Introduction to Cognitive Science: Attention

Adapted from Albert Ali Salah's slides

The firehose of experience

Focus on relevant information

Attention in Cognitive Science

- Monolithic process?
- specific locations
- in vision, audition, or touch;
- entire objects, not just locations;
- mental functions such as behavioral goals or tasks:

- To understand attention requires studying these multiple forms of selection, their similarities and dissimilarities, and their neural foundations.

Overview

- Attention and related concepts
- Human visual attention
- Computational models of attention
- Applications of computational models of attentional

What is attention?

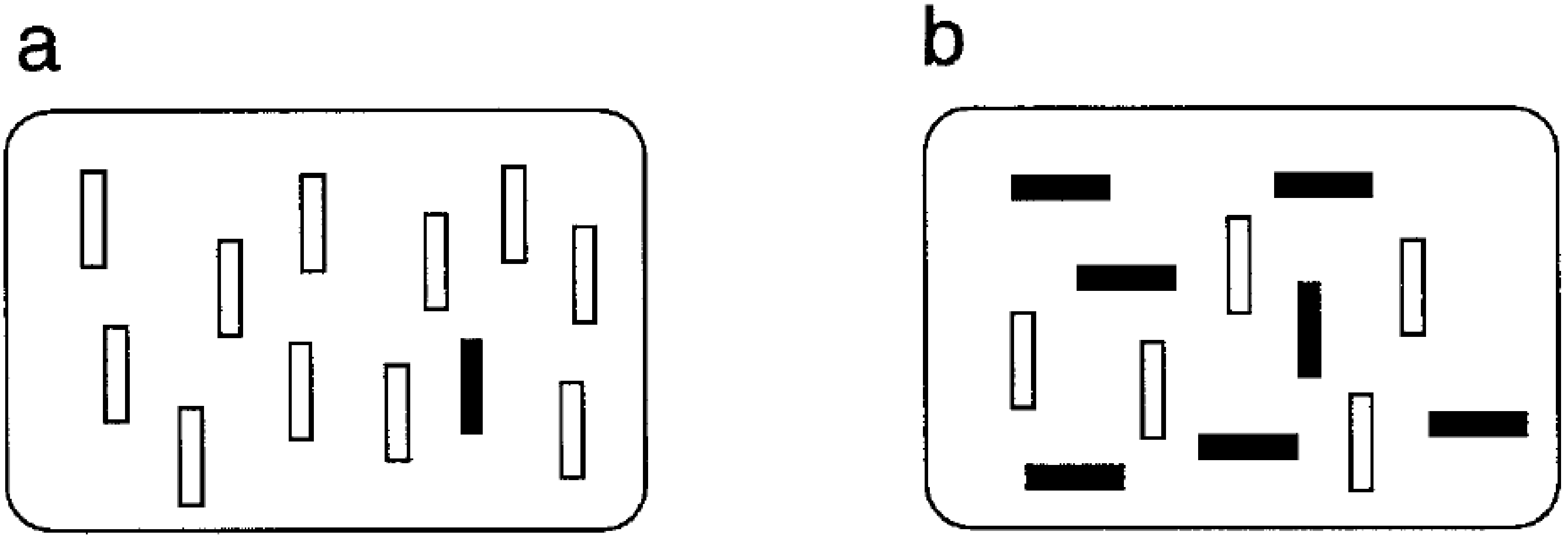
- “Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought. Focalization, concentration, of consciousness are of its essence. It implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state which in French is called distraction, and Zerstreutheit in German.”

William James, 1890

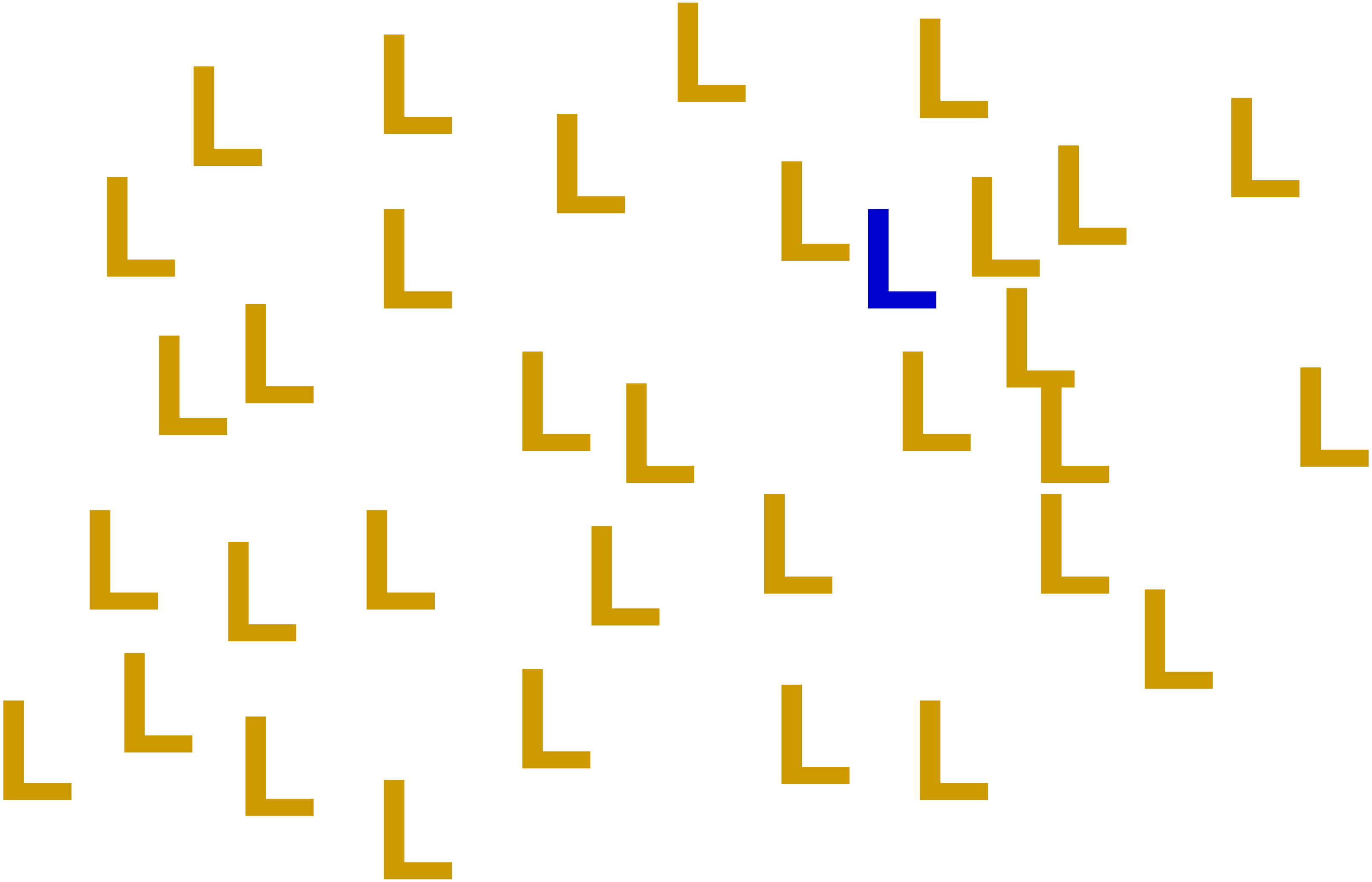
Why is attention important?

- [Cognitive Science] Posner & Petersen (1990): “The importance of attention is its unique role in connecting the mental level of description of processes used in cognitive science with the anatomical level common in neuroscience.”
- [Computer Science] Some cognitive tasks are very heavy in terms of computational resources. For instance visual search at all locations and across all scales is very time consuming. (Tsotsos et al., 1995)

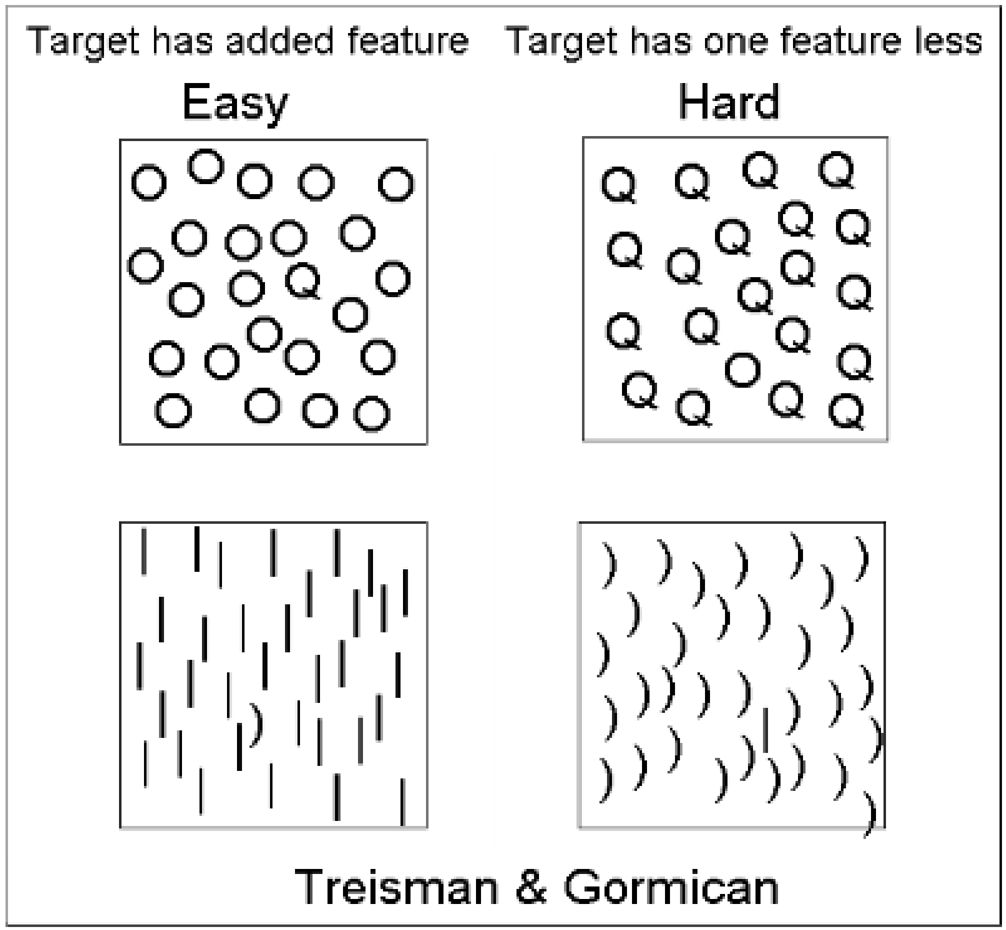
Competition between bottom-up and top-down inputs



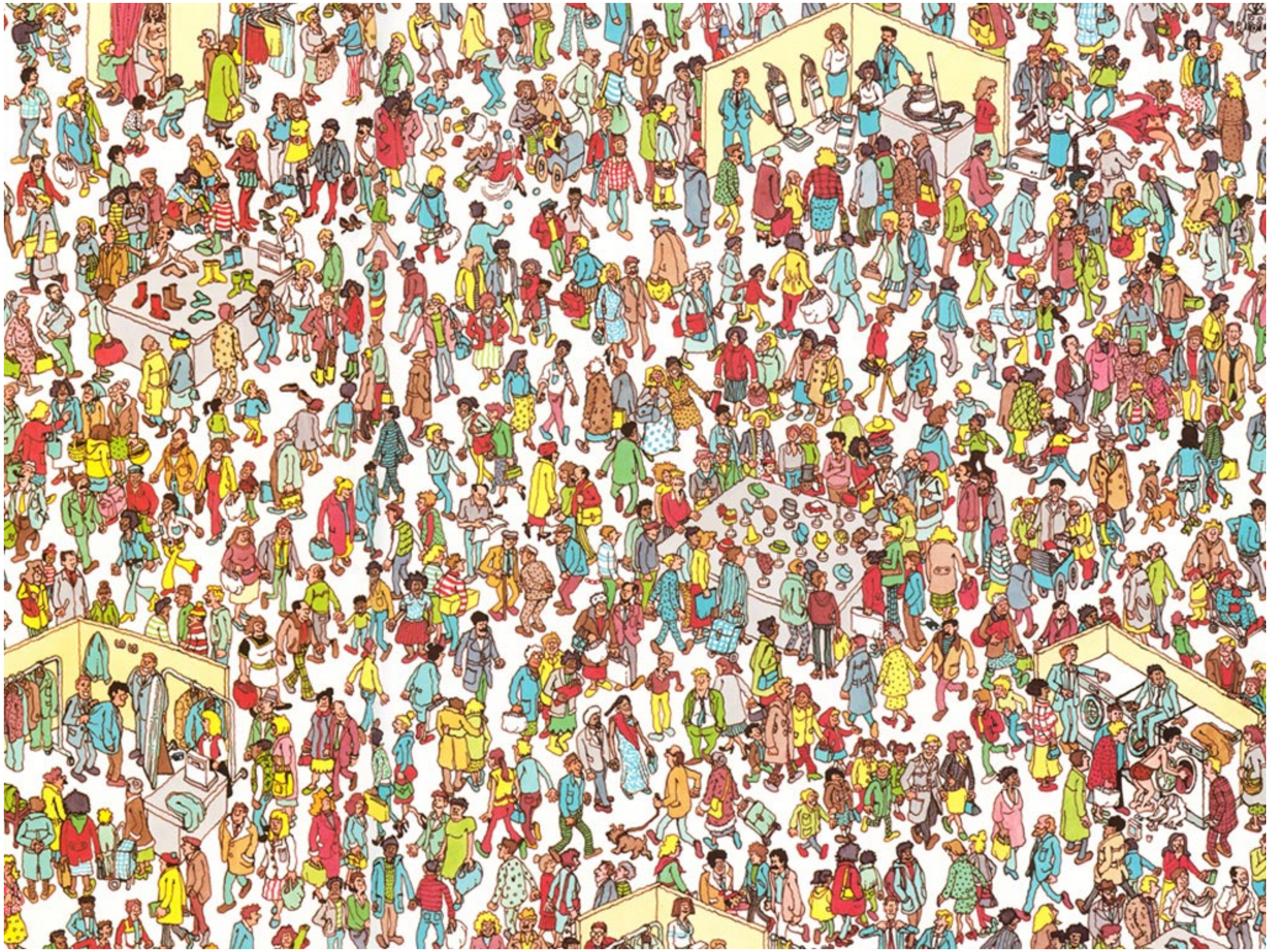
Saliency



Visual Search

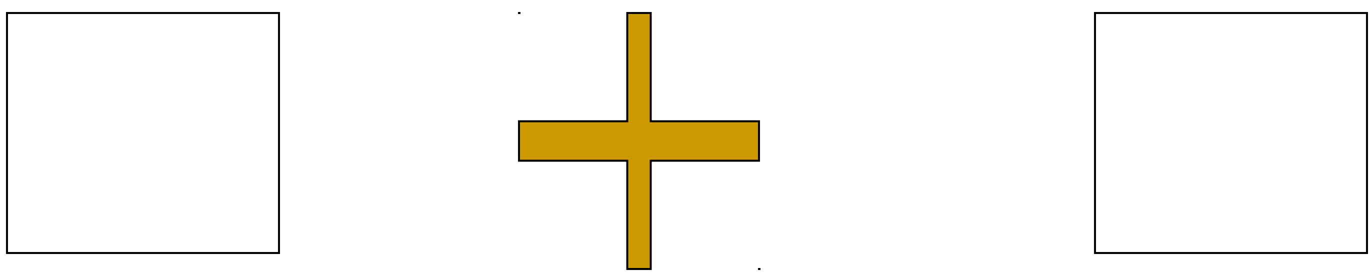


Where is Waldo? (or Wally?)

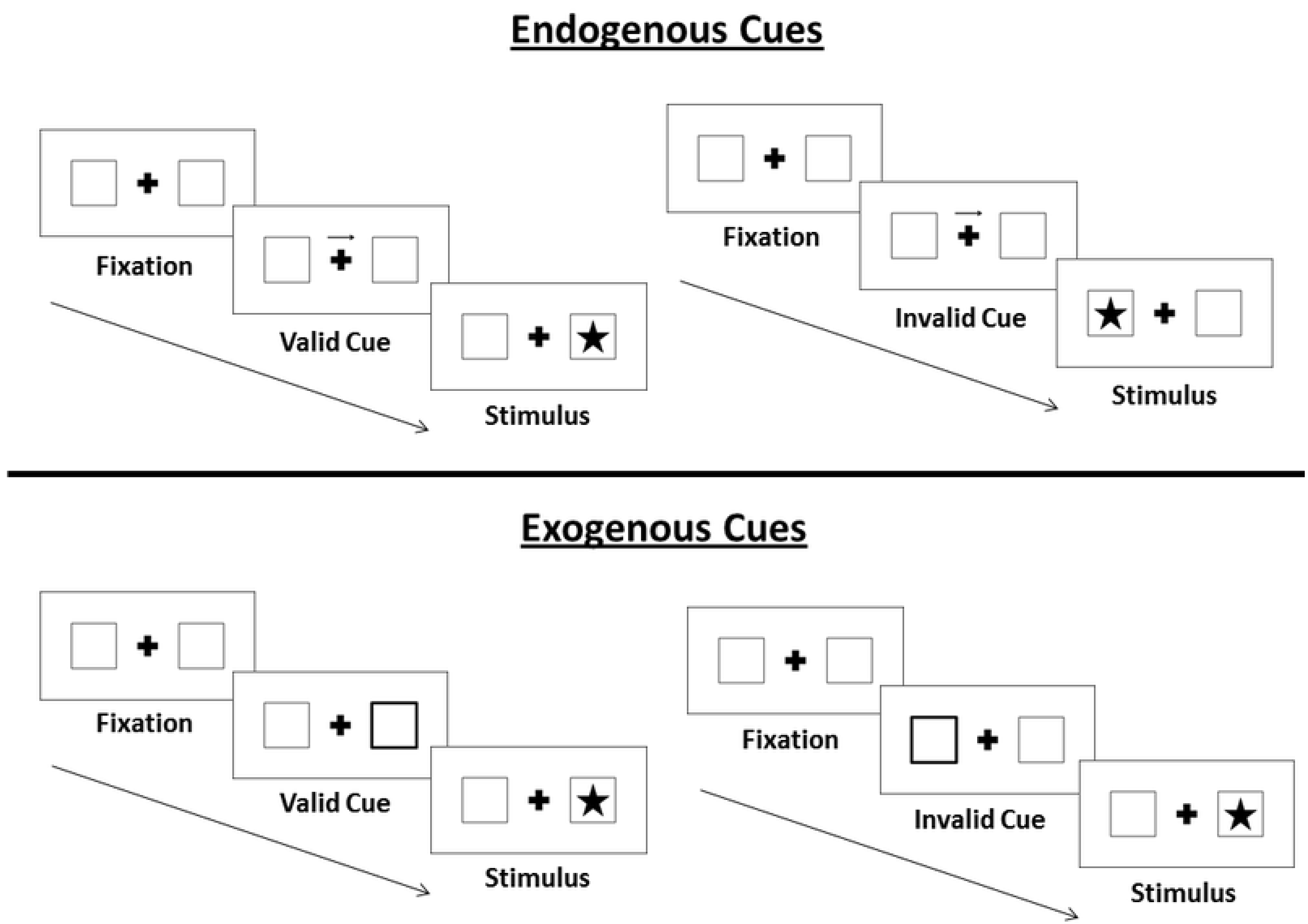


Orienting Attention

- Posner Cue - Target Paradigm:

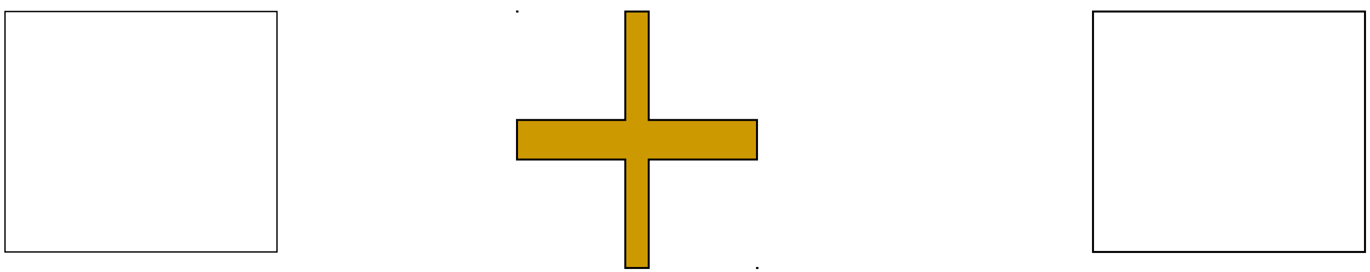


Subject presses a button as soon as x appears


Endogenous Cues
Fixation
Valid Cue
Stimulus
Fixation
Invalid Cue
Stimulus
Exogenous Cues
Fixation
Valid Cue
Stimulus
Fixation
Invalid Cue
Stimulus

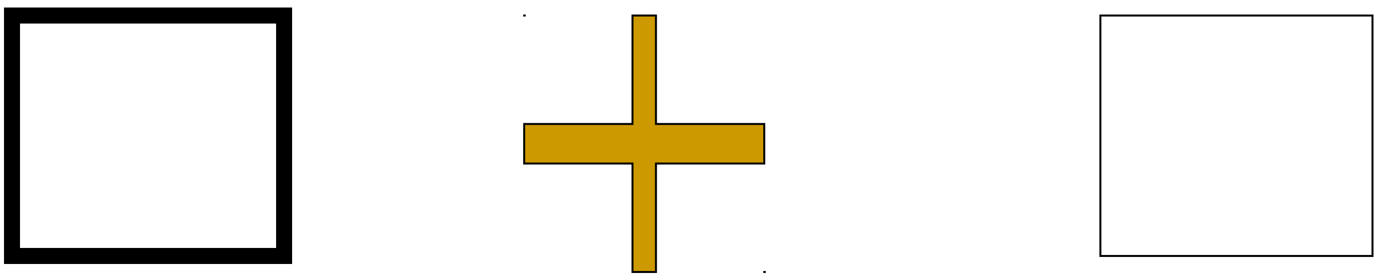
Orienting Attention

- Posner Cue - Target Paradigm:



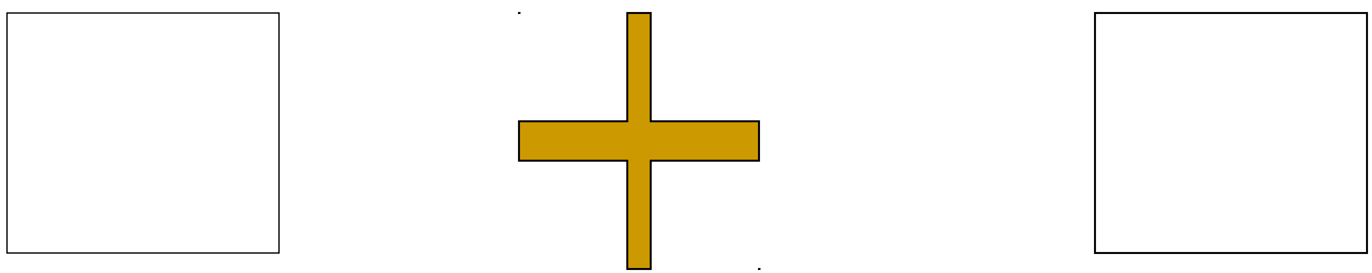
Orienting Attention

- Posner Cue - Target Paradigm:



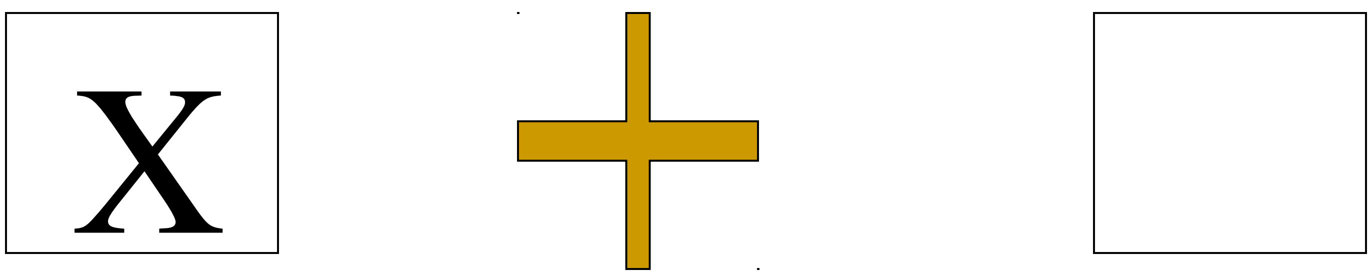
Orienting Attention

- Posner Cue - Target Paradigm:



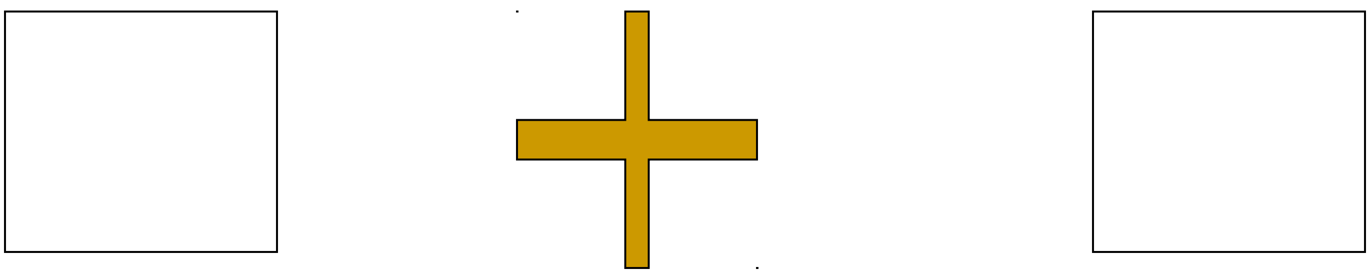
Orienting Attention

- Posner Cue - Target Paradigm:



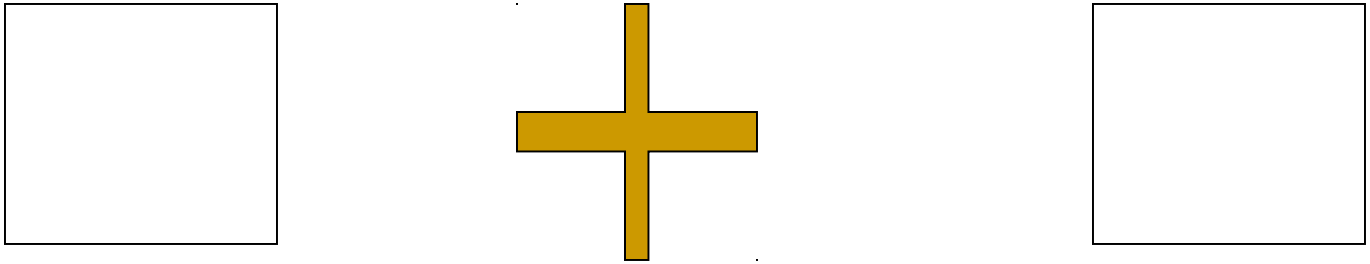
Orienting Attention

- Posner Cue - Target Paradigm:



Orienting Attention

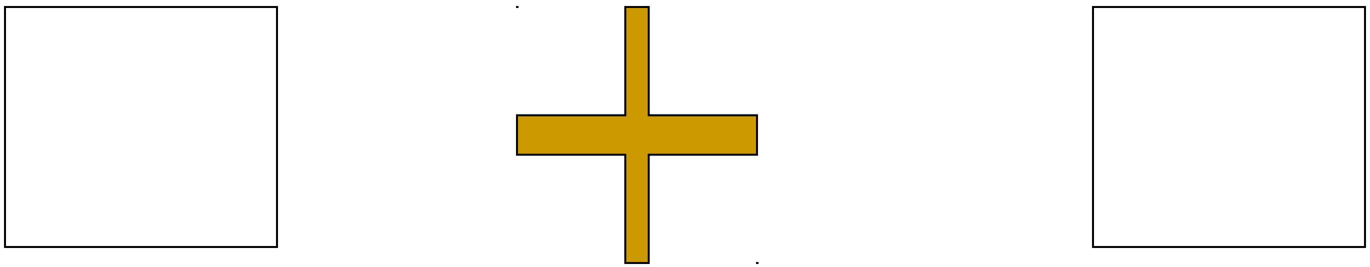
- Posner Cue - Target Paradigm:



That was a validly cued trial because the
x appeared in the box that flashed

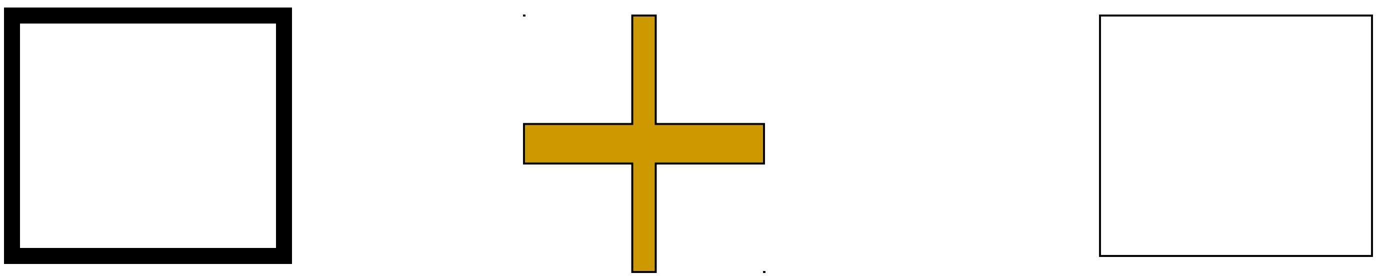
Orienting Attention

- Posner Cue - Target Paradigm:



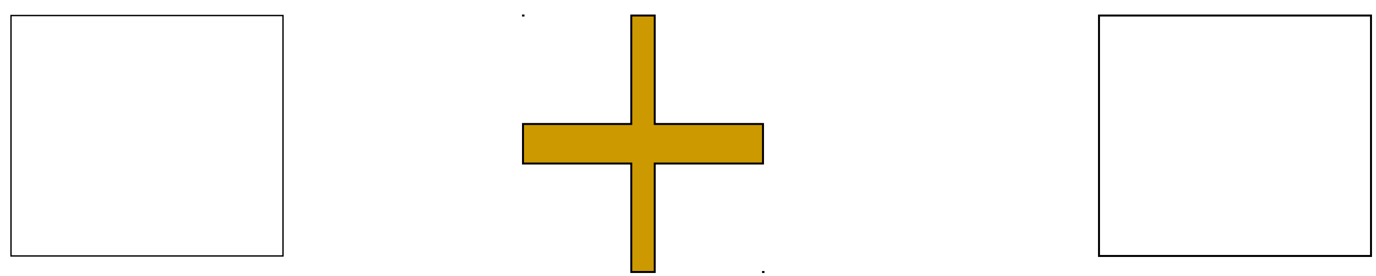
Orienting Attention

- Posner Cue - Target Paradigm:



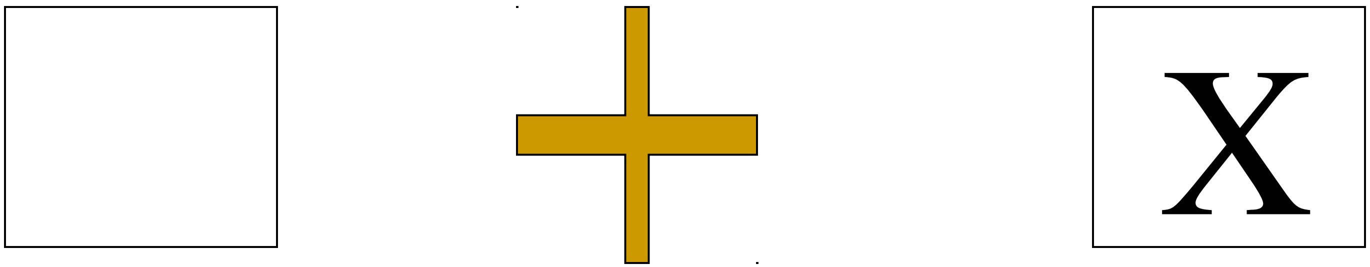
Orienting Attention

- Posner Cue - Target Paradigm:



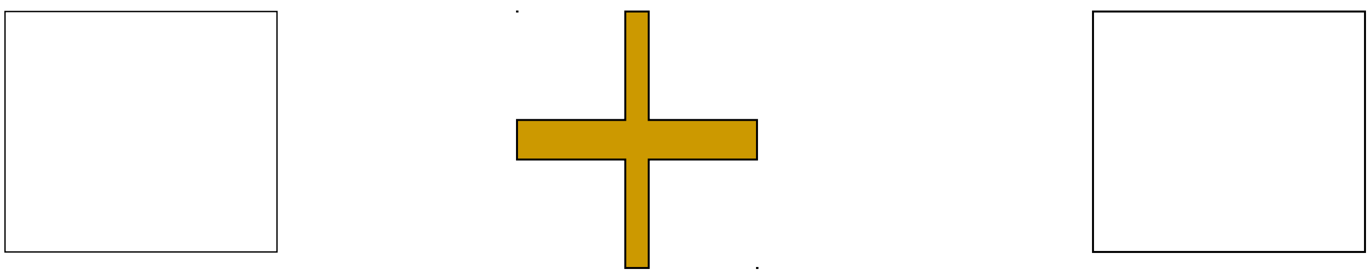
Orienting Attention

- Posner Cue - Target Paradigm:



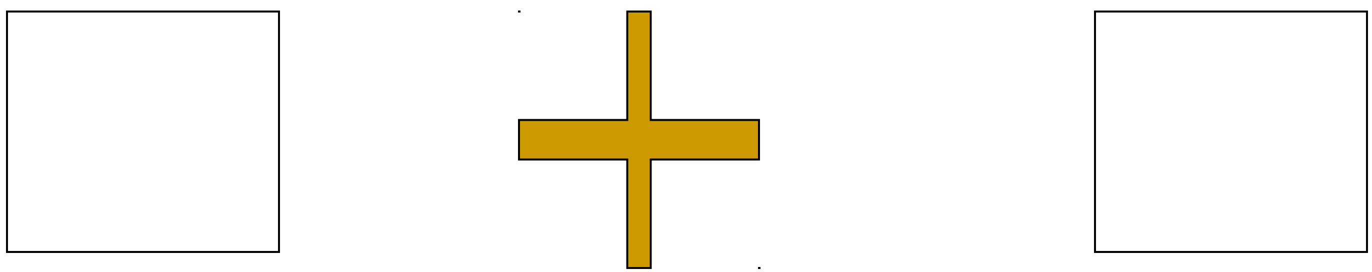
Orienting Attention

- Posner Cue - Target Paradigm:



Orienting Attention

- Posner Cue - Target Paradigm:

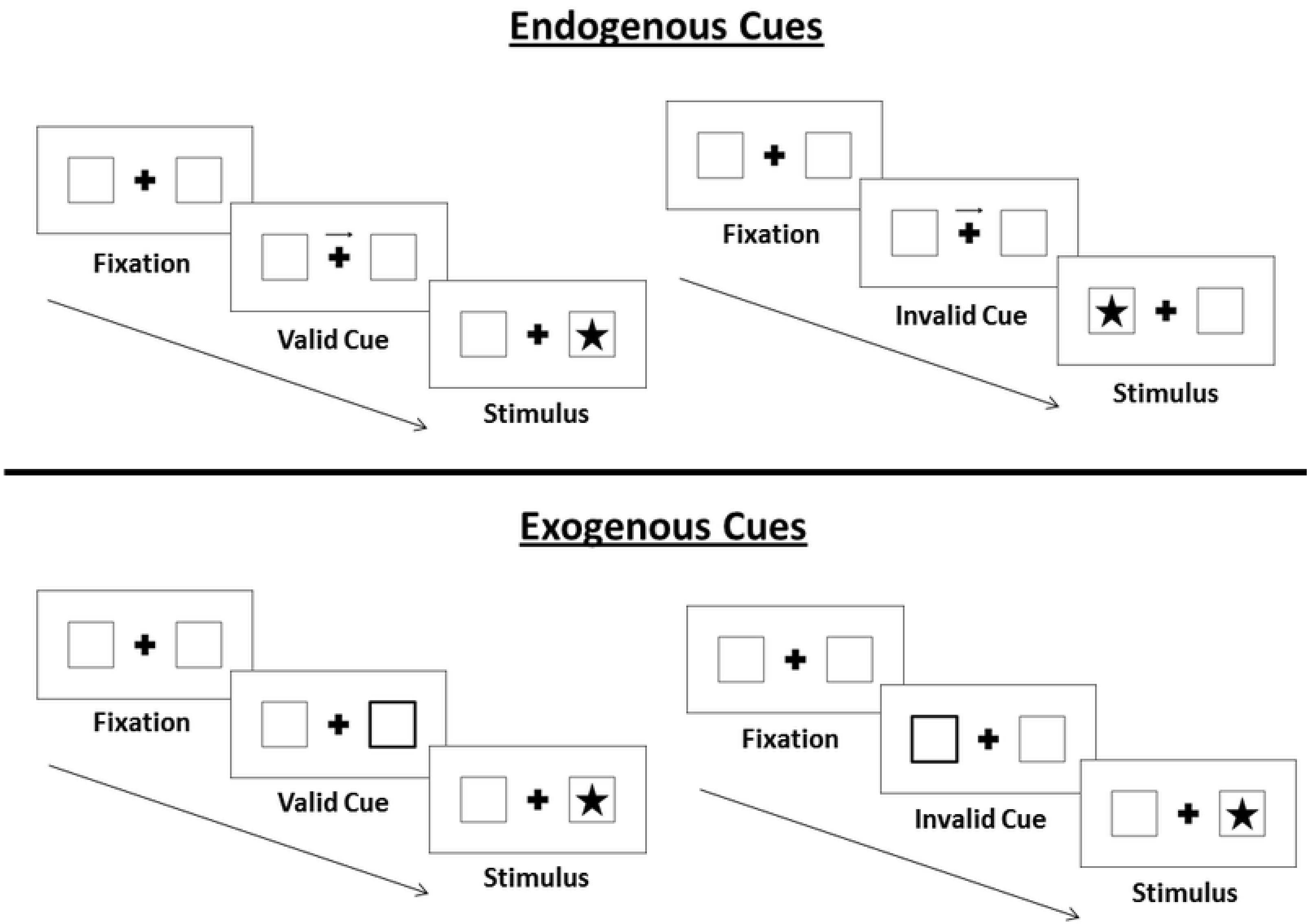


That was an invalidly cued trial because the x appeared in the box that *didn't* flash

Orienting Attention

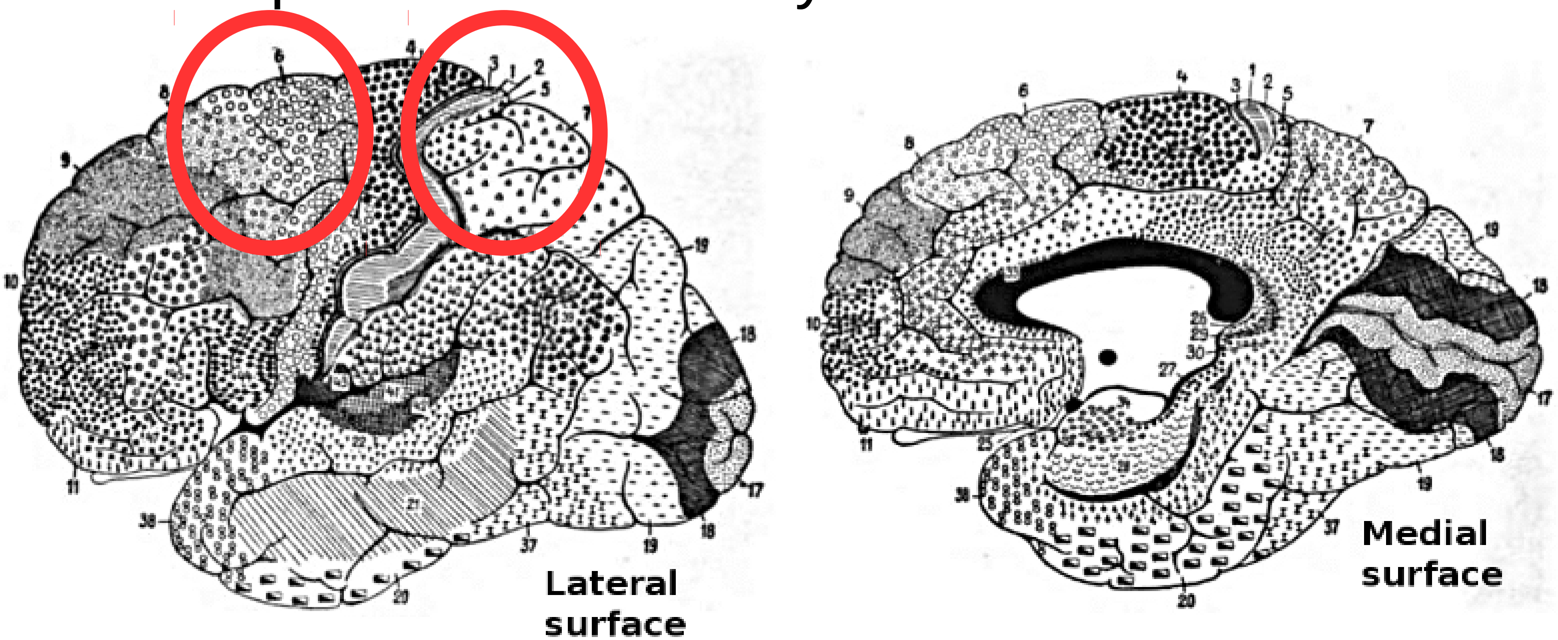
- Posner Cue - Target Paradigm:

$$\text{Attention Effect} = \text{Valid RT} - \text{Invalid RT}$$


Endogenous Cues
Fixation
Valid Cue
Stimulus
Fixation
Invalid Cue
Stimulus
Exogenous Cues
Fixation
Valid Cue
Stimulus
Fixation
Invalid Cue
Stimulus

PET studies

- superior parietal regions are involved in the exogenous, bottom-up control of spatial attention and that
- superior frontal regions are involved in the endogenous, top-down control of spatial attention
- Involve spatial selection only.

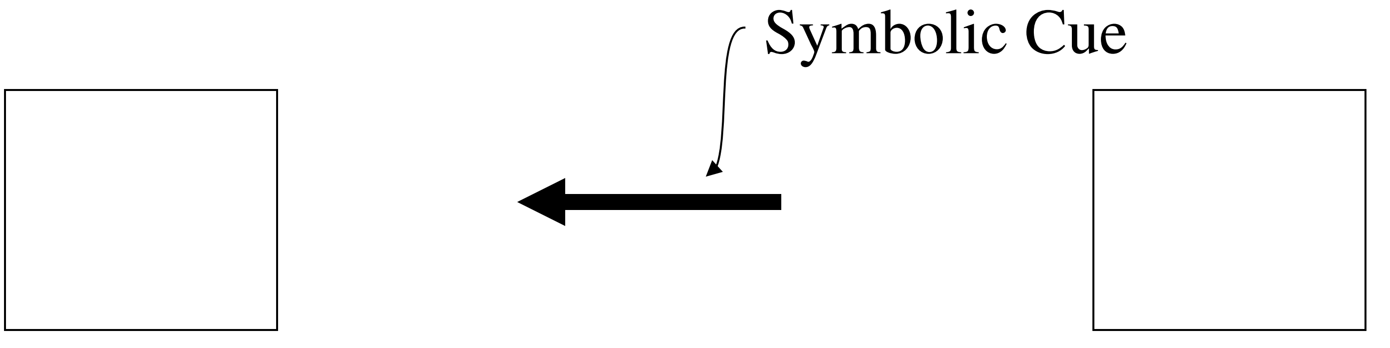


Major distinctions

- Voluntary vs. reflexive attention

Voluntary Orienting

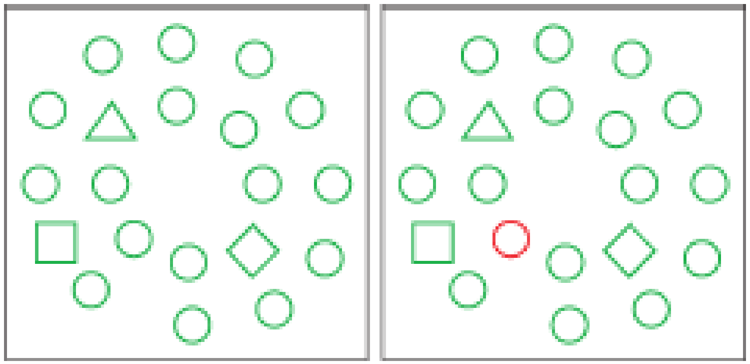
- *Symbolic* cues <u>may</u> orient attention towards *another* location.
- *Stimulus* cues orient attention to the *stimulated* location.



Reflexive Orienting

- Attention can be automatically “summoned” to a location at which an important event has occurred:
  - Loud noise
  - Motion
  - New Object

Transients



(b) Attentional capture

- We call this *attentional capture (or pop-out)*

Major distinctions

- Voluntary vs. reflexive attention
- Overt vs. covert attention
  - Eye movements: 3-5 times per second
  - Overt and covert attention are closely coupled

Eye Tracking

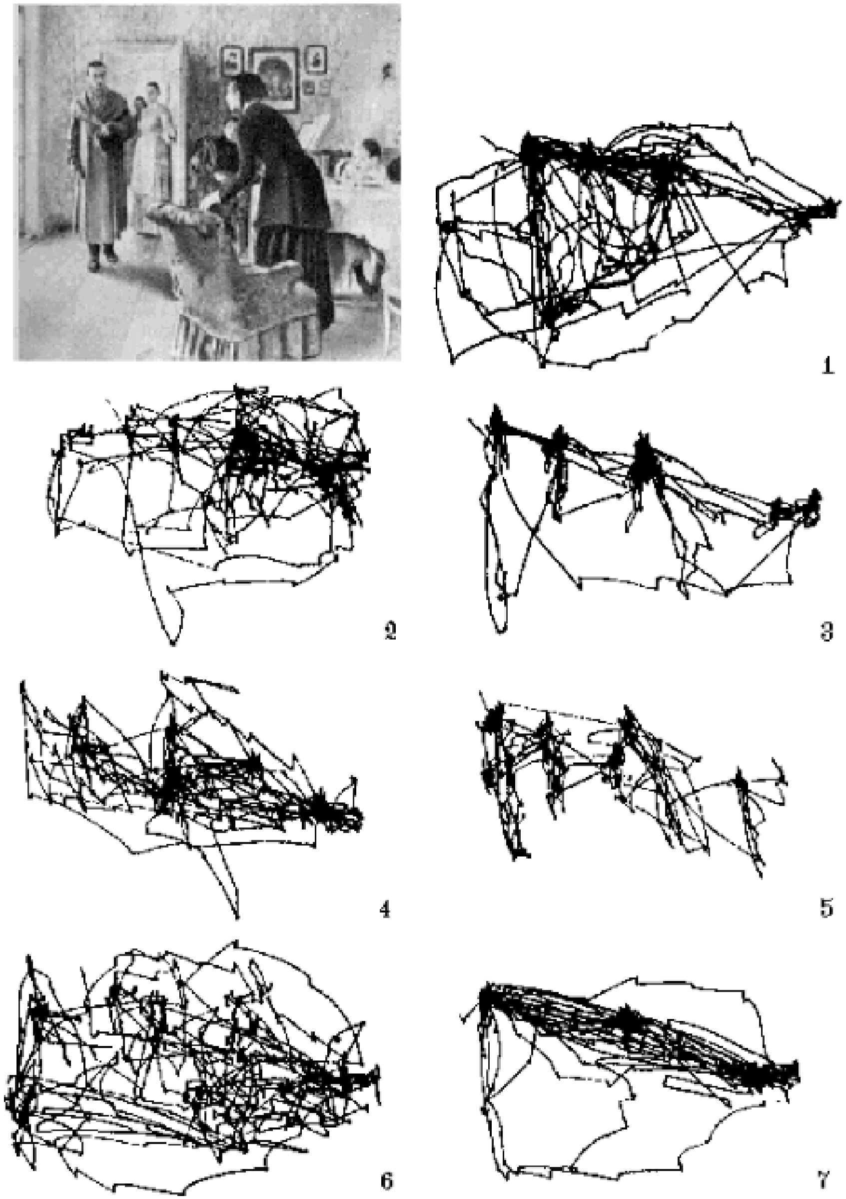
http://www.interactivevideo.com/eye_tracking.html

Major distinctions

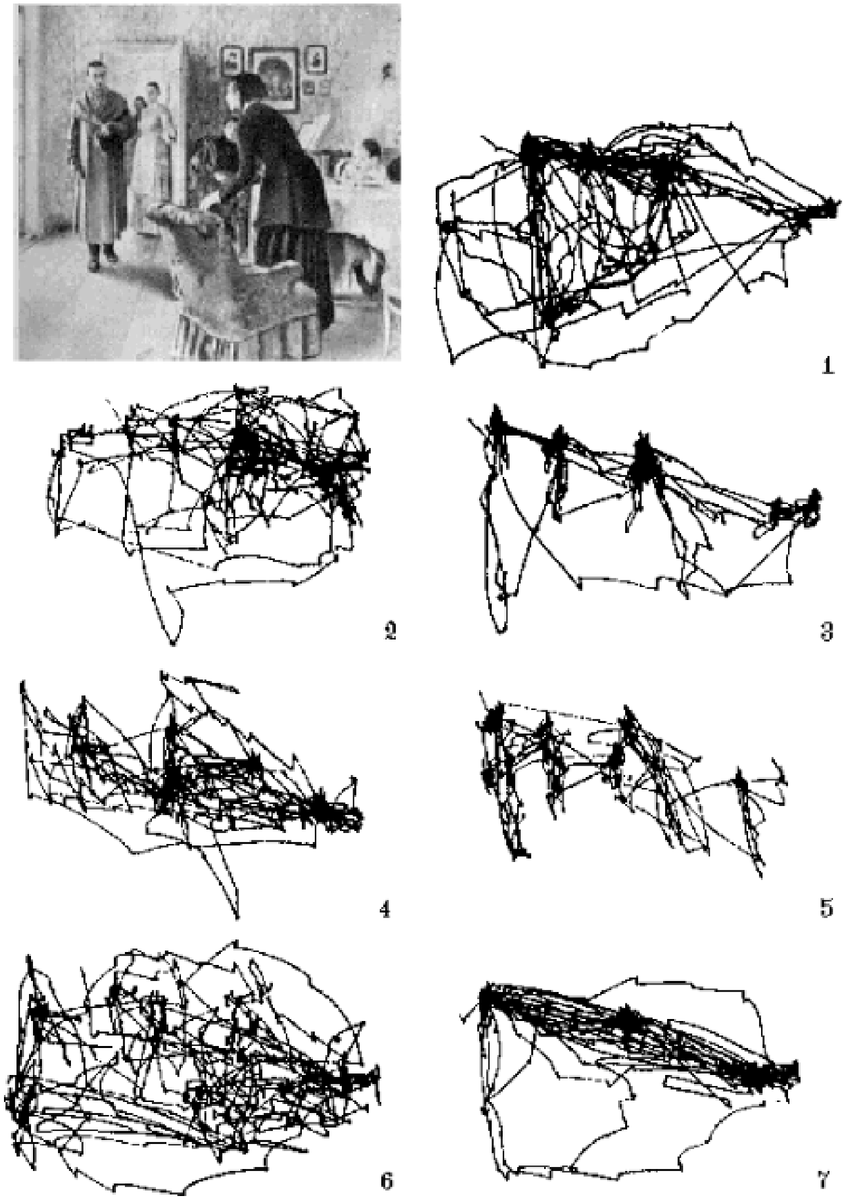
- Voluntary vs. reflexive attention
- Overt vs. covert attention
  - Eye movements: 3-5 times per second
  - Overt and covert attention are closely coupled
- Top-down vs. bottom-up

What draws the attention?

- Low-level features



Yarbus, 1967



1. Free examination.
2. Estimate the material circumstances of the family
3. Give the ages of the people
4. Surmise what the family had been doing before the arrival of the "unexpected visitor"
5. Remember the clothes worn by the people
6. Remember the position of the people and objects in the room
7. Estimate how long the "unexpected visitor" had been away from the family

Yarbus, 1967

What draws the attention?

- Low-level features (bottom-up cues)
- Personal and current goals (top-down cues)

Evolution

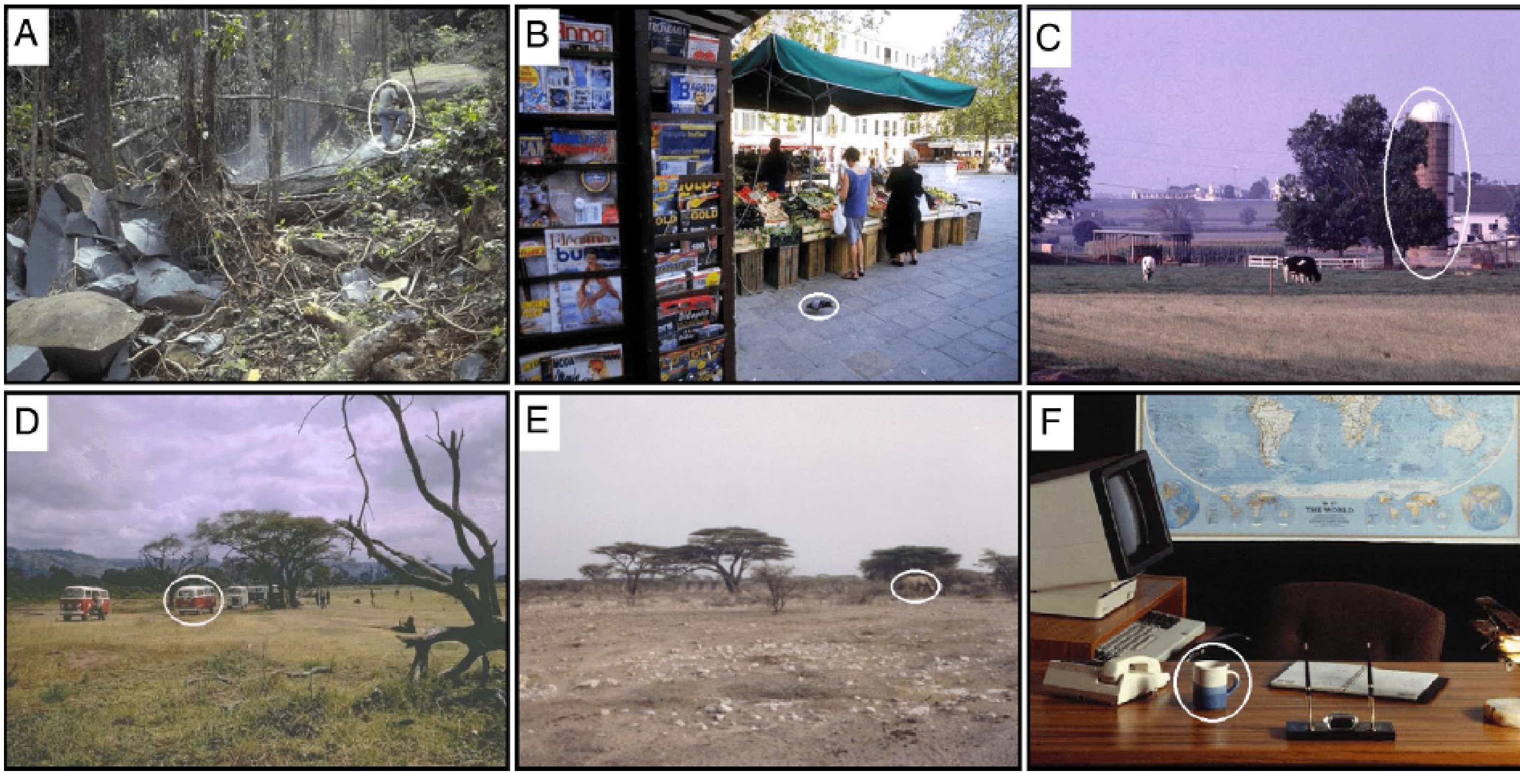


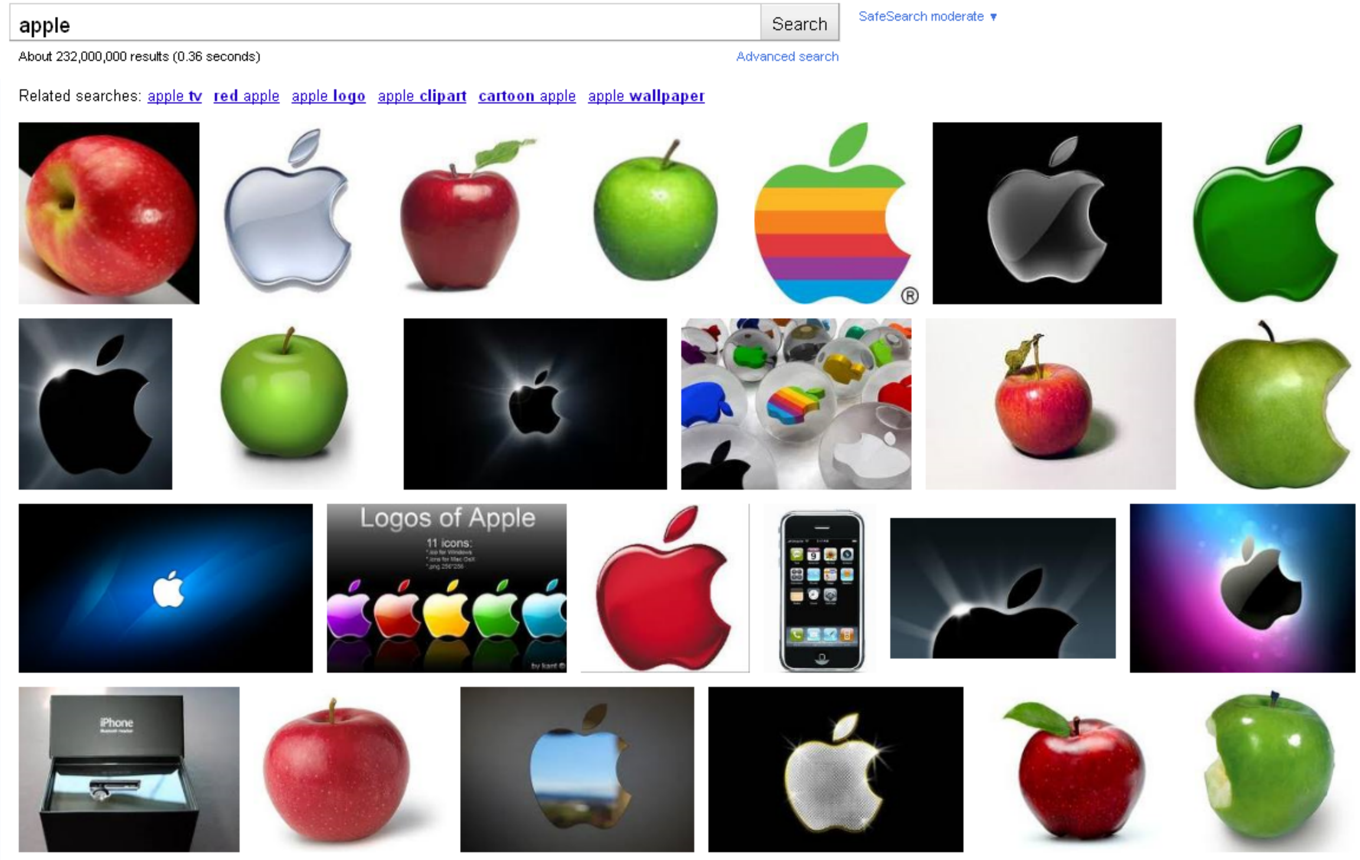
Fig. 2. Sample stimuli with targets circled. Although they are small (measured in pixels), peripheral, and blend into the background, the human (*A*) and elephant (*E*) were detected 100% of the time, and the hit rate for the tiny pigeon (*B*) was 91%. In contrast, average hit rates were 76% for the silo (*C*) and 67% for the high-contrast mug in the foreground (*F*), yet both are substantially larger in pixels than the elephant and pigeon. The simple comparison between the elephant and the minivan (*D*) is equally instructive. They occur in a similar visual background, yet changes to the high-contrast red minivan were detected only 72% of the time (compared with the smaller low-contrast elephant's 100% detection rate).

New, Cosmides, Tooby, PNAS 2007.

What draws the attention?

- Low-level features (bottom-up cues)
- Personal and current goals (top-down cues)
- Evolutionary bias

Natural images?



What draws the attention?

- Low-level features (bottom-up cues)
- Personal and current goals (top-down cues)
- Evolutionary bias
- Learned bias

Counting the balls

http://www.youtube.com/user/profsimons#p/a/u/0/vJG698U2Mvo

Major distinctions

- Voluntary vs. reflexive attention
- Overt vs. covert attention
  - Eye movements: 3-5 times per second
  - Overt and covert attention are closely coupled
- Top-down vs. bottom-up
- Early selection vs. late selection

Stages of Selection

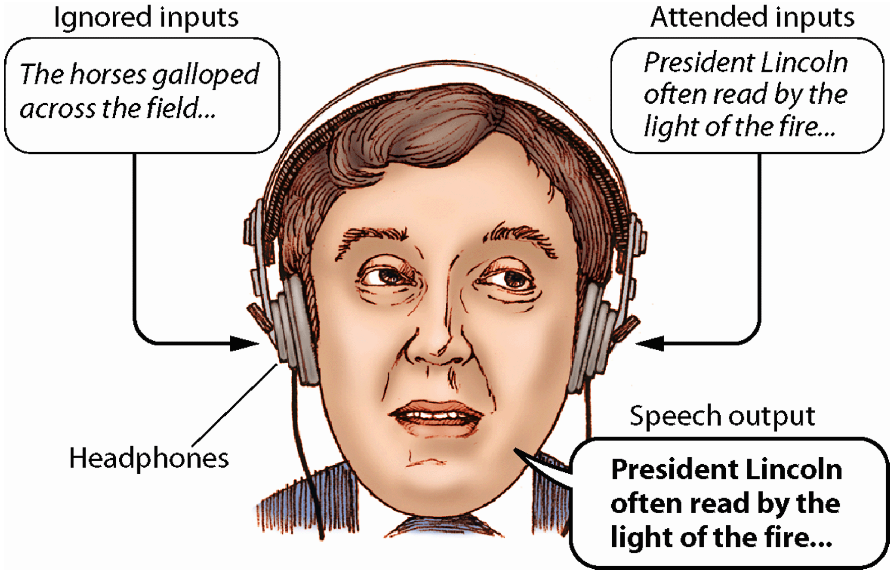
- Broadbent: *Early Selection* - a bottleneck exists early in the course of sensory processing that filters out all but the attended *channel*
- Alternative theory: *Late Selection* - the bottleneck exists not at the lowest stages, but at the highest - such as response planning, memory and consciousness

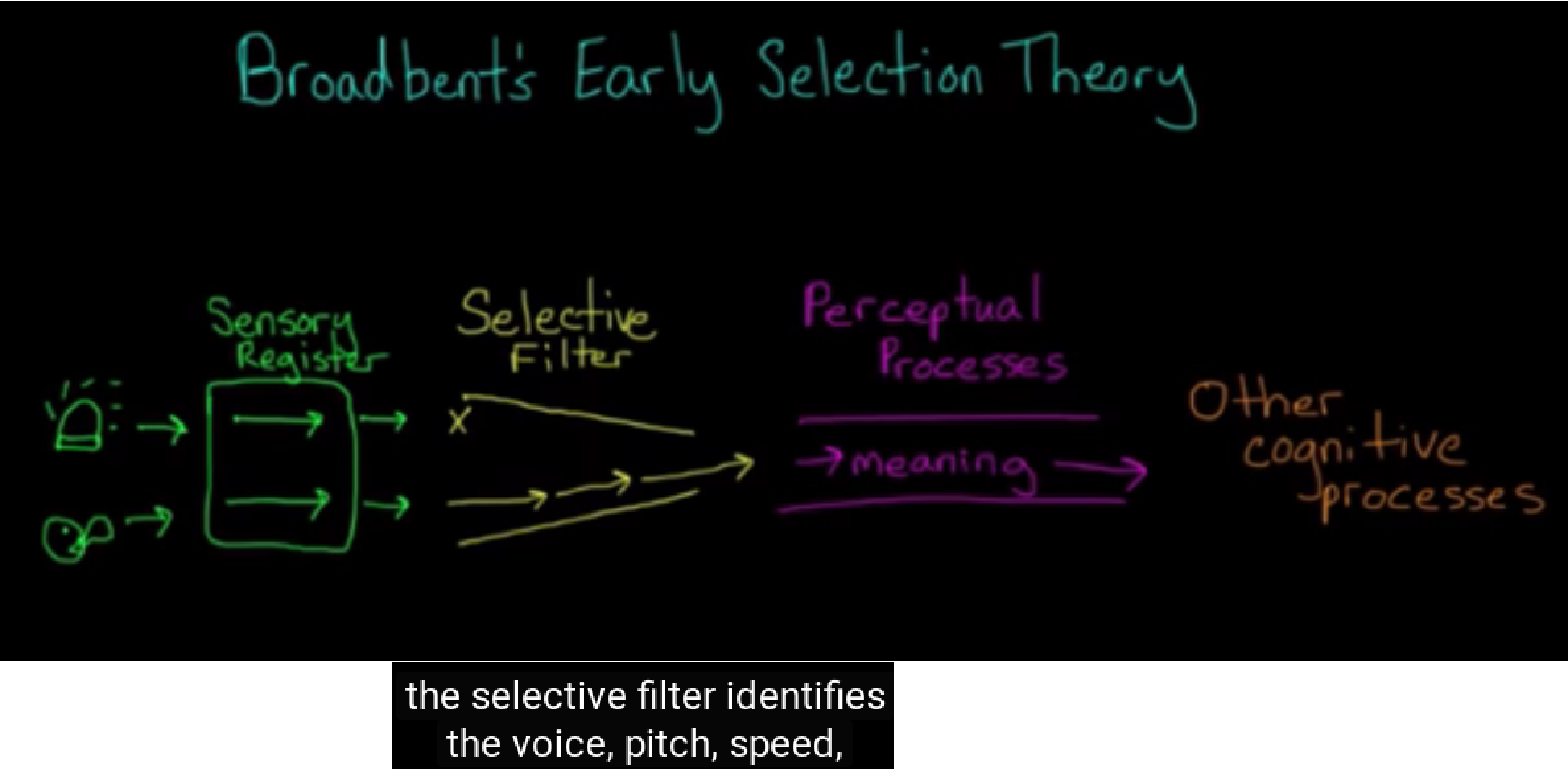
Testing selection

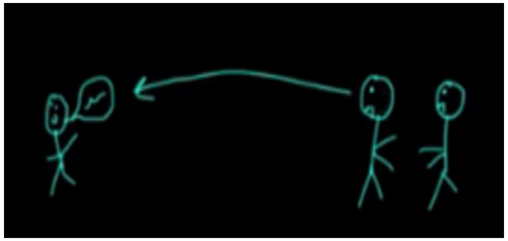
- If early selection hypothesis is correct...
  - Information (such as meaning of words) in unattended channel shouldn’t be processed for meaning
  - We should be able to find differences in brain activity in primary sensory areas (A1, V1)

Testing Selection

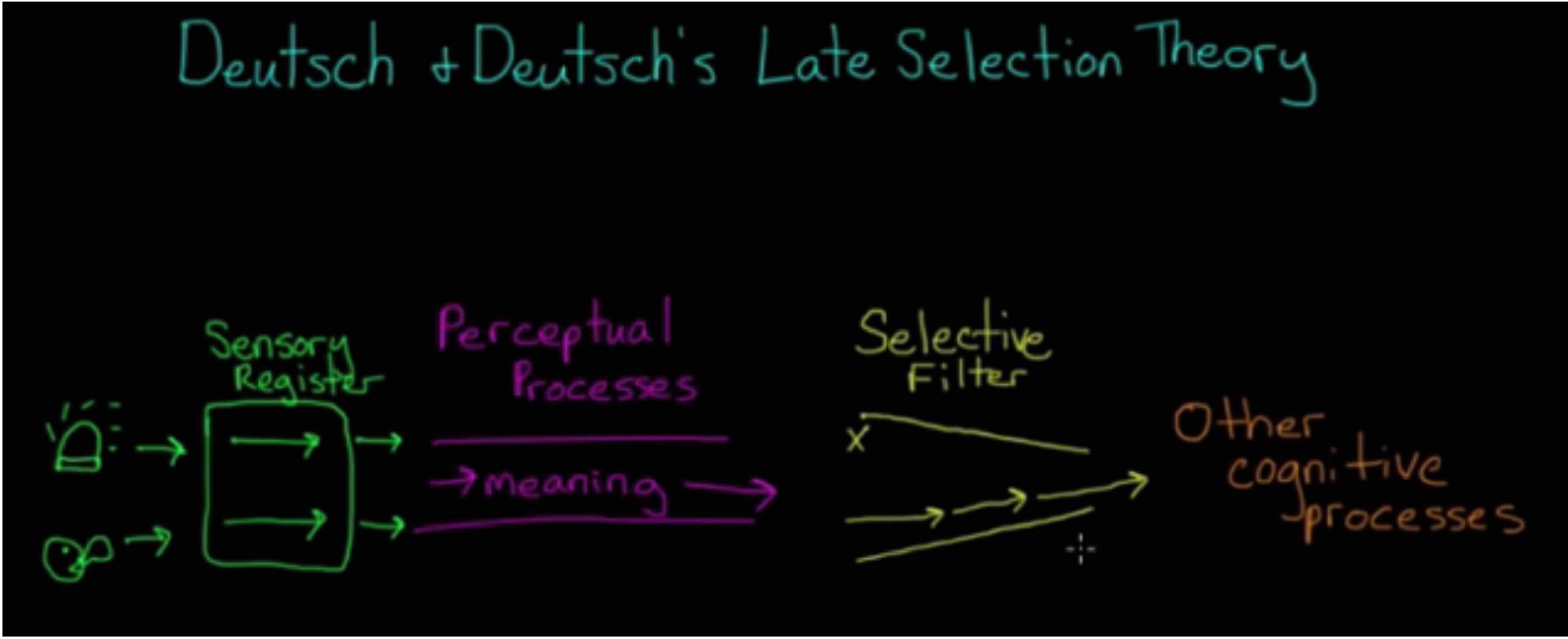
- *Shadowing Task*: ignore one input, repeat back the other
- Subjects are largely unaware of unshadowed message
- Certain words such as their name distract them
- Evidence exists for both early and late selection mechanisms
    - *early* reduction in “sensory gain” followed by *late* suppression of unselected information





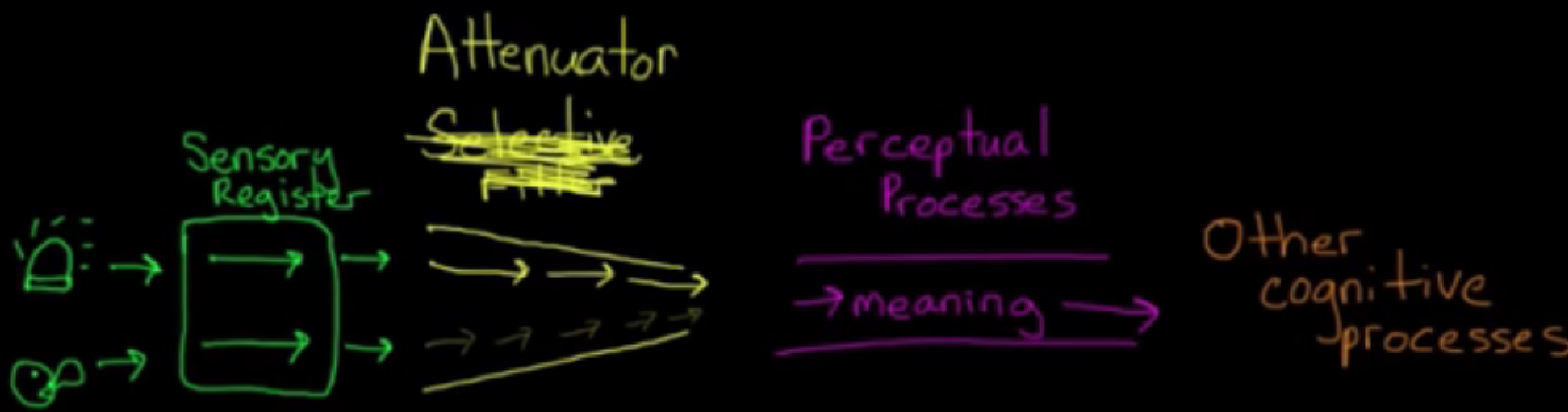


Theories of selective attention | Processing the Environment | MCAT | Khan Academy
https://www.youtube.com/watch?v=qpsaHE_uZic



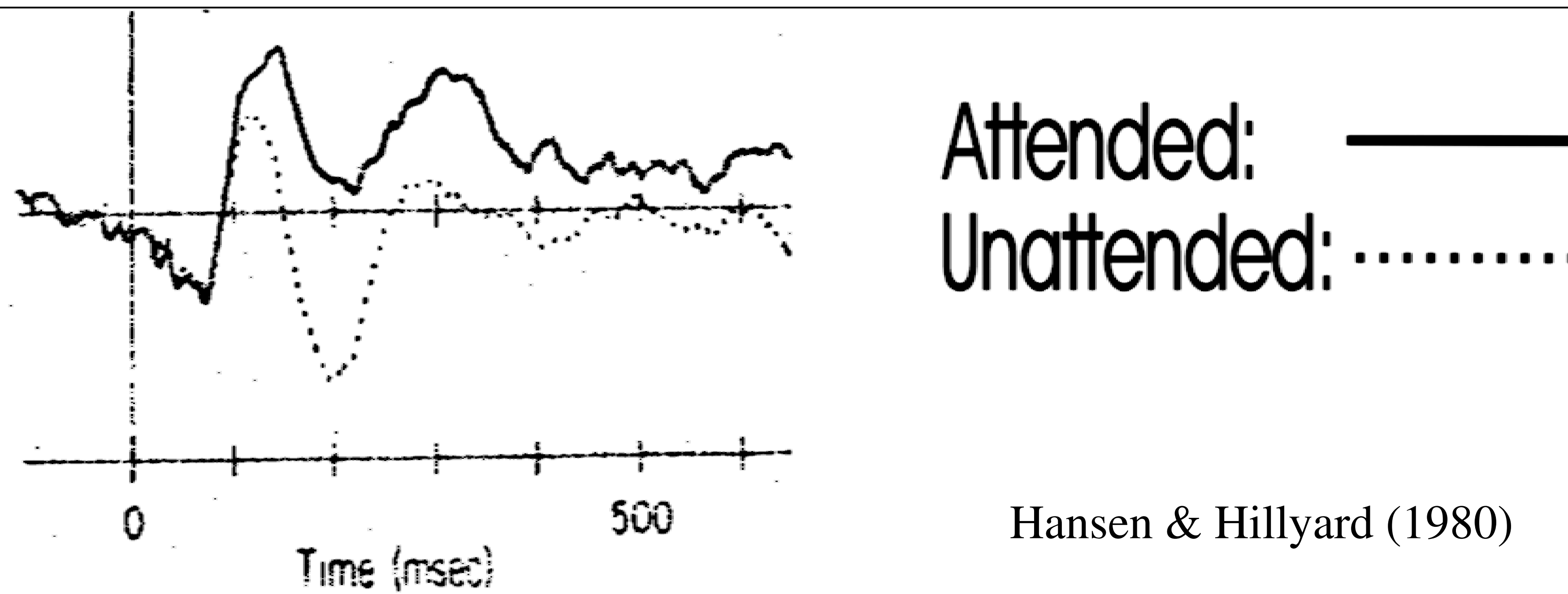
Limited resources and efficient mechanisms?

Treisman's Attenuation Theory

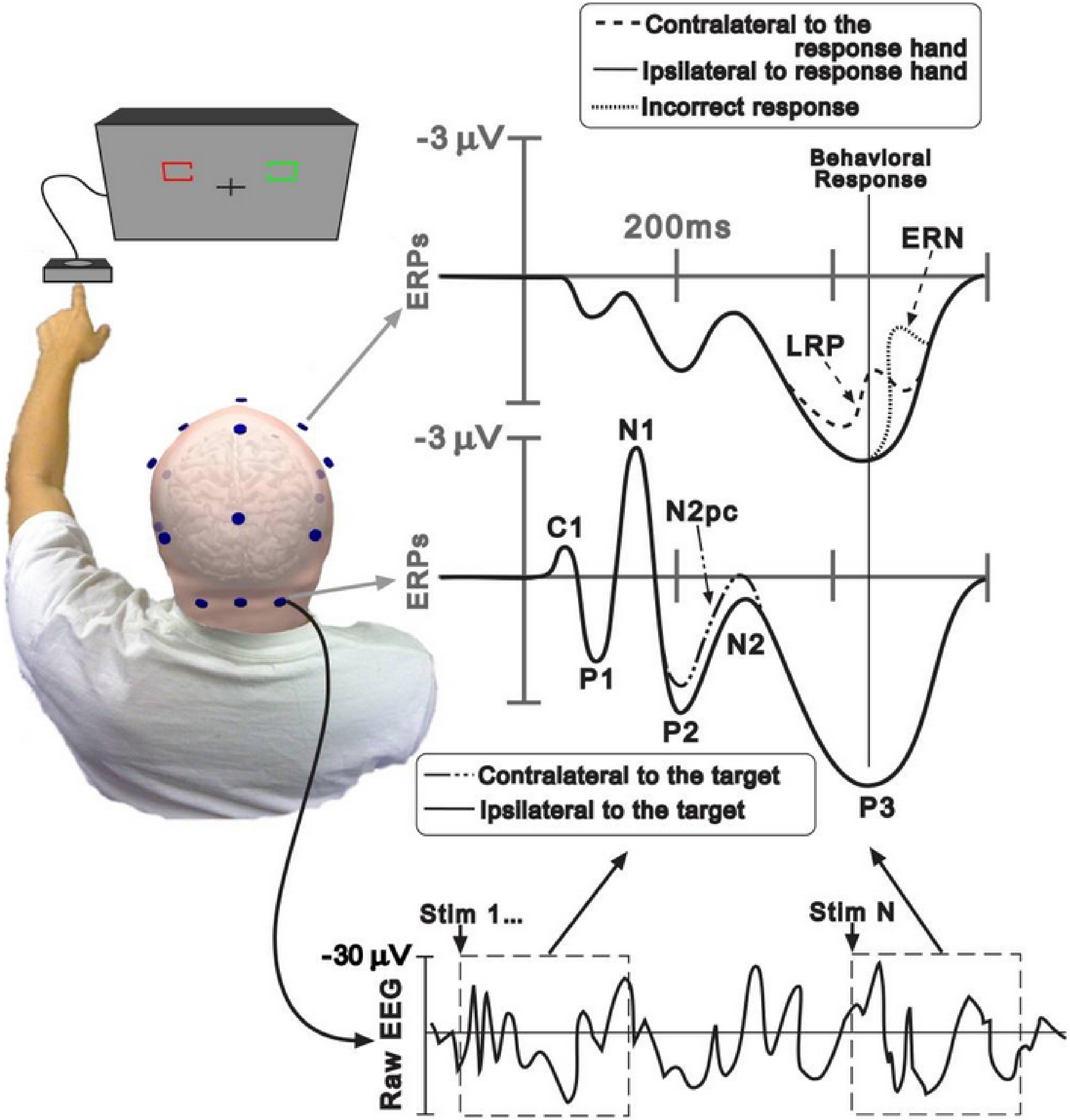


Testing Selection

- Electrical activity recorded at scalp (EEG) shows differences between attended and unattended stimuli in A1 within 90 ms



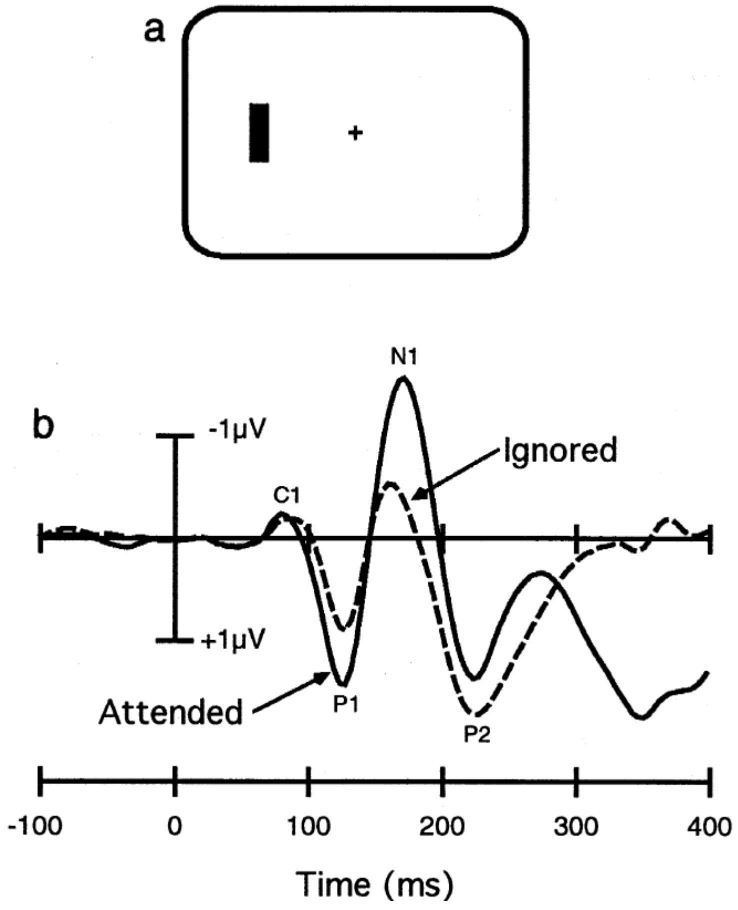
Hansen & Hillyard (1980)



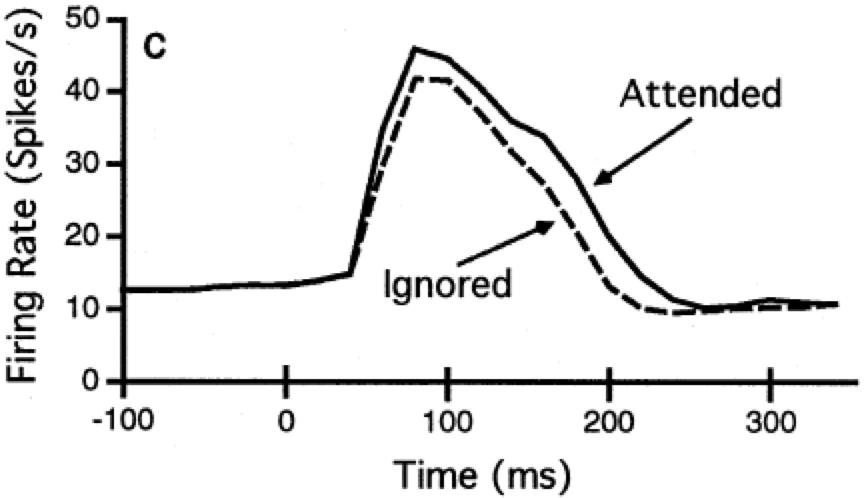
Contralateral to the response hand
Ipsilateral to response hand
Incorrect response
-3 μV
Behavioral Response
200ms
ERN
ERPs
LRP
-3 μV
N1
N2pc
C1
ERPs
N2
P1
P2
Contralateral to the target
Ipsilateral to the target
P3
Stim 1...
Stim N
-30 μV
Raw EEG

Event Related Potential (ERP)

- The C1 wave in area V1,
  - C1 wave is unaffected by spatial attention
- The C1 wave is followed by the P1 and N1 waves
  - P1 modulation reflects an effect of attention on sensory processing, supporting early selection models of visual–spatial attention. The N1 effect appears to reflect a modulation of visual discrimination processes.



Single-neuron recording



- Attention had no effect on responses in area V1, but in area V4 (an intermediate visual processing region) attended location stimuli evoked higher rates of neural firing than did ignored location stimuli
- Attention acts as a preset filter that controls the amplitude of the sensory response in intermediate-level areas of visual cortex.

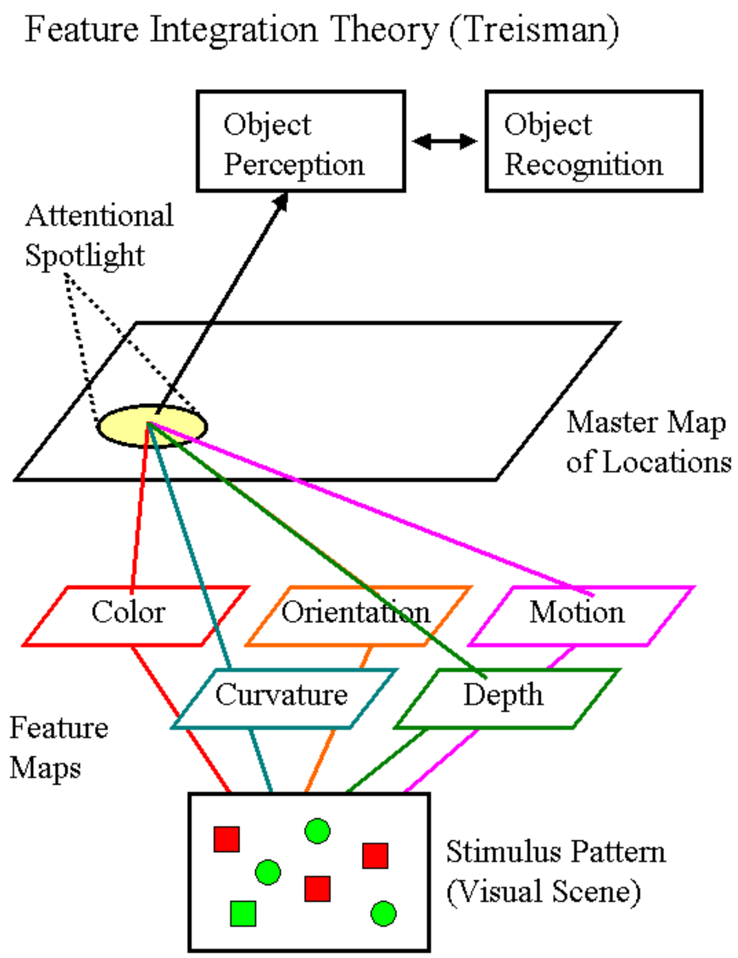
Saliency map

- The Saliency map integrates the normalized information from the individual feature maps into one global measure of conspicuity.
- Koch and Ullman, 1985 (p. 221), “Saliency at a given location is determined primarily by how different this location is from its surround in color, orientation, motion, depth etc.”
- Koch & Ullman hypothesized that a centralized retinotopic map can provide regulation of attention deployment (i.e. saliency map)
- Recent results show a number of retinotopically organized maps in the brain, particularly in midbrain, thalamus and occipital lobes, but also in temporal, parietal and even frontal cortices (Saygın & Sereno)

Attentional selection and inhibition of return

- Attentional spotlight attends to the most salient spot
- This spot is temporarily suppressed to allow the next salient spot to capture attention
  - Through “inhibition of return”
- Speeded local discriminations:
  - visual processing at recently attended locations might be slower
- Inhibition of return is shown to be object-bound

Feature Integration Theory



http://www.luc.edu/faculty/asutter/FIT.gif

Computational Models of Attention

- How to define SALIENCY?
  - Psychophysical models of human attention
    - Treisman, Sykes, Gelade, 1977
    - Duncan, Humphreys, 1989
    - Chun, Wolfe, 1996
  - Empirical approaches
    - Le Meur et al., 2006
  - Information theoretic approaches
    - Jaegersand, 1995
    - Gao, Vasconcelos, 2005

Empirical approach

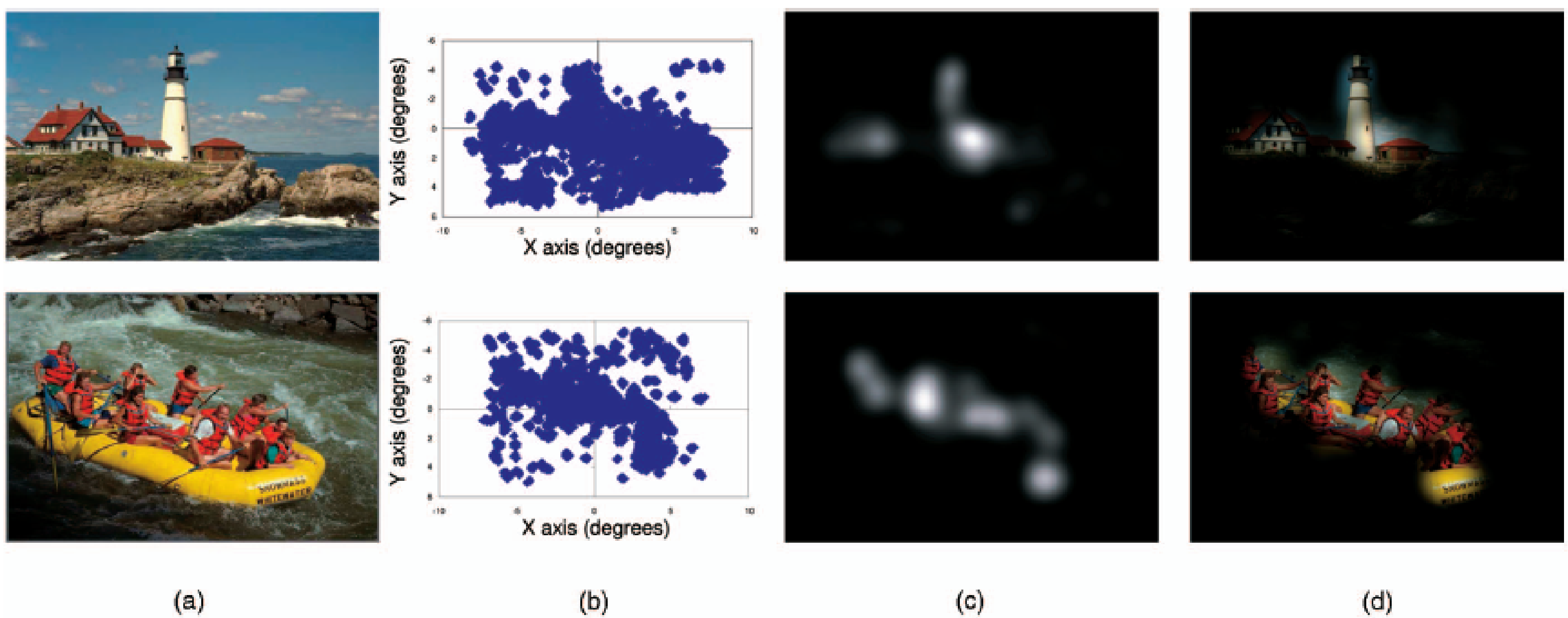
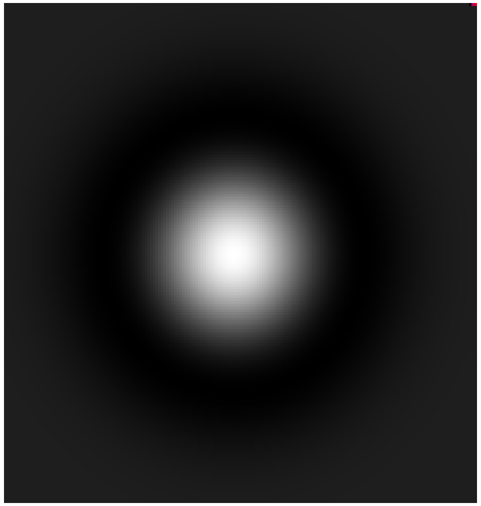


Fig. 2. (a) The original picture, (b) the spatial distribution of human fixations for 14s of viewing time, (c) fixation density map obtained by convolved the spatial distribution with a 2D Gaussian filter, and (d) highlighted human RoI (Regions of Interest) obtained by redrawing the original picture by leaving in the darkness in the nonfixated areas.

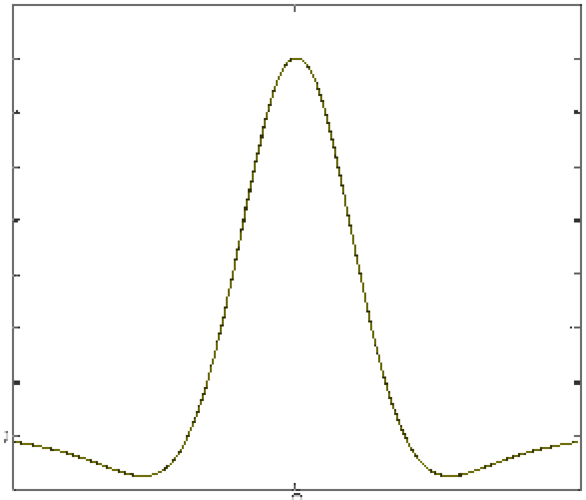
Le Meur et al., 2006

Centre-surround cell responses

$$\mathrm{DoG}_{\sigma,\gamma}(x,y) = A_c \frac{1}{\gamma^2} e^{-\frac{x^2+y^2}{2\gamma^2\sigma^2}} - A_s e^{-\frac{x^2+y^2}{2\sigma^2}}$$

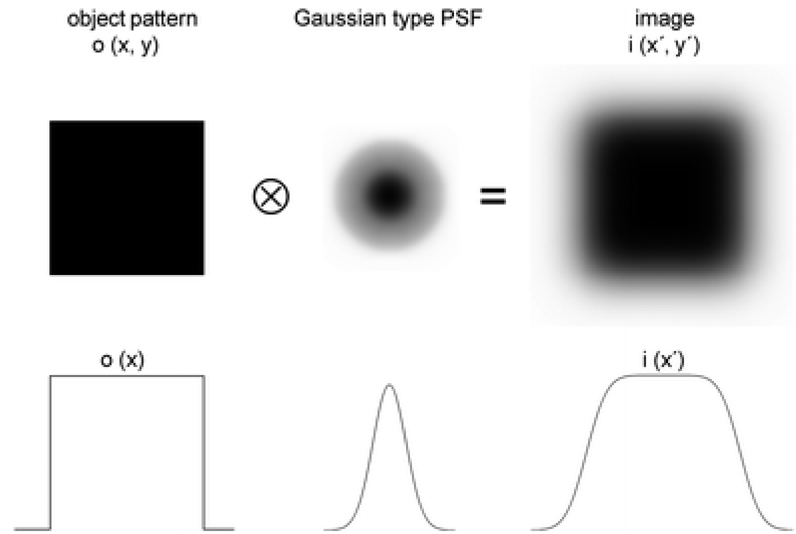


Intensity map of a
two-dimensional DoG function

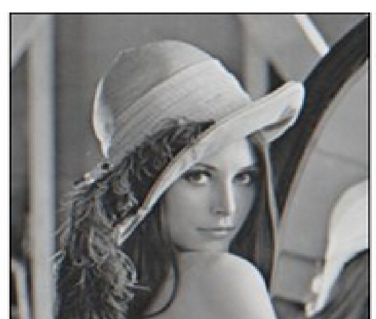


One-dimensional DoG profile

CONVOLUTION



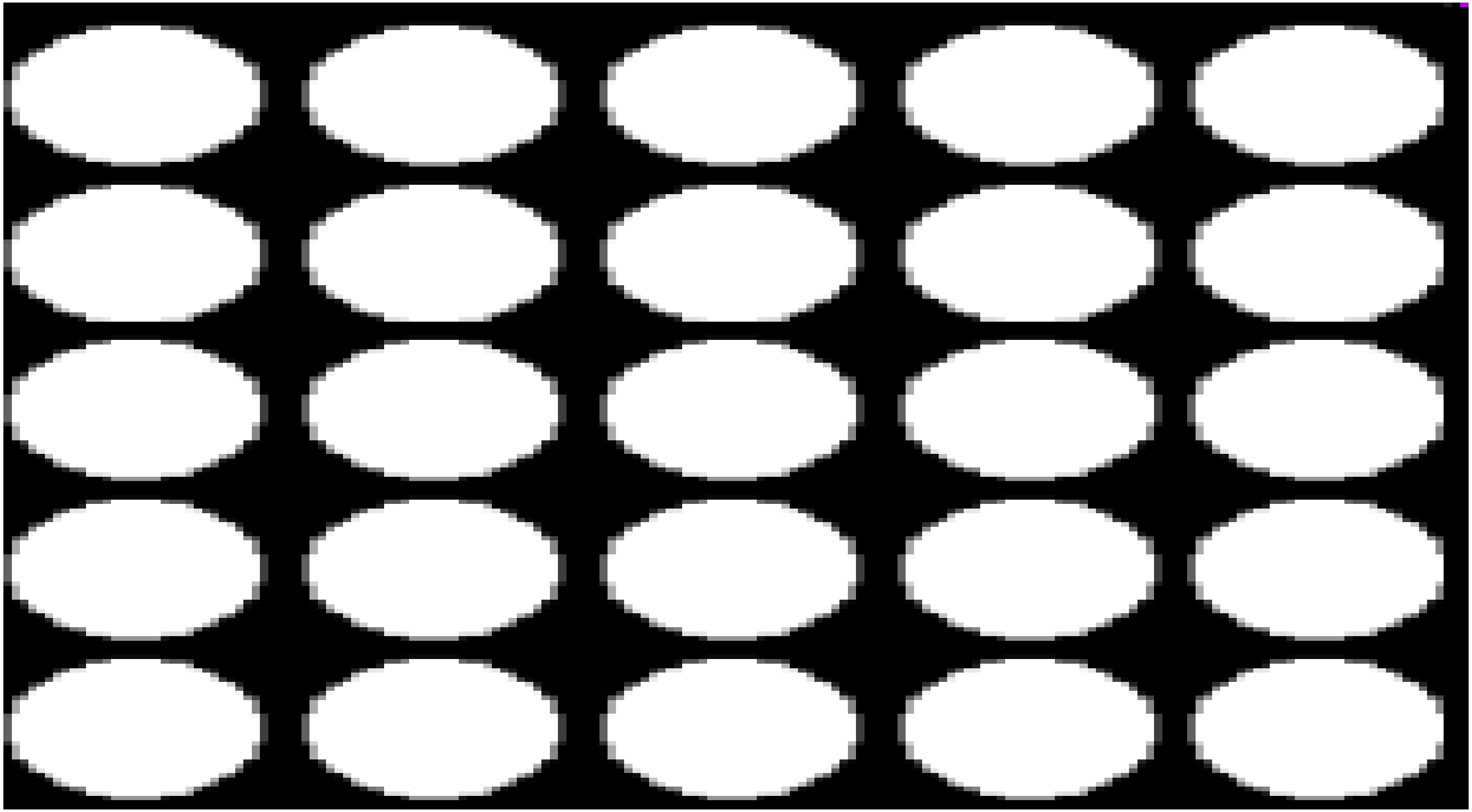
$\sigma = 1$



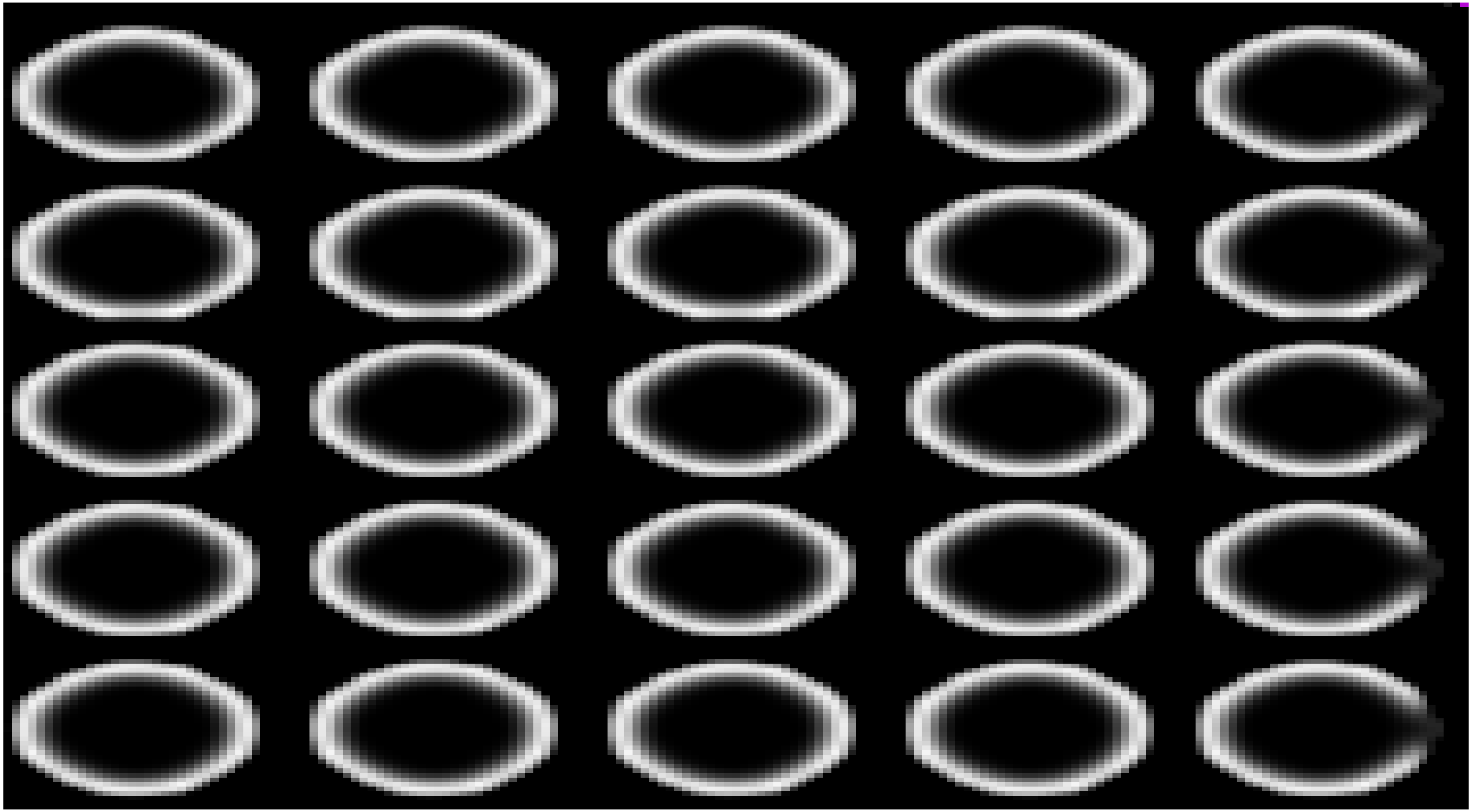
$\sigma = 2$

difference

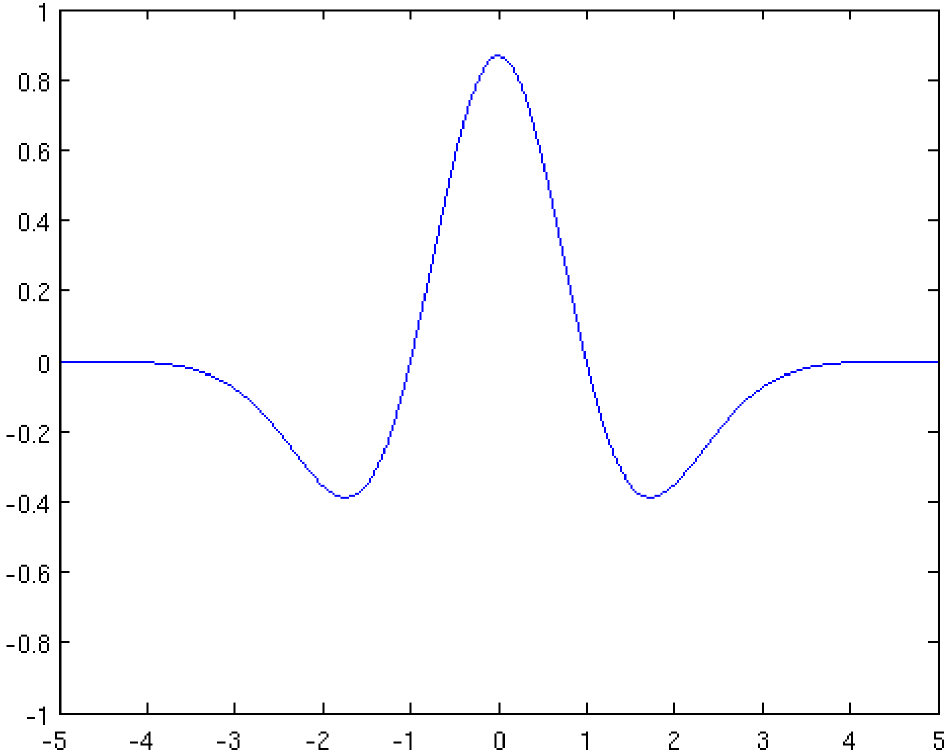
The effect of centre-surround



This will be the result

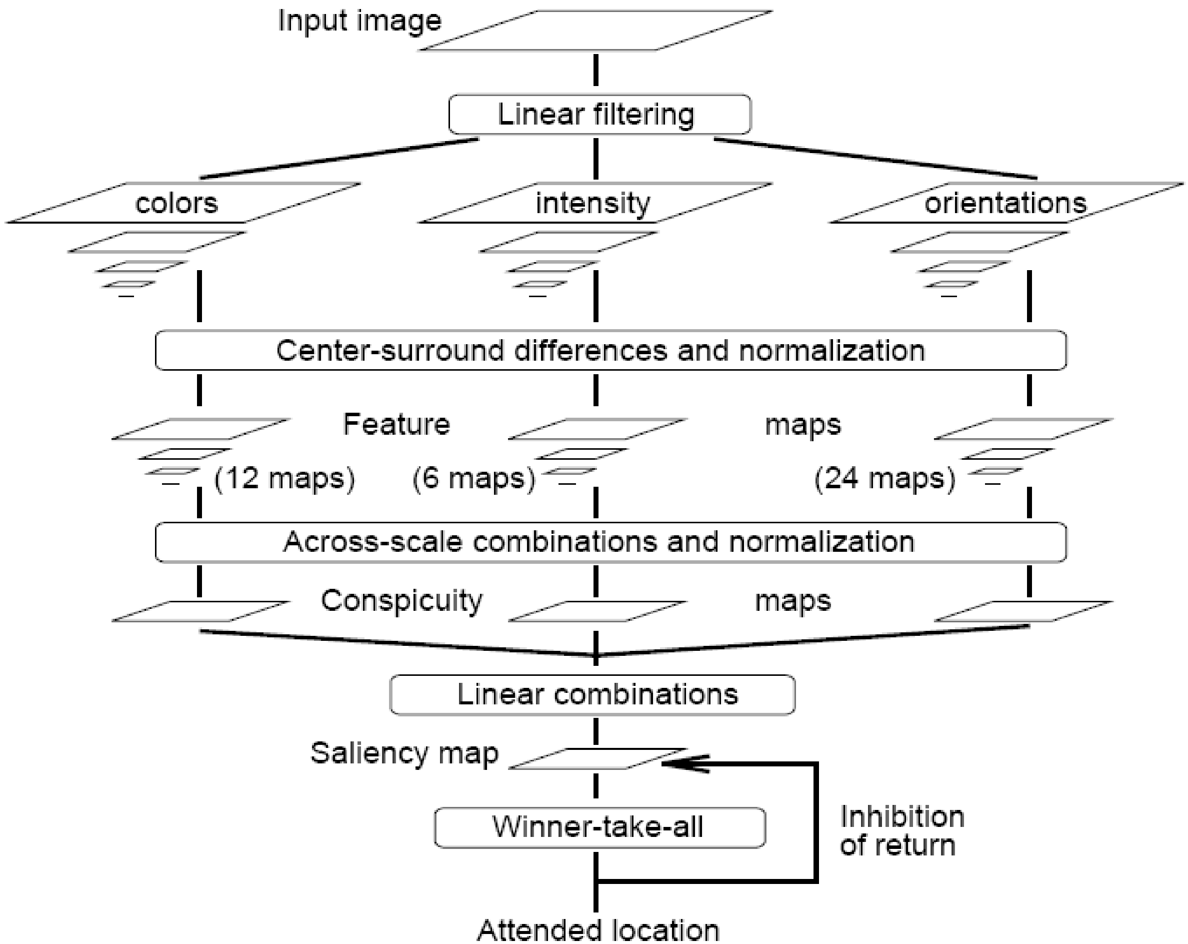


Mexican hat

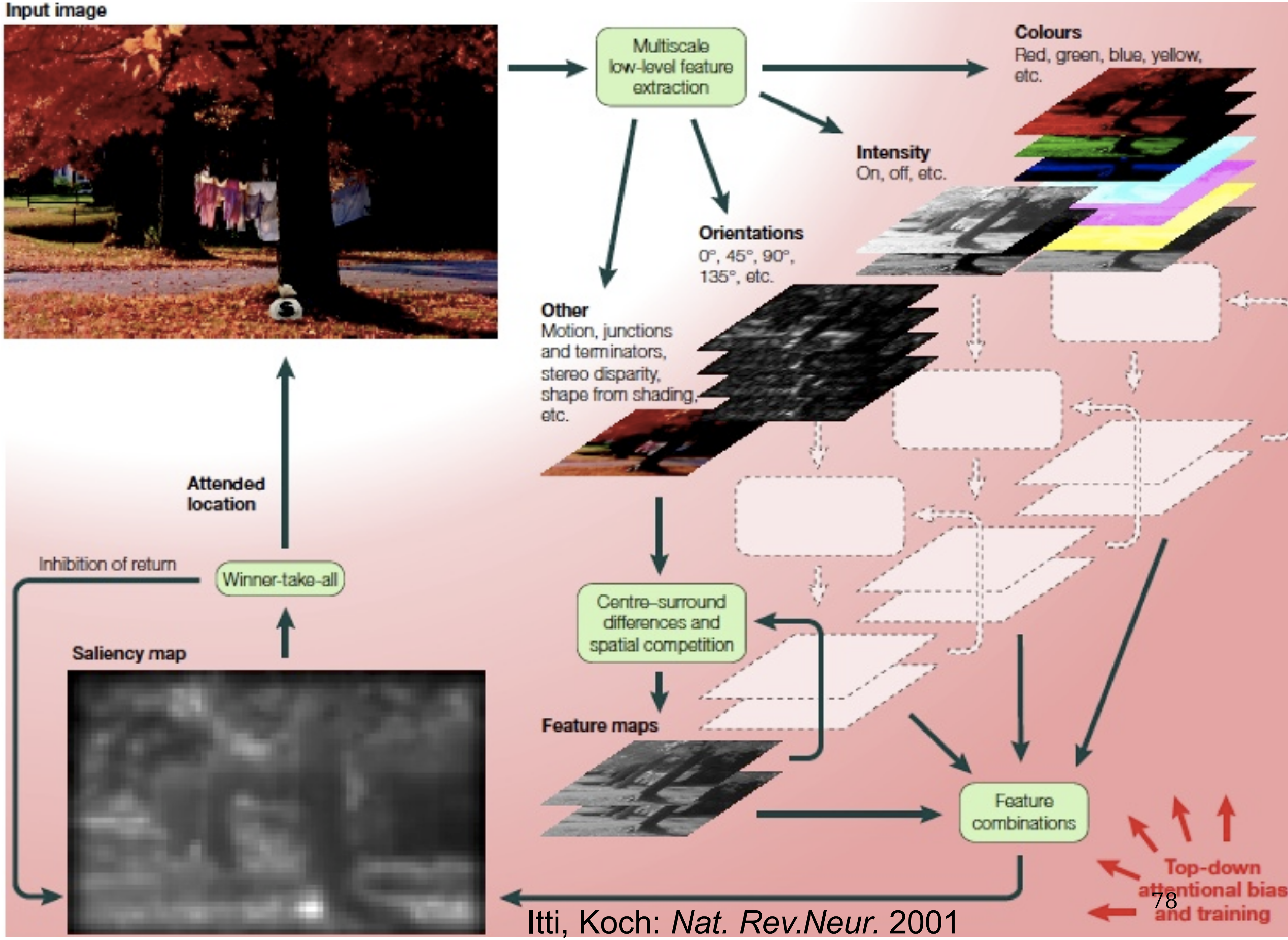


$$\psi(t) = \frac{2}{\sqrt{3\sigma}\pi^{\frac{1}{4}}}\left(1 - \frac{t^2}{\sigma^2}\right) e^{\frac{-t^2}{2\sigma^2}}$$

Computational Model for FIT



Itti, Koch, Niebur, IEEE Trans. PAMI, 1998.



Itti, Koch: *Nat. Rev.Neur.* 2001

References

- Josh Mc Dermott – Attention lecture notes
- Matthew Tata – lecture notes
- Itti, Koch, 2001
- Knudsen, 2007

The source code to compute saliency maps is freely available at http://ilab.usc.edu/toolkit/downloads.shtml.