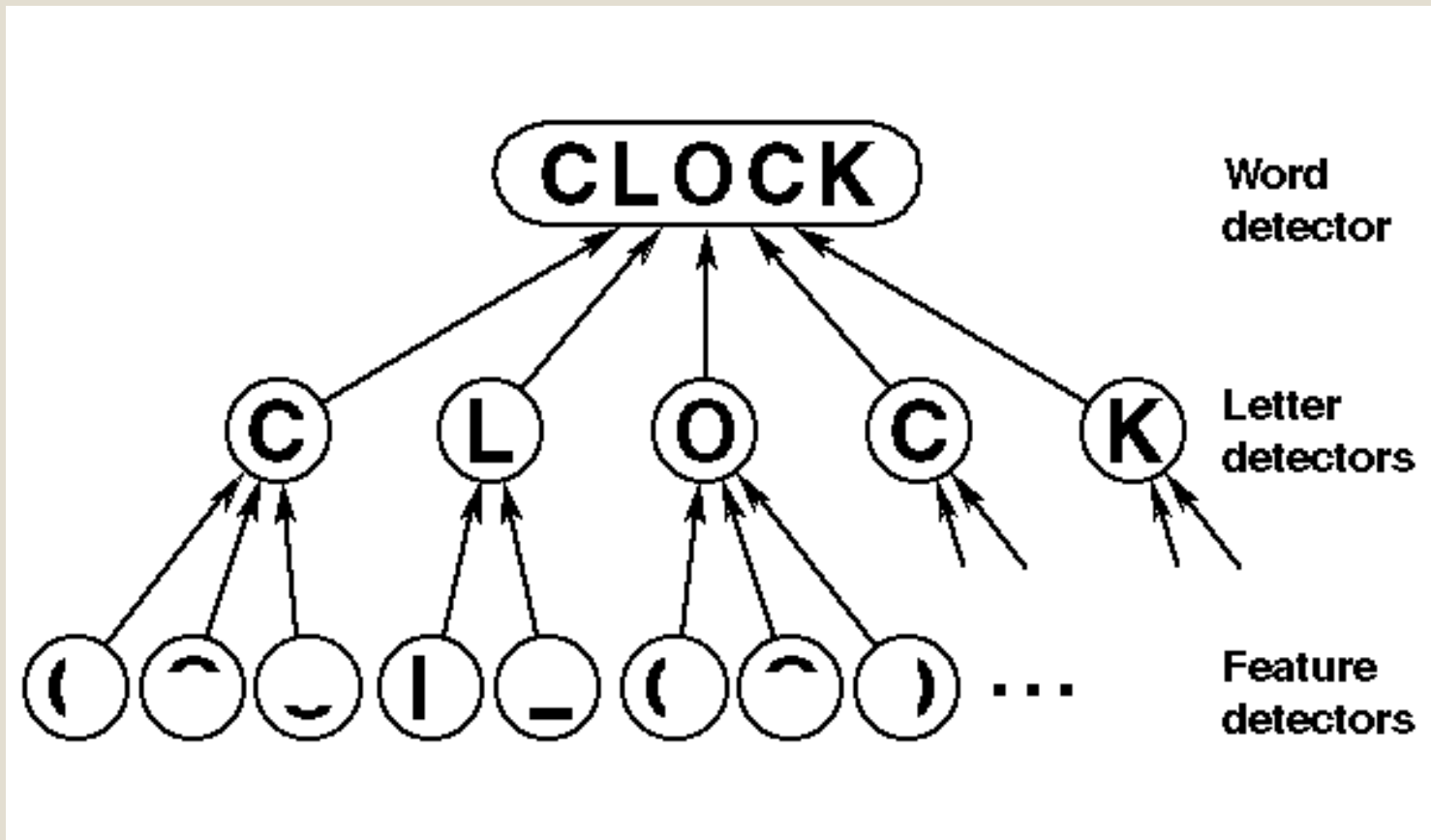


Principles of Network models

- Top-down and bottom up activation
- Activation and Inhibition
- Probabilistic activation

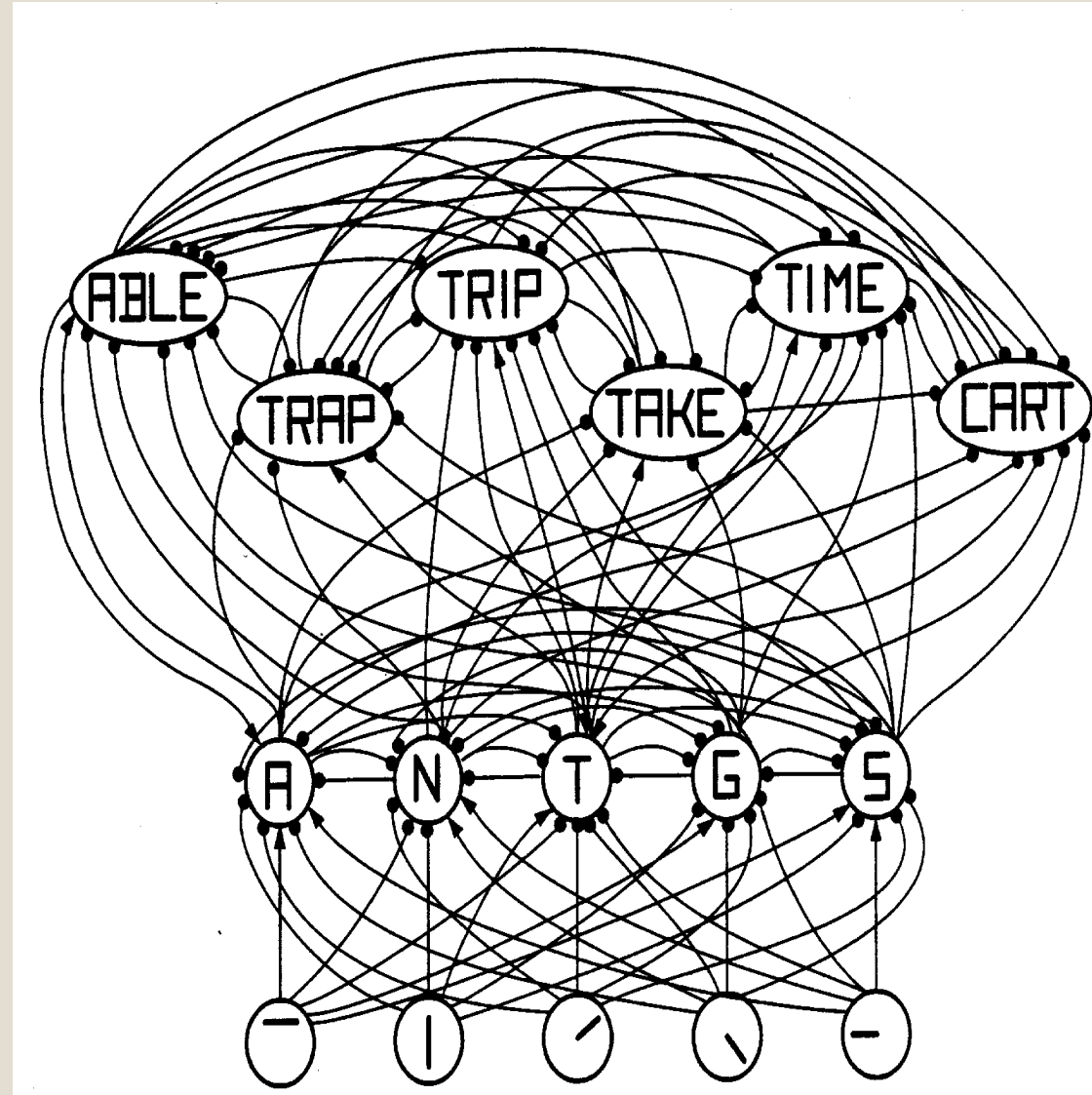


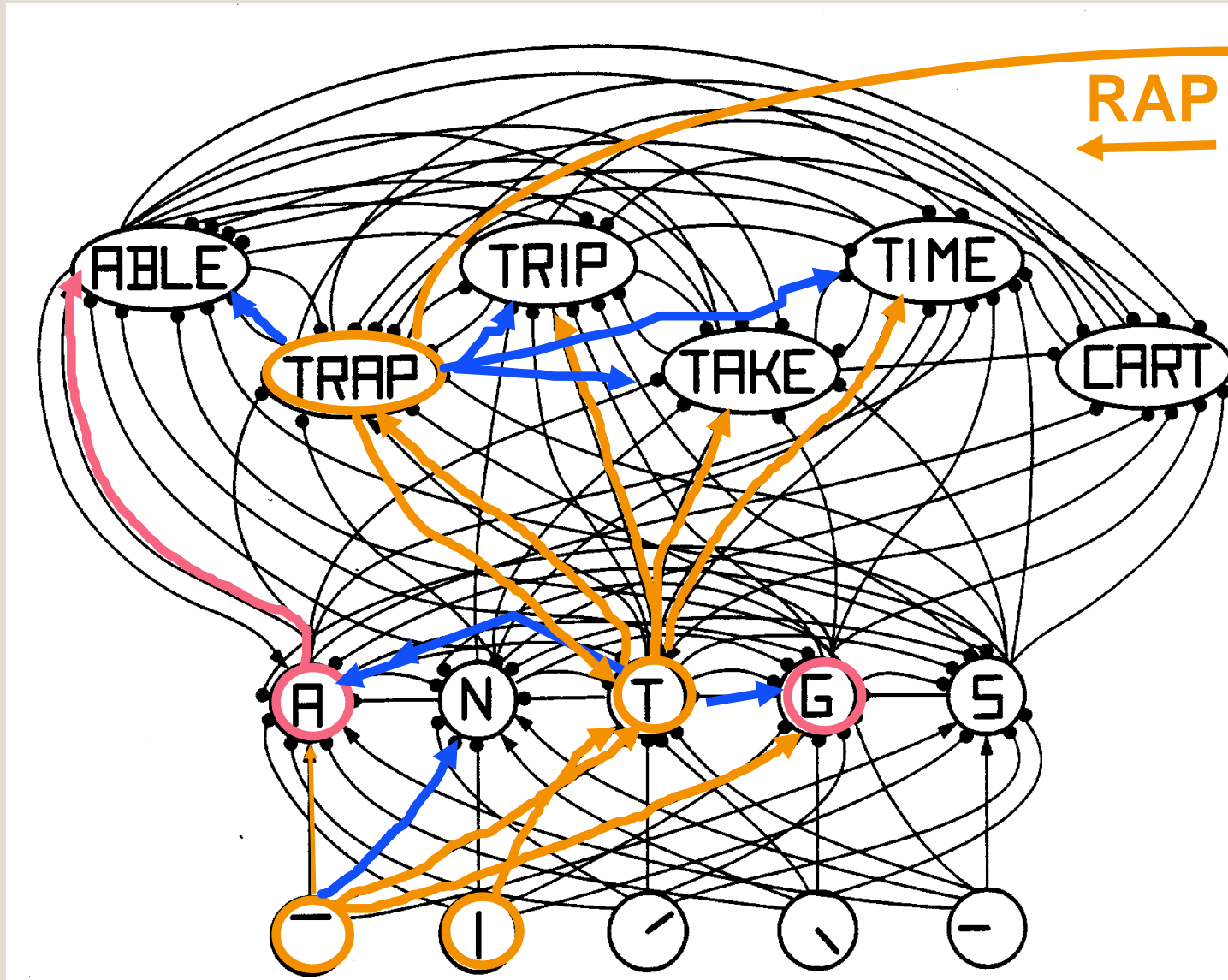
Feed-Forward Model



Feed-back Model

Activation & Inhibition





The robustness of an interactive activation model allows us to perceive words even when large chunks of them are missing.

The activation for the letters we see is still higher than for any other possible letter.



Dreams, ghosts, apparitions

- The expectations and top-down influence of the brain can allow for the experience of many things that aren't there.
- These often disappear with more information.

