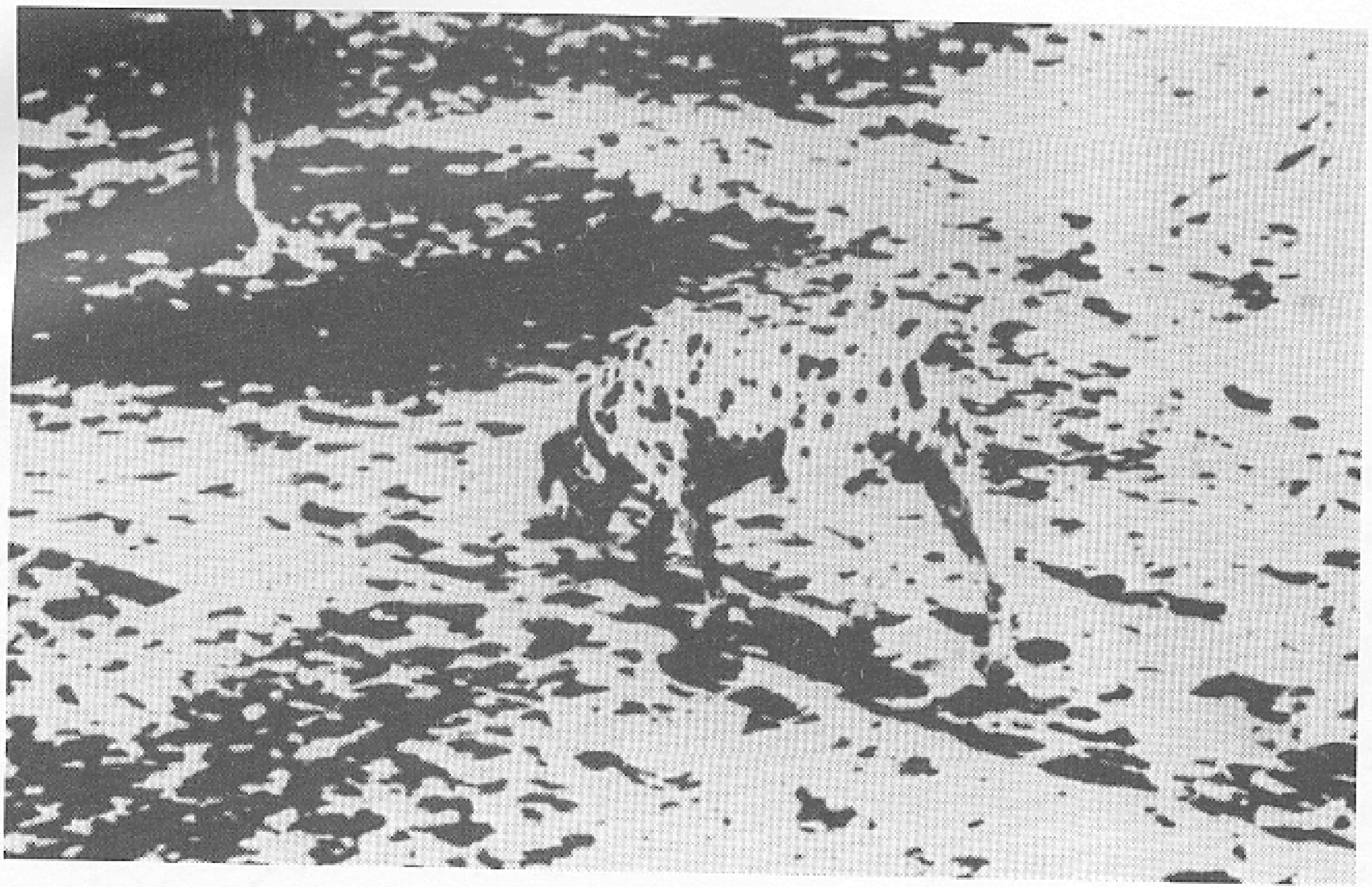
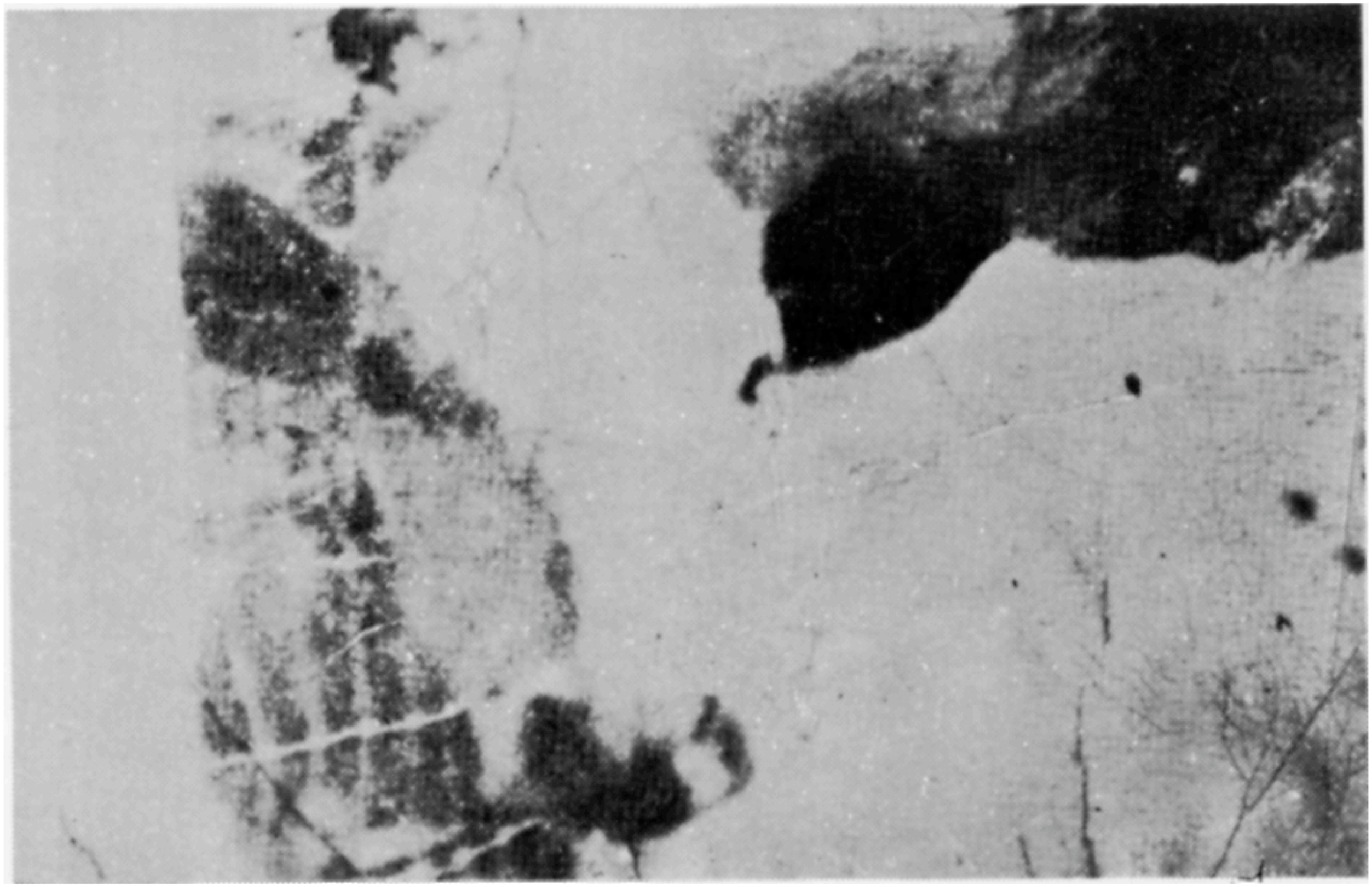
Bottom-up and Top-down Perception

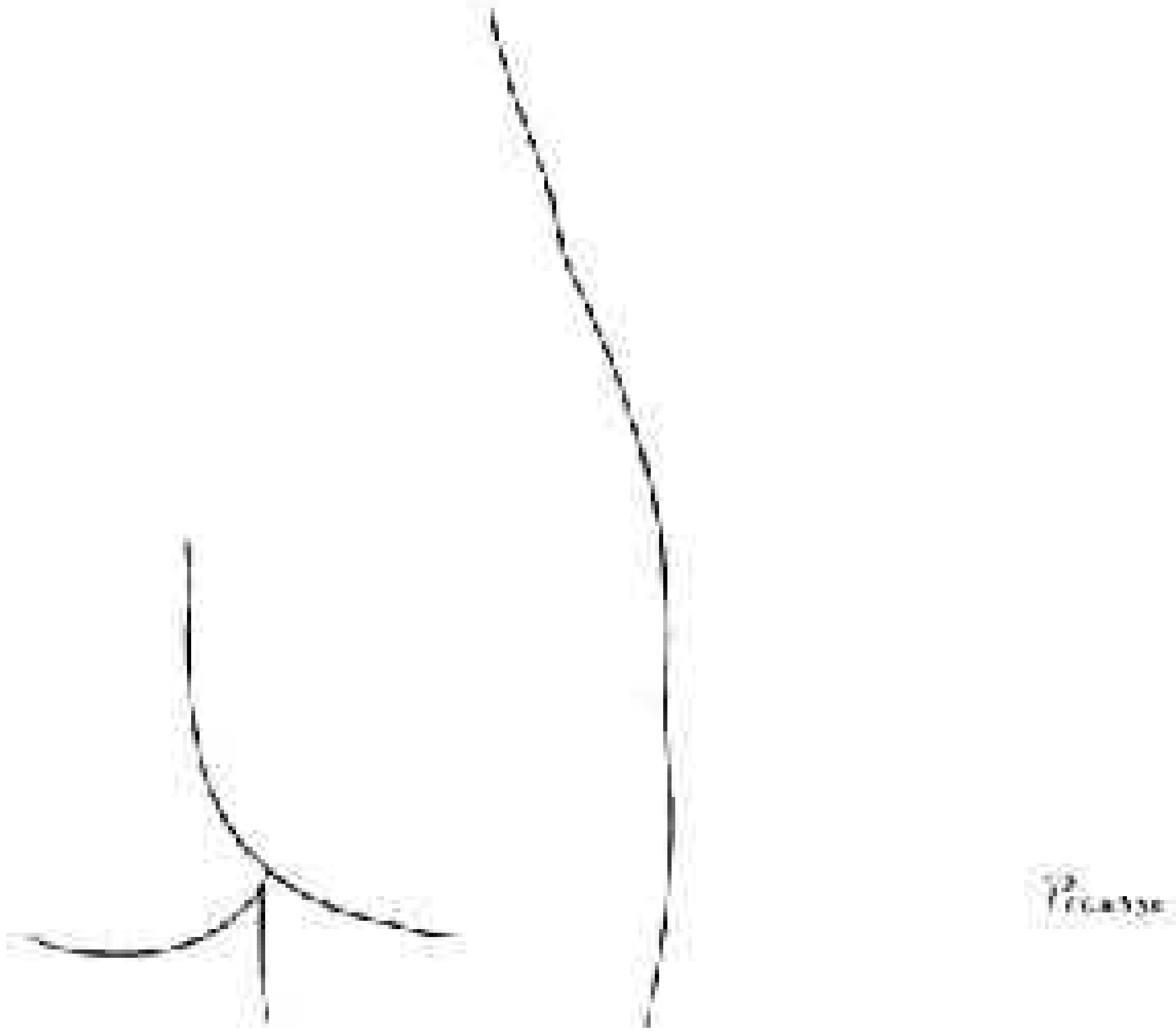
- Bottom-up perception
  - Physical characteristics of stimulus drive perception
  - Realism
- Top-down perception
  - Knowledge, expectations, or thoughts influence perception
  - Constructivism: we structure the world
  - "Perception is not determined simply by stimulus patterns; rather it is a dynamic searching for the best interpretation of the available data." (Gregory, 1966)
- Interactive
  - Perception is driven by both at the same time

Knowing where an object is can make an otherwise invisible object appear





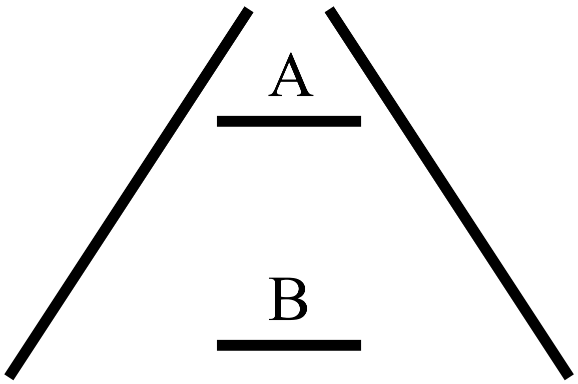
Matisse

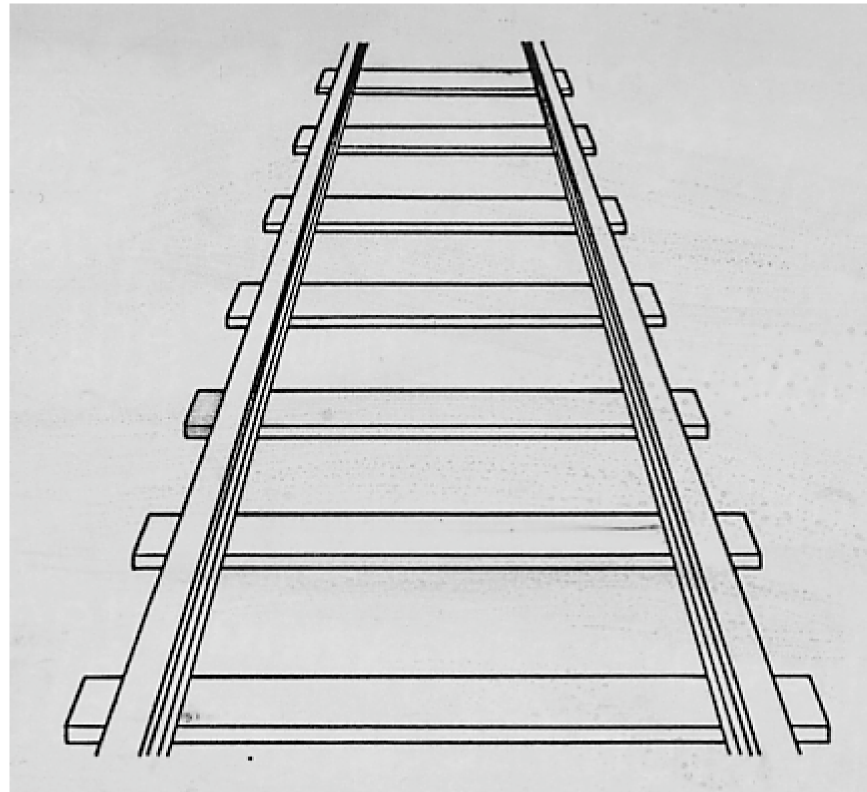

Picasso

Perceptual Illusions

- Why study illusions?
  - Illusions reveal constraints/biases on perception
    - Constraints are perceptual assumptions that we make
      - Usually correct but occasionally wrong
      - When wrong, illusion results
    - Illusions come from helpful processes
      - Without constraints, no perception at all!
  - Explore human contribution to perception by dissociating real world from our perception of it
- Case Studies
  - Railroad tracks illusion
  - Apparent Motion
  - Stereo depth perception

The Railroad Tracks Illusion



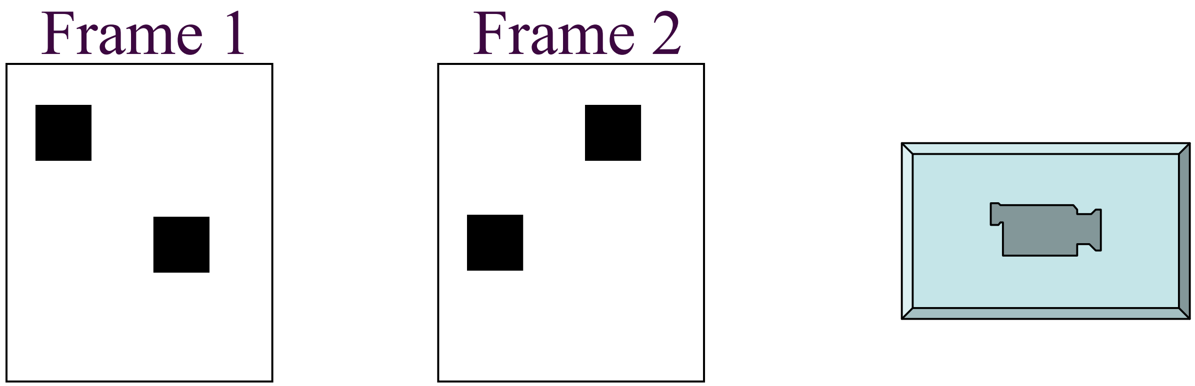


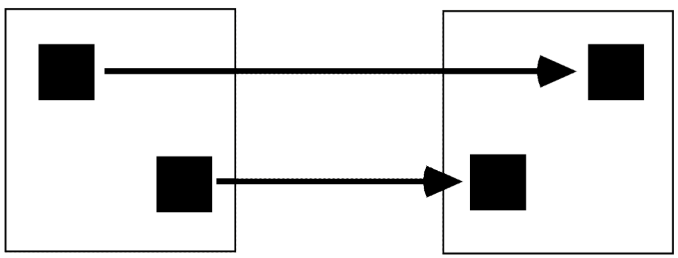
Assumption: the scene is taken from a 3-D world

Apparent Motion

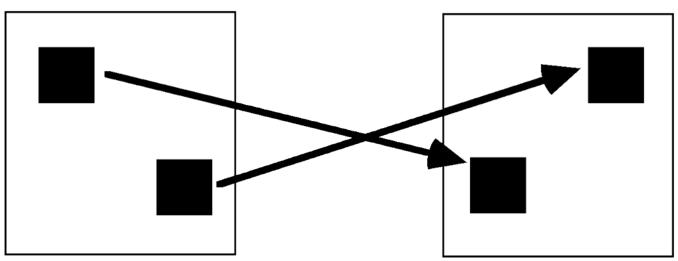
- Motion Perception
  - Importance for perceptual organization
  - Dedicated brain areas
- Apparent Motion
  - Motion from sequentially presented still frames
  - Assume objects in one frame are the same as those in the other frame, just moved
  - Challenge: How to determine which objects correspond to each other across frames

One-to-one Mapping Constraint

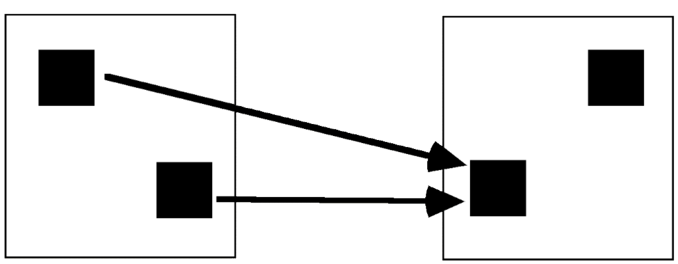




Yes, horizontal motion



Yes, vertical motion

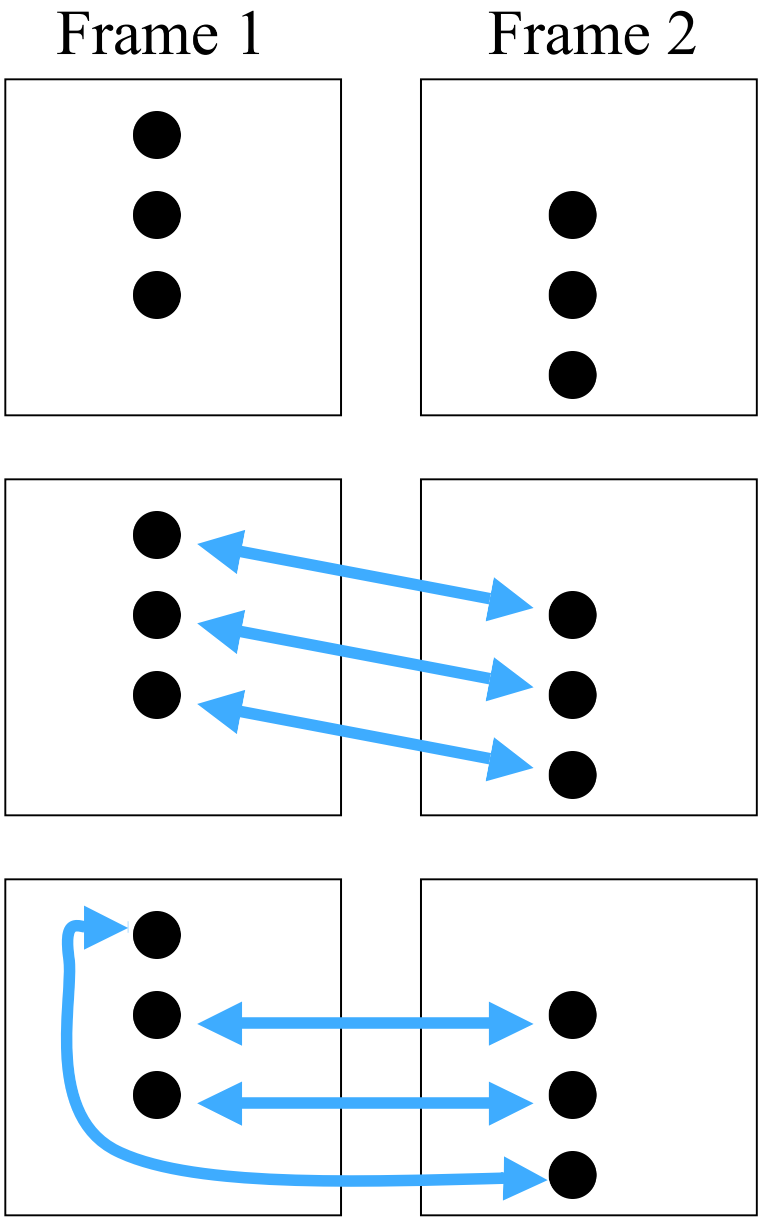


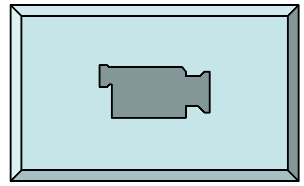
No, violates 1-to-1 mapping

Constraints on Motion Perception

- Proximity
  - Parts A and B tend to be the same object if they are close
- Shape similarity
  - Parts A and B tend to be the same object if they are similar in their shape
- Color and size similarity
- One-to-one mapping constraint
  - Two parts at Time T should not correspond to one part at Time T+1
  - Global coherence: Correspondences all influence each other

Ternus Effect

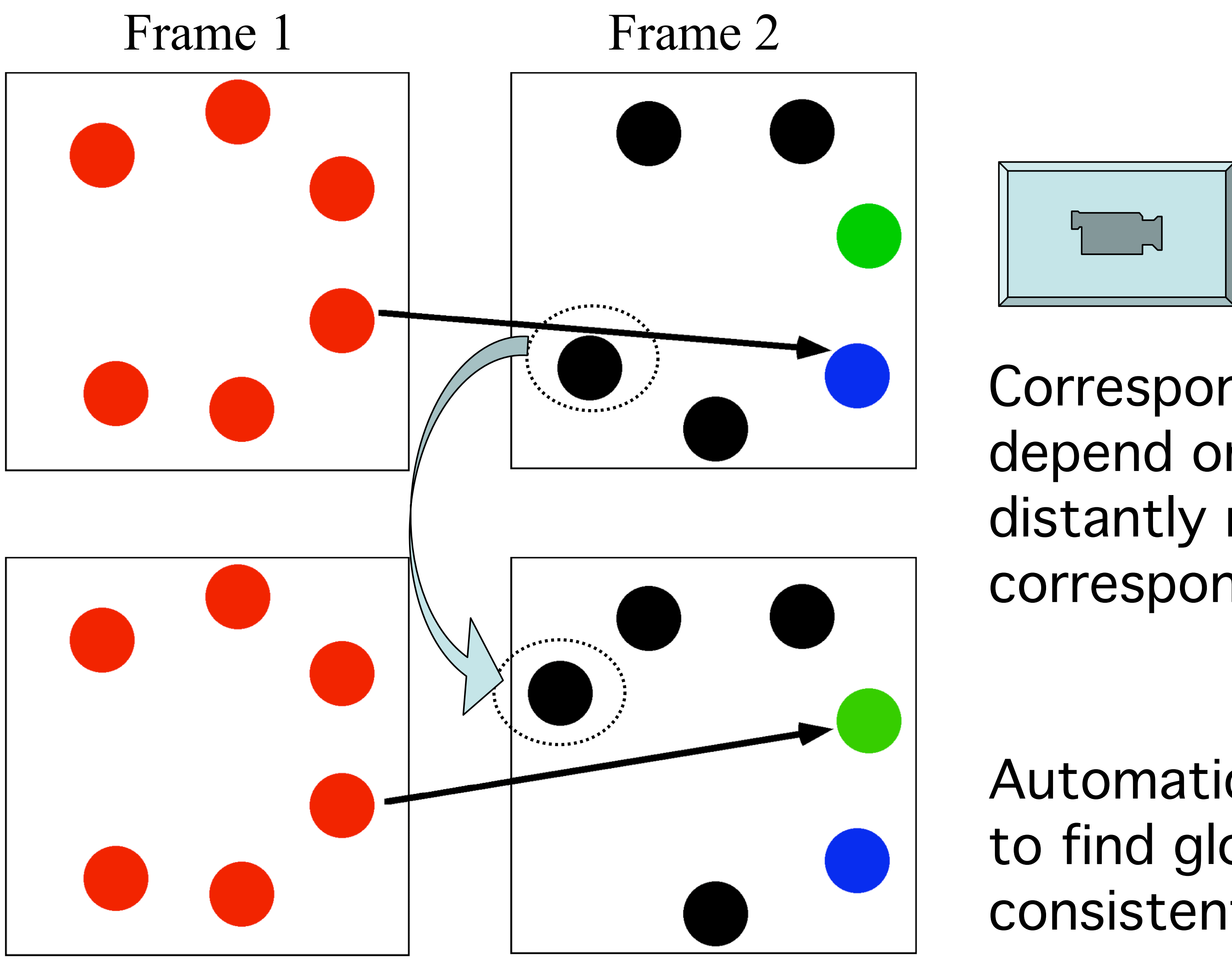




Globally coherent correspondences (Long pause)

Locally determined correspondences (Short pause)

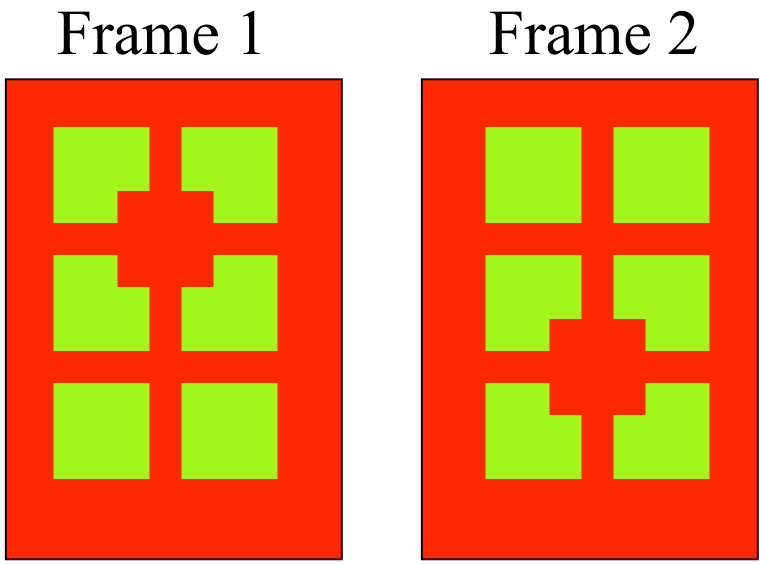
Globally Coherent Motion

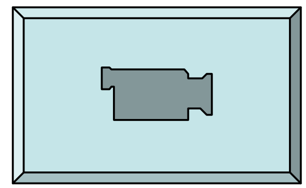


Correspondences depend on distantly related correspondences

Automatic tendency to find globally consistent solutions

Illusory Motion of Illusory Contours





Illusory square moves, so the generation of illusory contours occurs before the generation of apparent motion.

If contours were generated only after motion is perceived, then people would see a pac-man (which requires no illusory contours) rotating.

Constraint Satisfaction Network for Apparent Motion Perception

Nodes represent correspondences between elements across frames
Activity represents strength of correspondence
Neural network does not learn: connections are hard-wired
Activation/inhibition spreads according to constraints:

- Shape, color, size, location similarity: if corresponding elements are similar, then activity increases
- Motion similarity: Excitation between two nodes if similar directions of motion are implied by them
- Consistency
  - Consistent nodes excite one another
  - Inconsistent nodes inhibit one another
  - Consistent = one-to-one mapping
  - Inconsistent = two-to-one mapping
    - As the consistency constraint increases, increase the negative weight that connects all inconsistent nodes, and the positive weight that connects consistent nodes
- Match
  - Bias for each cell to have a correspondence
    - If an object has no match, increase the activation of all nodes that connect that object to other objects

Constraint Satisfaction Network for Apparent Motion Perception
(Dawson, 1991; Ullman, 1979)

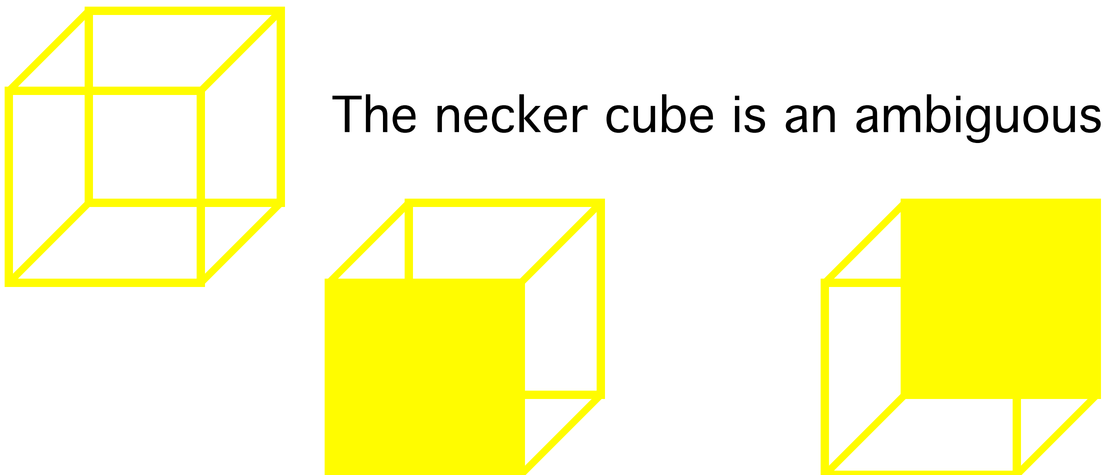
Processing in model

Time = number of cycles of activation passing

Soft-constraints (neural networks need not be tabula rasas)

Activation passing leads to increased harmony over time

- Harmony = consistency between nodes



The necker cube is an ambiguous object

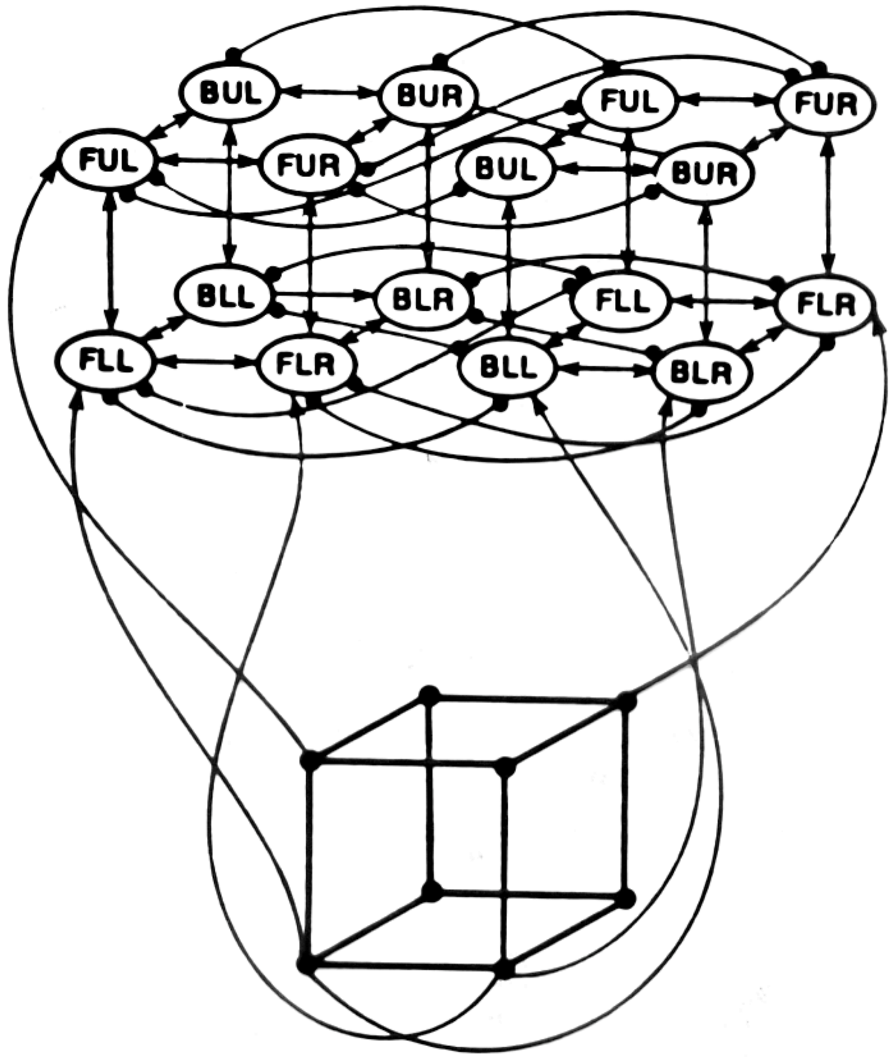
Each interpretation is internally consistent and harmonious

Networks settle into one of two consistent interpretations

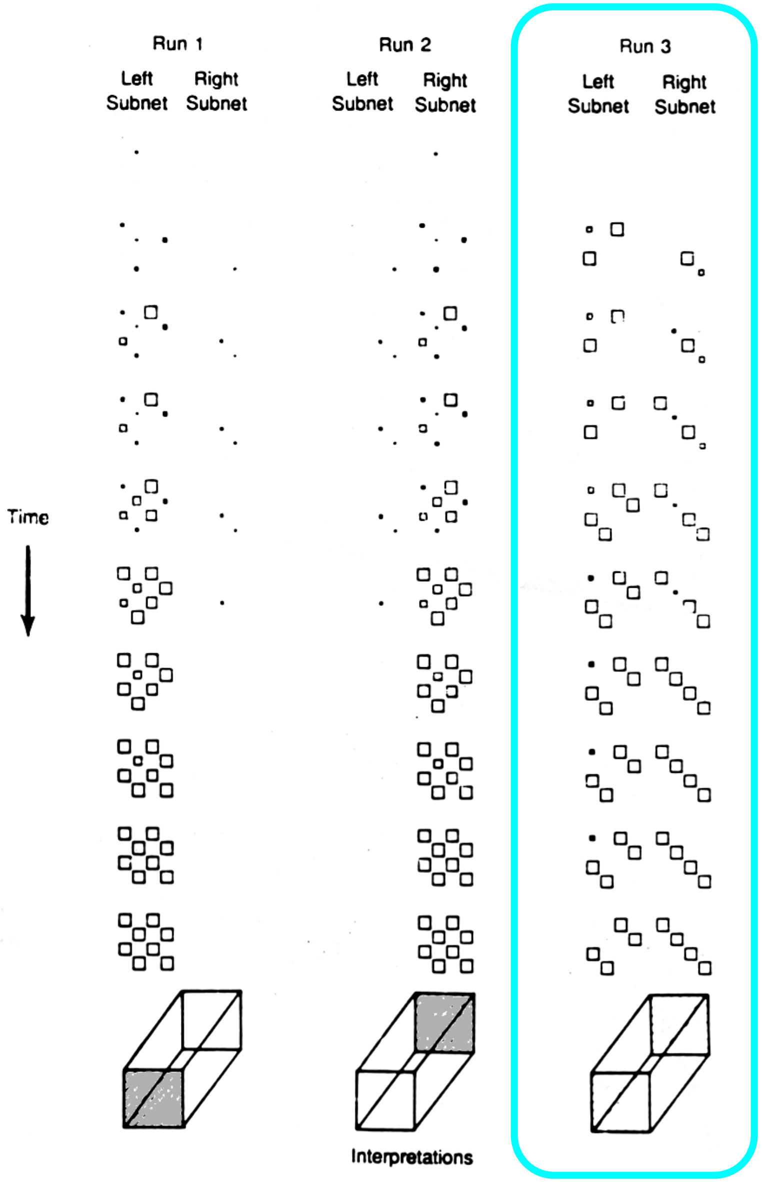
Constraint Satisfaction Network for Necker Cube Perception

Inhibitory

Excitatory

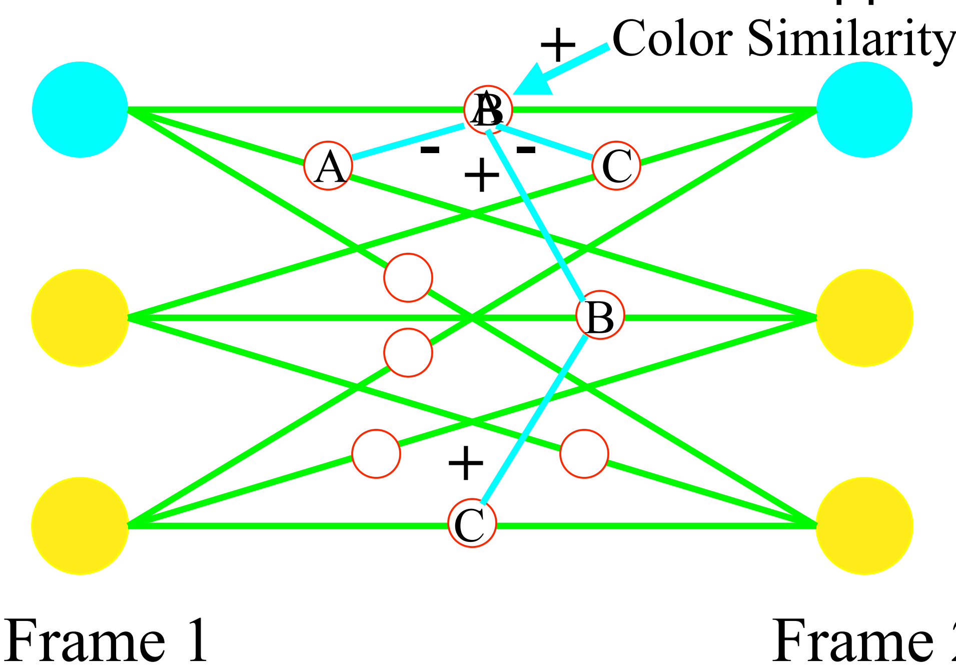


Constraint Satisfaction Network for Necker Cube Perception



Unlikely

Constraint Satisfaction Network for Apparent Motion



Inconsistent nodes if 2-to-1 mapping

Activity of B_{t+1} = Activity of B_t - Activity of C_t

Consistent nodes if not 2-to-1 mapping

Activity of B_{t+1} = Activity of B_t + Activity of C_t

Excitatory and inhibitory links are hard-wired according to constraints, not learned

Activity of a node is based on:

N objects per scene -> N*N nodes

Similarities between elements connected by the node

The activity of other nodes

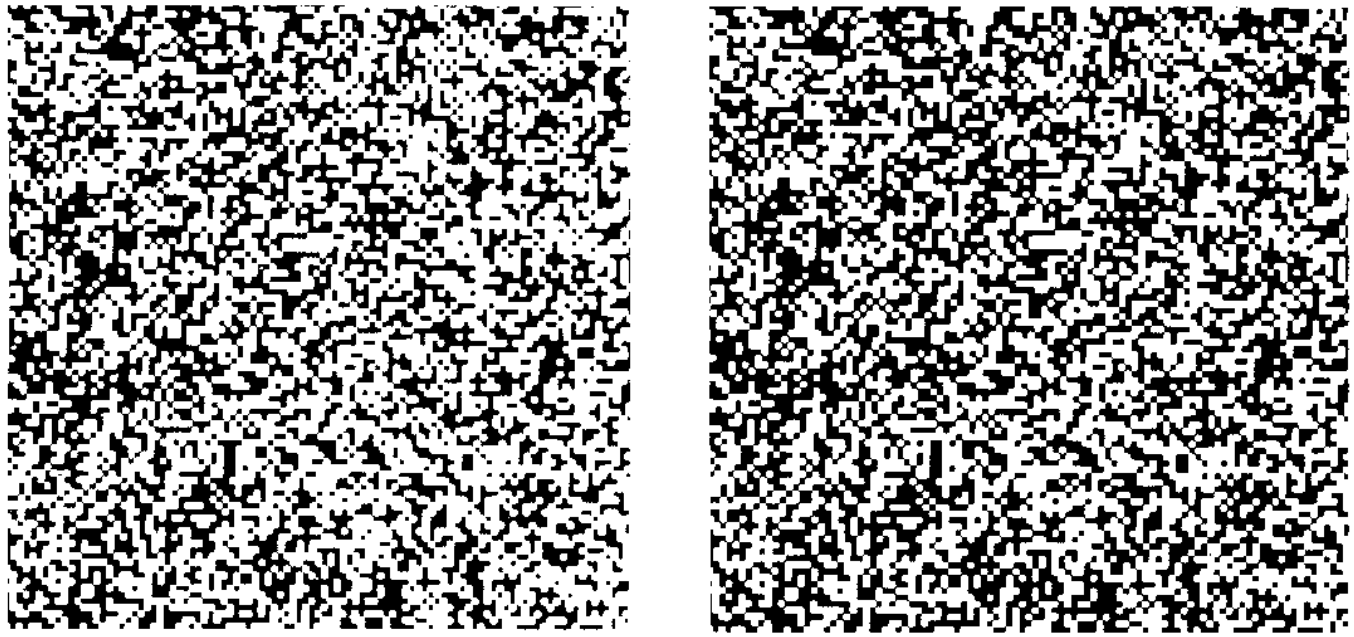
Applications of the Apparent Motion Network

- Similarity matters
  - Similar objects are more likely to correspond to each other
- Network finds consistent correspondences
- Hysteresis
  - Once a stable percept is found, it resists change
  - Adding randomness helps appropriate restructuring
- Predicts distribution of responses
  - Make model stochastic by adding randomness to nodes
  - Even with randomness, stable percepts are found
- Applicability to other areas
  - Stereo depth perception (Marr & Poggio, 1979)
  - Analogical reasoning (Goldstone, 1994; Holyoak & Thagard, 1989)
  - Translation between languages and conceptual systems (Goldstone & Rogosky, 2002)

The Correspondence Problem in Depth Perception

- Stereopsis as a major depth cue
  - Left and right eyes see different images
  - Differences in positions of objects in two eyes tells us about their depth
  - Correspondence problem: What element in the left eye corresponds to what element in the right eye?
- Analogy to apparent motion
  - Frame 1: Frame 2 :: Left eye image : Right eye image
  - For both apparent motion and stereopsis, for two images elements to correspond means that they come from the same real-world object
  - Constraints: location similarity, shape similarity, 1-to-1 mapping, smoothness

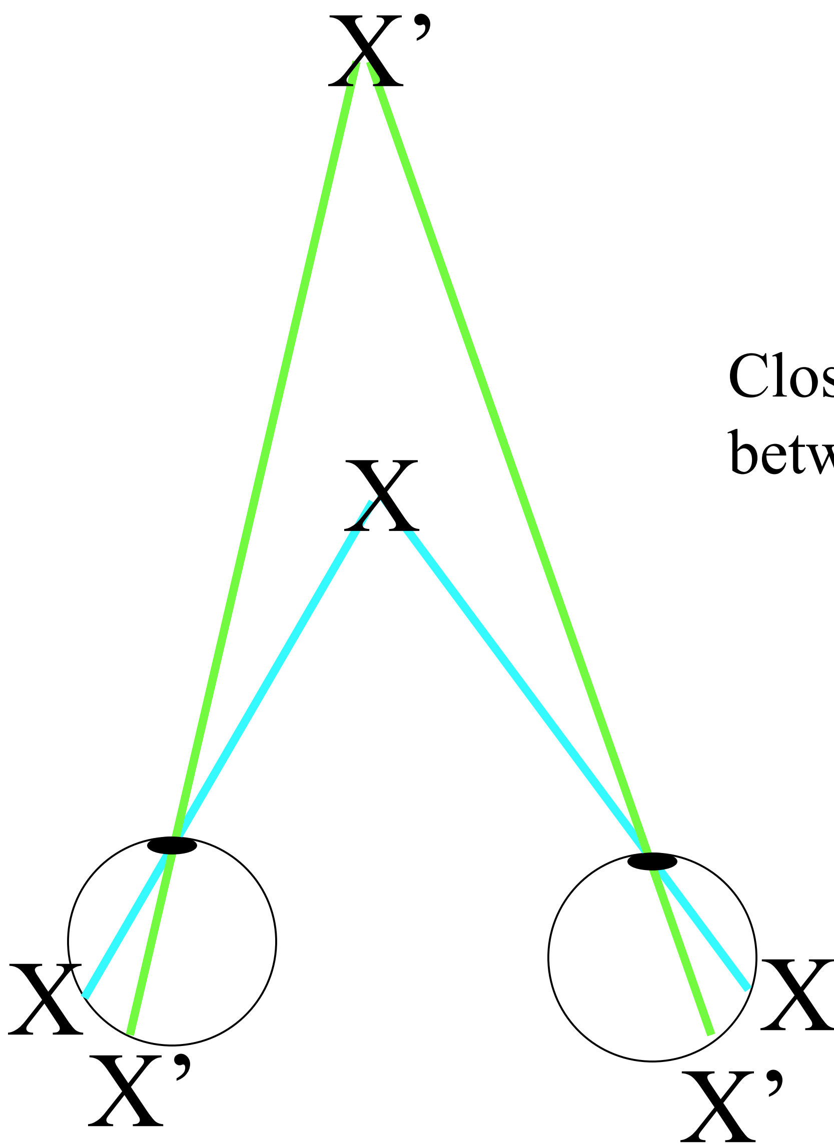
Random-dot stereograms (Julesz, 1971)



1	0	1	0	1	0	0	1	0	1
1	0	0	1	0	1	0	1	0	0
0	0	1	1	0	1	1	0	1	0
0	1	0	A	A	B	B	X	0	1
1	1	1	B	A	B	A	Y	0	1
0	0	1	A	A	B	A	Y	1	0
1	1	1	B	B	A	B	X	0	1
1	0	0	1	1	0	1	1	0	1
1	1	0	0	1	1	0	1	1	1
0	1	0	0	0	1	1	1	1	0

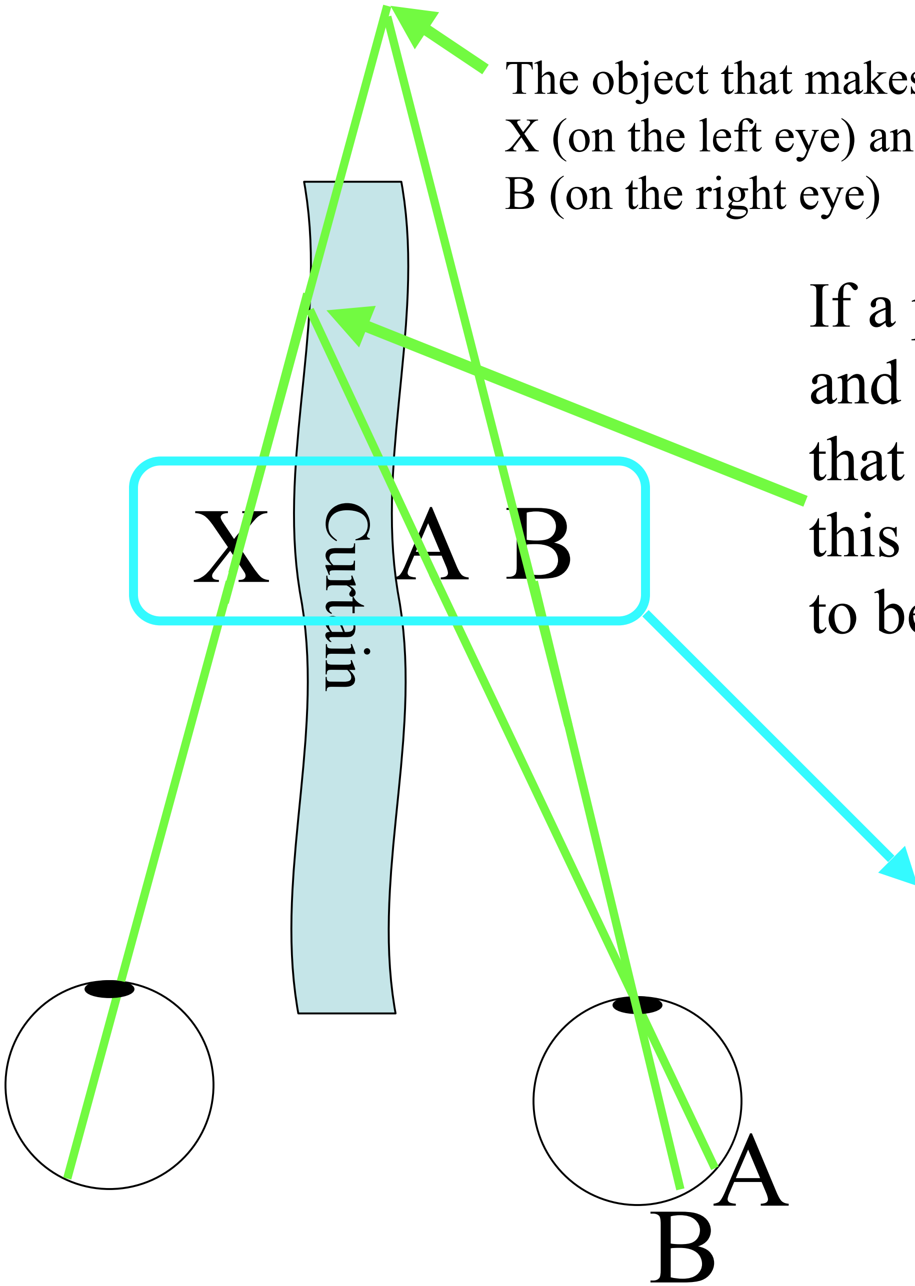
1	0	1	0	1	0	0	1	0	1
1	0	0	1	0	1	0	1	0	0
0	0	1	1	0	1	1	0	1	0
0	1	0	Y	A	A	B	B	0	1
1	1	1	X	B	A	B	A	0	1
0	0	1	X	A	A	B	A	1	0
1	1	1	Y	B	B	A	B	0	1
1	0	0	1	1	0	1	1	0	1
1	1	0	0	1	1	0	1	1	1
0	1	0	0	0	1	1	1	1	0

Retinal Disparities



Closer object = greater disparity between retinal images

Illusions in Stereopsis



The object that makes Image X (on the left eye) and Image B (on the right eye)

If a person sees X on the left eye, and A on the right eye and assumes that they come from the same object, this is where the object would need to be.

The closer two images are that are assumed to come from the same object, the closer that object is assumed to be.

Marr & Poggio (1976)

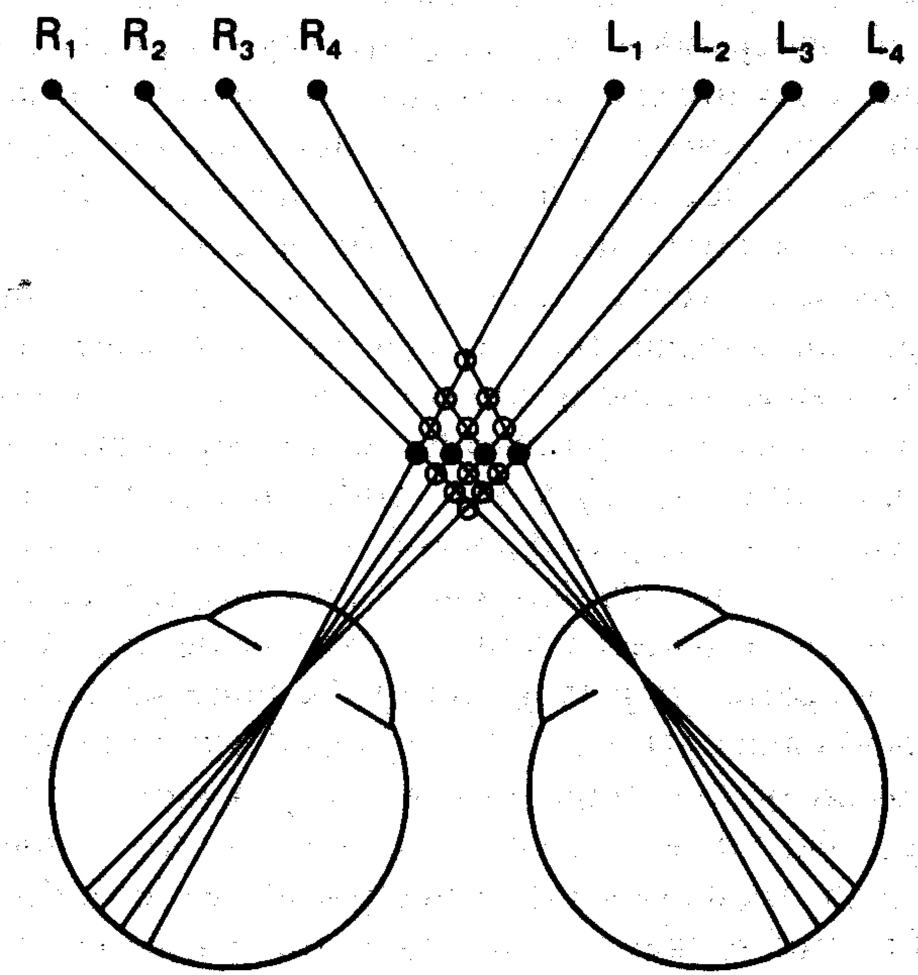


FIG. 2.21. Two eyes viewing a simplified stereogram in which each eye sees just 4 dots. Each of the left eye's dots (L_1 to L_4) could match any of the right eye's dots (R_1 to R_4), so that the number of possible matches, shown with filled and open circles, is very large. The visual system chooses the matches which are shown with filled circles. (Adapted from Marr & Poggio, 1976, and reproduced from Bruce & Green, 1985, with permission.)

Marr & Poggio (1976)

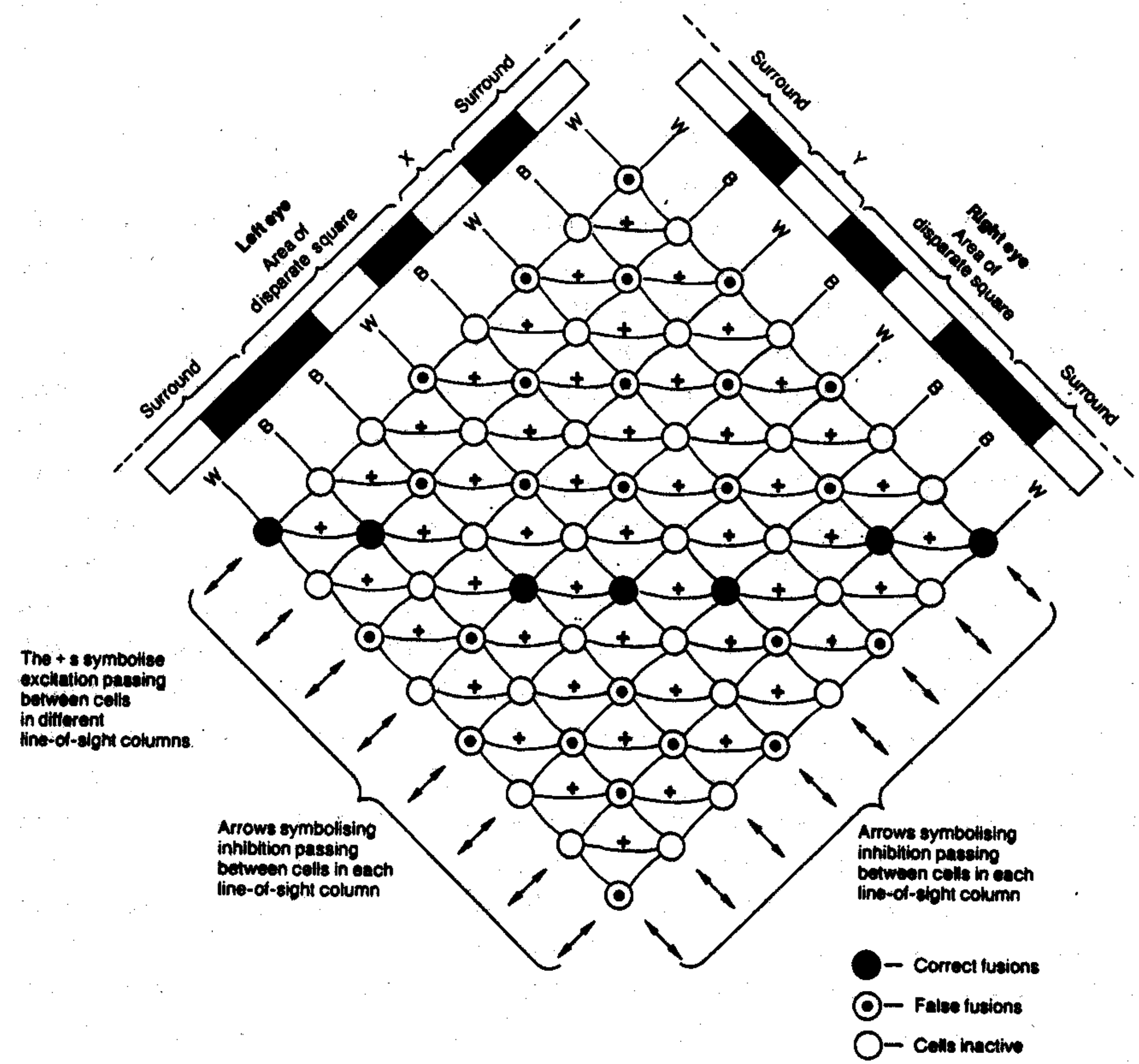


FIG. 2.26. A portion of a neural network to solve random-dot stereograms. (Reproduced from Frisby, 1979, with permission.)

The Mueller-Lyer Illusion

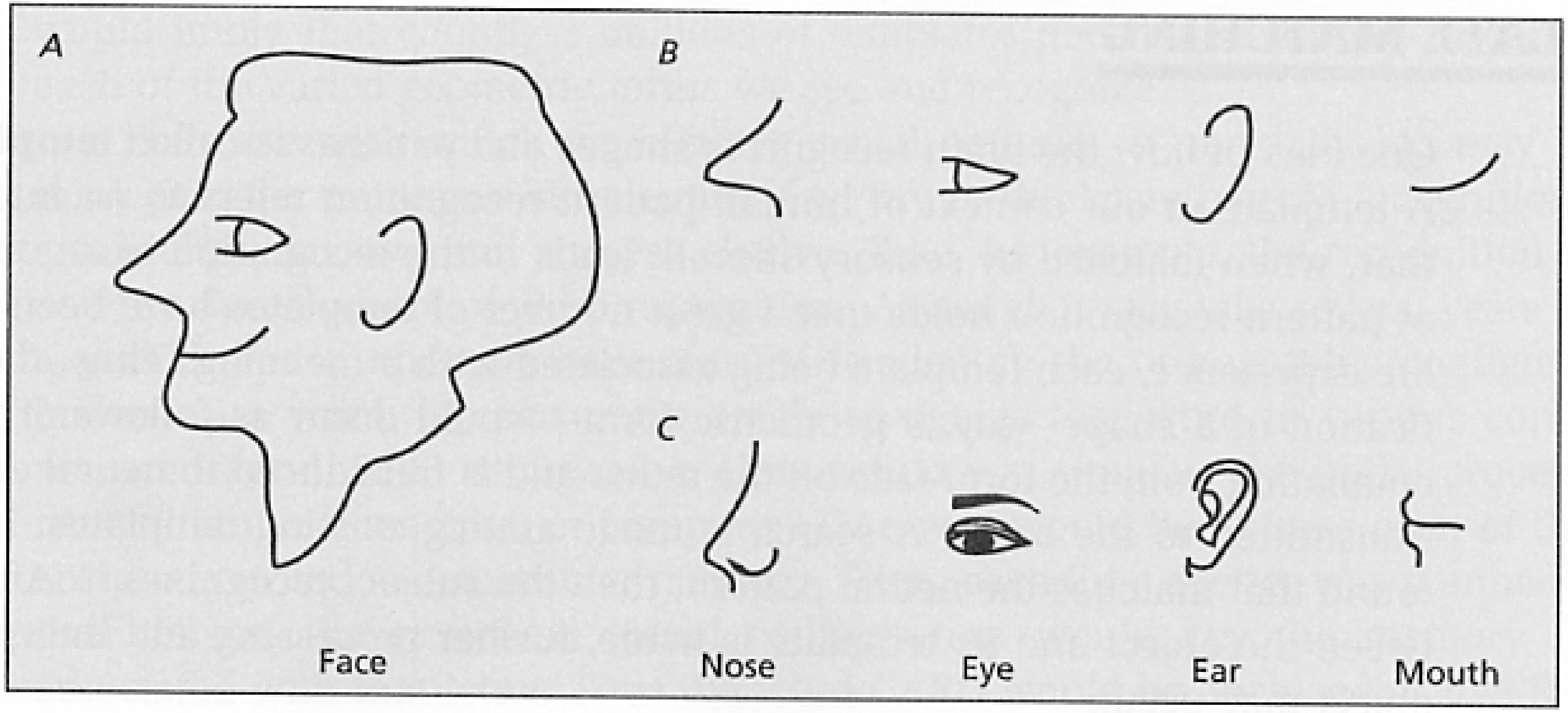
Cognitive Impenetrability

Context effects on Perception

- Parts simultaneously constrain each others' perception
- Perception: finding a globally harmonious interpretation of many parts
- Examples
  - McGurk effect: vision automatically influences audition
  - ABC/11 12 13. Fruit faces
  - For holistic objects, perception of parts influenced by context (Farah, 1992)
  - The word superiority effect (McClelland & Rumelhart, 1981; Wheeler, 1972)
- Perceptual chunks
  - Unitization: creating a larger chunk by reliably pairing parts (Goldstone, 2000; Shiffrin & Lightfoot, 1998)
  - Differentiation: breaking down a single chunk into pieces if the pieces vary independently (Burns & Shepp, 1978; Goldstone & Steyvers, 2001; Smith, 1978)
  - Perceptual modules: large within-chunk dependencies, small between-chunk dependencies

A

B


A
B
C
Face
Nose
Eye
Ear
Mouth

Archimbaldo

Context effects on perception (Farah, 1992)

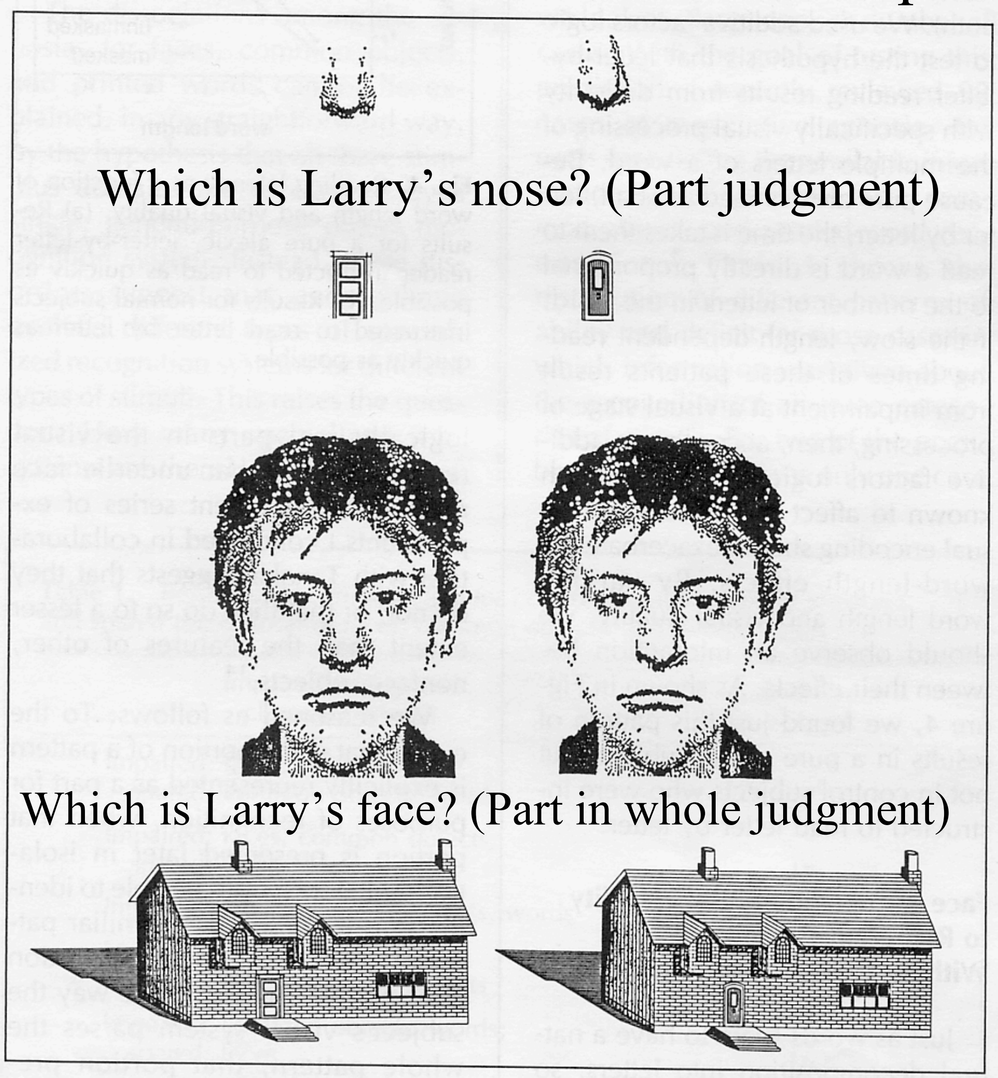
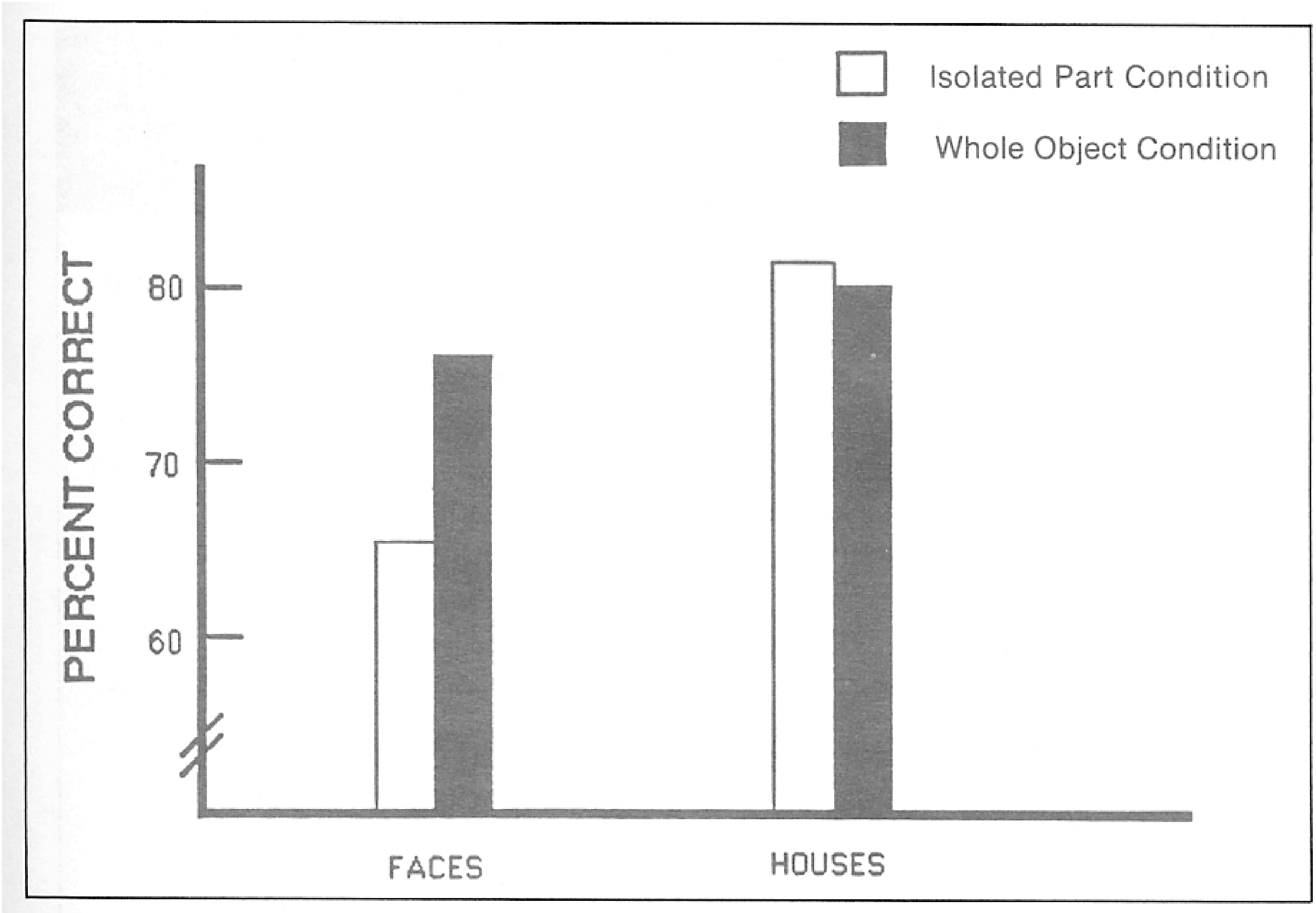


Fig. 5. Examples of pairs of test items from an experiment on the recognition of faces and houses. Subjects studied whole items individually and learned to identify them by name (e.g., Larry's face or Larry's house). The test was administered in a two-alternative forced-choice format, either for an isolated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



Word Superiority Effect

WORK

---K
Or ?
---D

XXXK

---K
Or ?
---D

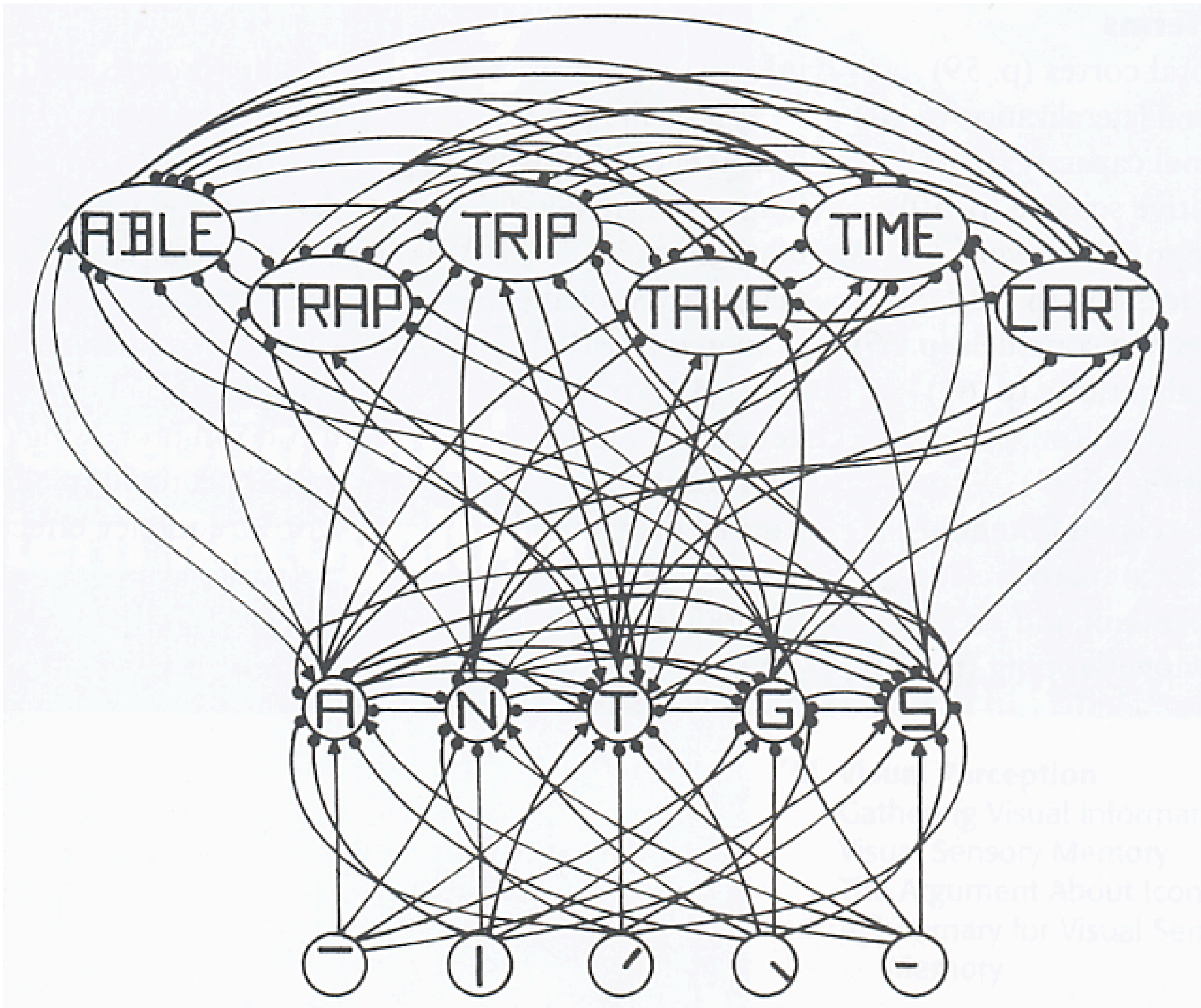
Subjects are more likely to choose the correct letter when it is in the context of a word than when it is isolated.

Context improves people's sensitivity, not just bias

Aspects of the Word Superiority Effect (WSE)

- Letters better identified in words than in non-words or by themselves
- Words as perceptual units
- Sensitivity, not bias, effect
  - Bias to respond with letter that would form a real word cannot explain WSE because both letter choices form a real word
- Pseudo-word superiority effect too
  - E in MAVE is better identified than E in VMAE
- Pattern mask is important for WSE

Interactive Activation Model
McClelland & Rumelhart (1981)



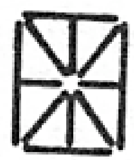
Interactive Activation Model (IAT)

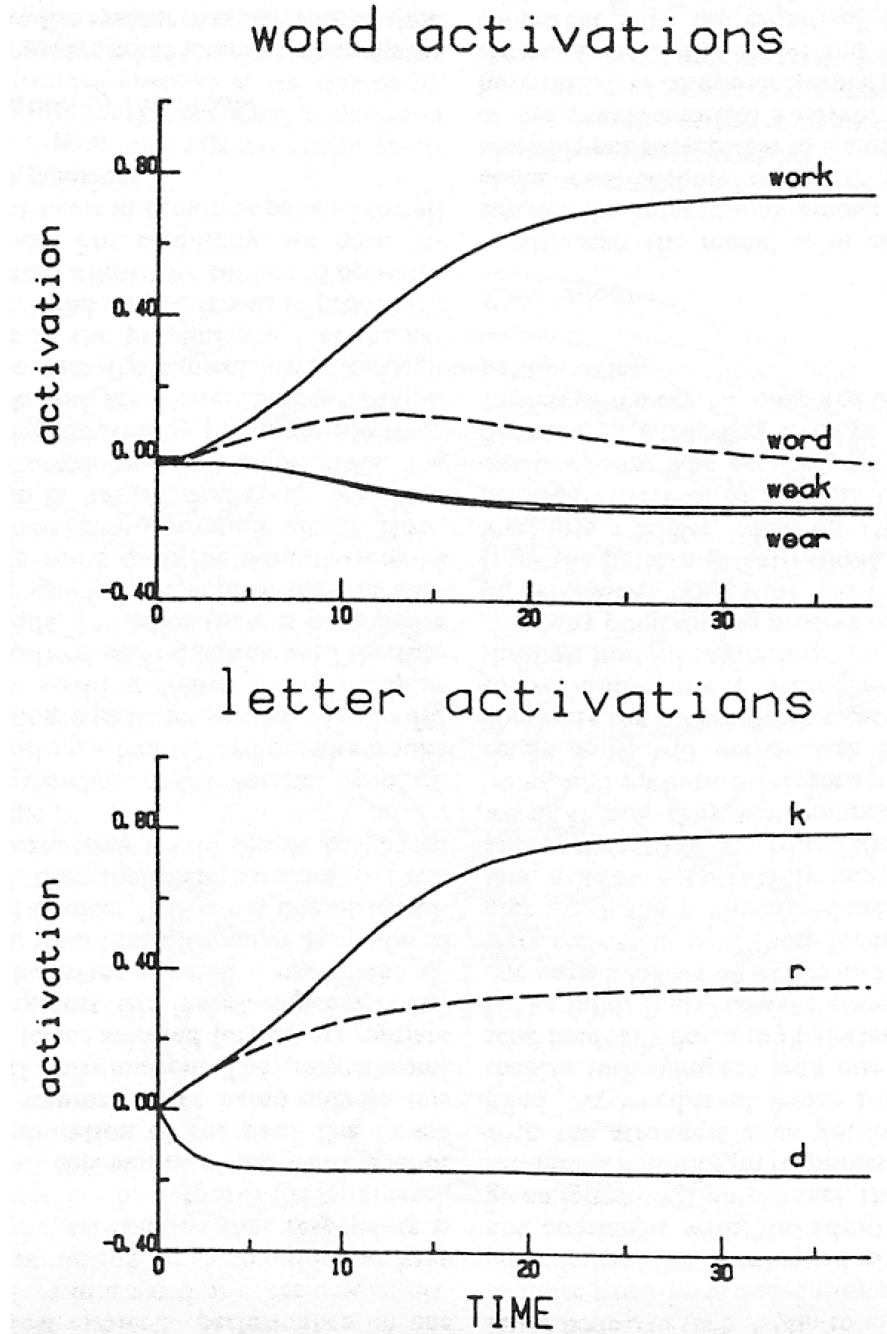
- Cascading activation
  - Feature-level processing not complete before higher-levels start
  - Top-down <u>and</u> bottom-up is not viciously circular
  - Contrast to standard information processing
- Architecture
  - Feature, Letter, and Word level units
  - Activation between levels, inhibition within levels
  - Lateral inhibition
    - Good for creating discrete edges, category, decisions
    - Digitalization
- Processing: flow of activation/inhibition along links

ABCDEFGHI

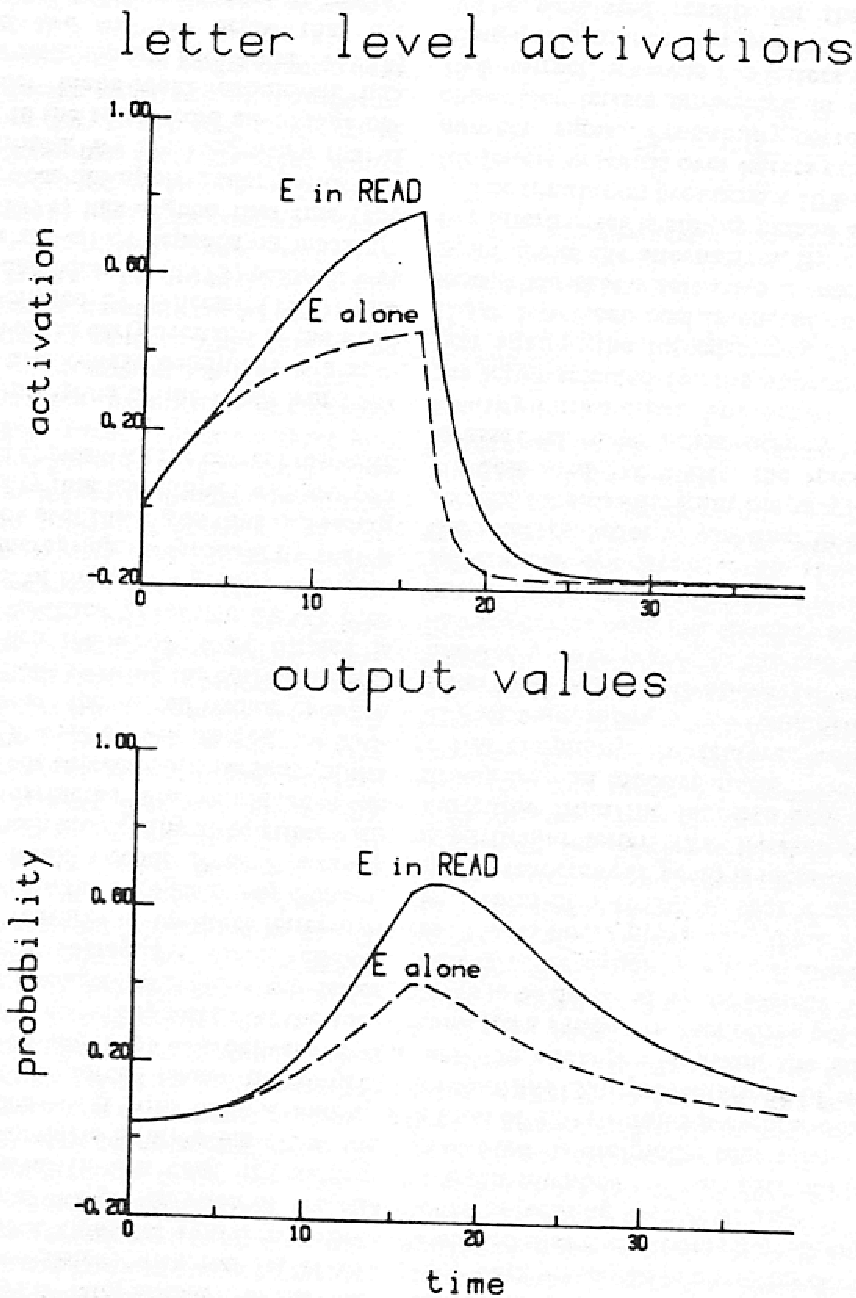
JKLMNOPQR

STUVWXYZ





word activations
activation
0.80
0.40
0.00
-0.40
0
10
20
30
work
word
weak
wear
letter activations
activation
0.80
0.40
0.00
-0.40
0
10
20
30
k
r
d
TIME



letter level activations
activation
E in READ
E alone
1.00
0.60
0.20
-0.20
0
10
20
30
output values
probability
E in READ
E alone
1.00
0.60
0.20
-0.20
0
10
20
30
time

$$n_i(t) = \square_{j} \square_{ij} e_j(t) \square \square_{k} \square_{ik} i_k(t)$$

n= net input, □=weight of excitatory input, e =activation of incoming excitatory node.

□=weight of inhibitory input, i = activation of incoming inhibitory node.

Net input to node is based on consistent and inconsistent inputs.

$$E_i(t) = n_i(t)\left(M - a_i(t)\right) if\, n_i(t) > 0$$

E_i = effect on node i, M = maximum activation possible, a_i = activation of node i.

Effect on node is based on input, but has a ceiling at M. The closer the current activity of the node is to M, the less the effect of positive input will be.

$$E_i(t) = n_i(t)\left(a_i(t) - m\right) if\, n_i(t) < 0$$

m = minimum activation possible.

If input is negative, then the floor is at m. If current activity is already at floor, then input has no effect.

$$a_i(t+\Delta t) = a_i(t) - \Theta_i\left(a_i(t) - r_i\right) + E_i(t)$$

Θ=rate of decay, r = resting level of unit

New activity is based on old activity, and decay to a resting level, and the effect of the input to the node.

$$\bar{a}_i(t) = \int_{-\infty}^{t} a_i(x) e^{-(t-x)r} dx$$

a-bar = running average of activation, decay of old information.
A cumulative average across time of a unit's activity will be its strength.
More recent activity levels matter more than older activity levels, and the decay rate of old information is based on r.

$$s_i(t) = e^{\square a_i(t)}$$

Si = response strength of unit i, □=steepness of function relating activation to response.

Exponential functions serve to emphasize differences between larger quantities, which is important because activations are capped at 1. The difference between .8 and .9 should be greater than between .7 and .8.

$$p(R_i, t) = \frac{s_i(t)}{\sum_{j \in L} s_j(t)}$$

P(Ri,t) = probability of responding with unit i's response
L = set of nodes at same level as i.

Luce choice rule: if you have N alternatives that are mutually exclusive (can only do one of them), then this rule assures that the probabilities will add up to one, and the probability of making a response is based on its relative strength.

$$p\left(R_i\right) = \frac{e^{\square s_i}}{\sum_{j \in L} e^{\square s_j}}$$

Generalized Choice Rule

$S_1 = 0.7 \qquad S_2 = 0.4$

If □=1, $P(S_1)$= 2.01/(2.01+1.49) = .574
$P(S_2)$= 1.49/(2.01+1.49) = .426

If □=10, $P(S_1)$= .953
$P(S_2)$= .047

As □ increases, the choice becomes increasingly deterministic
Probability matching (small □) versus reward maximization (big □)

the "rich get richer" effect

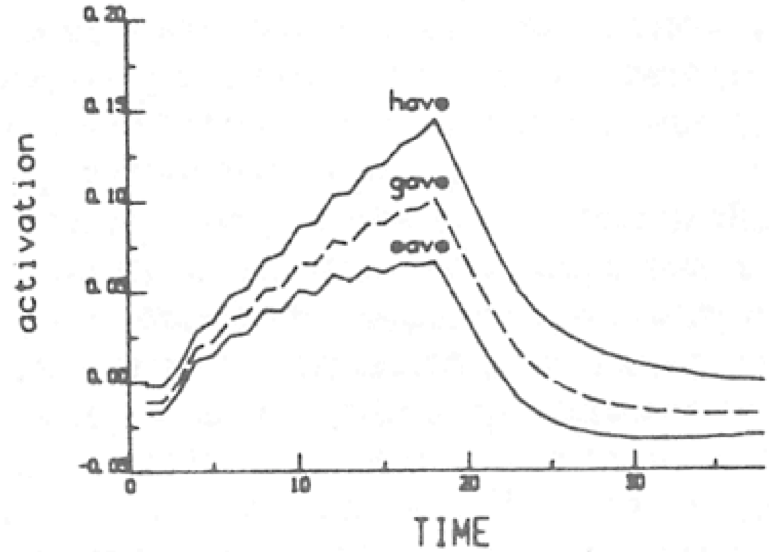


Figure 11. The rich-get-richer effect. (Activation functions for the nodes for *have*, *gave*, and *save* under presentation of *MAVE*.)

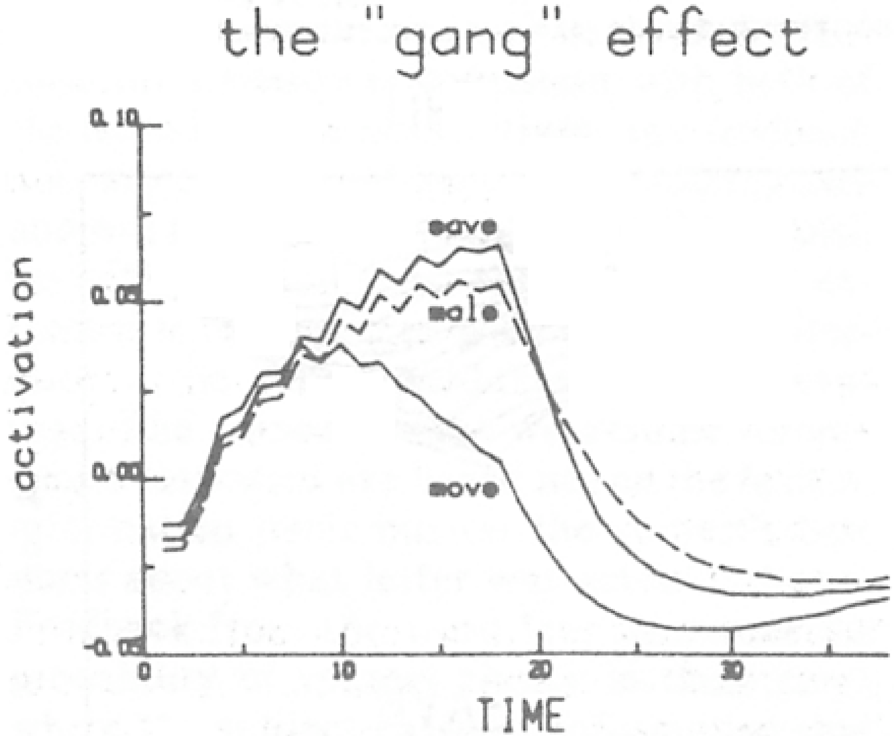


Figure 12. The gang effect. (Activation functions for *move*, *male*, and *save* under presentation of *MAVE*.)