Visual Attention and Ecological Theory of Perception

DT2350 Human Perception for Information Technology

Giampiero Salvi

KTH/CSC/TMH giampi@kth.se

HT 2015

Literature

Goldstein, E. (2009/2014) **Sensation and Perception**

- Chapter 6: Visual Attention
- Chapter 7: Taking Action

Weinschenk, S.M. (2011) **100 Things Every Designer Needs to Know About People**

- Chapter 45: People pay attention only to salient cues
- Chapter 8: People can miss changes in their visual field
- Chapter 7: People see cues that tell them what to do with an object (already in lecture 2)

most references in the book

Preamble: Studying Humans

Most of the studies cited in the book involve humans facing stimuli

- problem 1: partial view on perception
- problem 2: often artificial conditions
- the models we obtain can predict those specific situations

it is also desirable to test the global system

Opportunity: Simulate Humans

- make the complexity of real problems emerge
- test perception/cognition theories
- extra effect: create artifacts that interact with us in a more natural way
- limitation: we have to simplify to some extent

NOTE: perceptual studies will always be important

Introducing the iCub

- platform for cognitive studies
- large European project (RobotCub) involving researchers from engineering to neuroscience to psychology
- eight exemplars in labs in Europe, USA and Japan
- sensors: vision, hearing, tactile, proprioception, balance
- cognitive models implemented as control software

Outline

Outline

Why Do We Need Visual Attention?

The senses send to the brain huge amounts of information

- avoid overloading the brain
- dimensionality reduction
- disregard irrelevant information

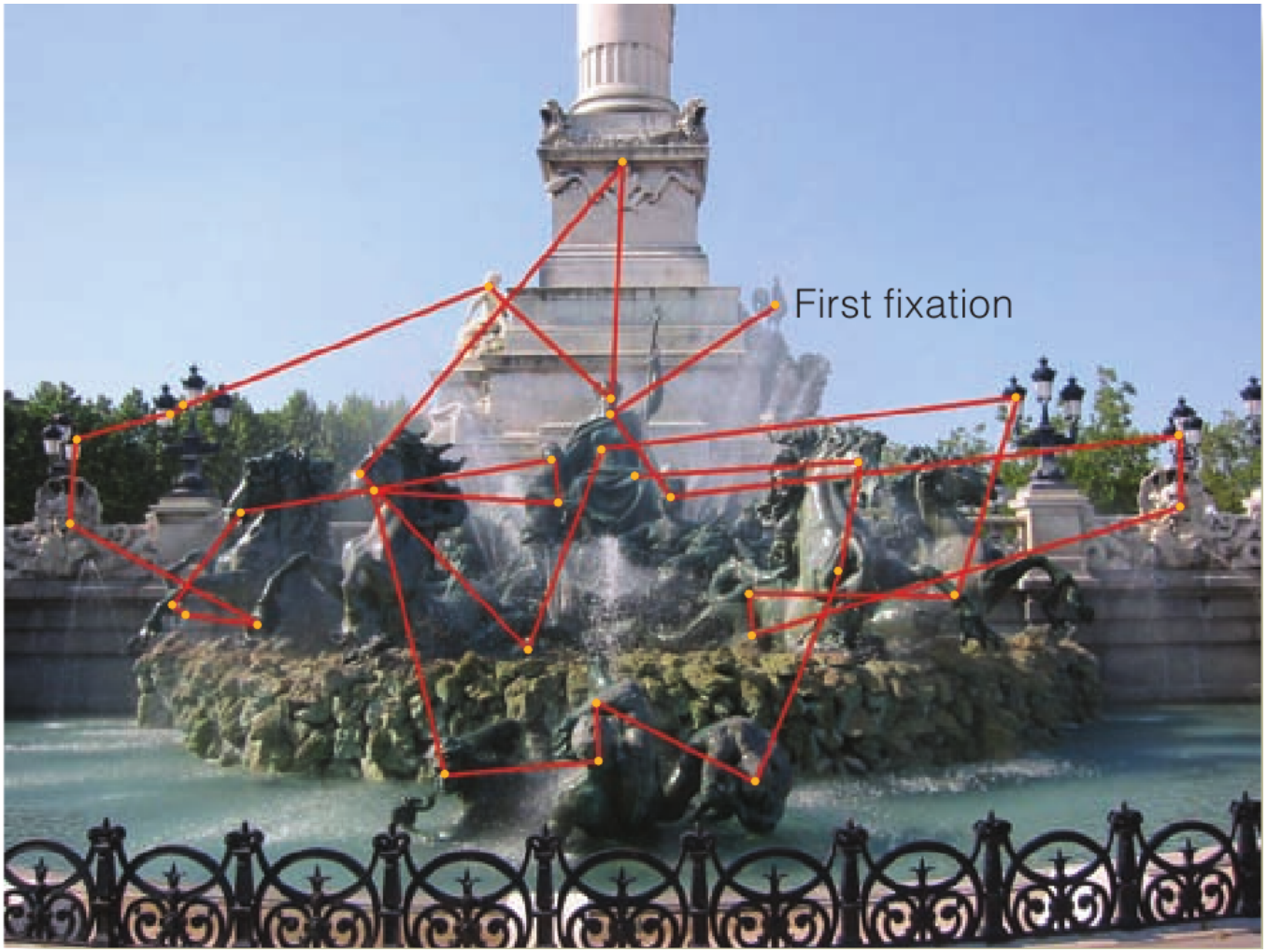
William James (1890) Principles of Psychology

Camera vs Retina

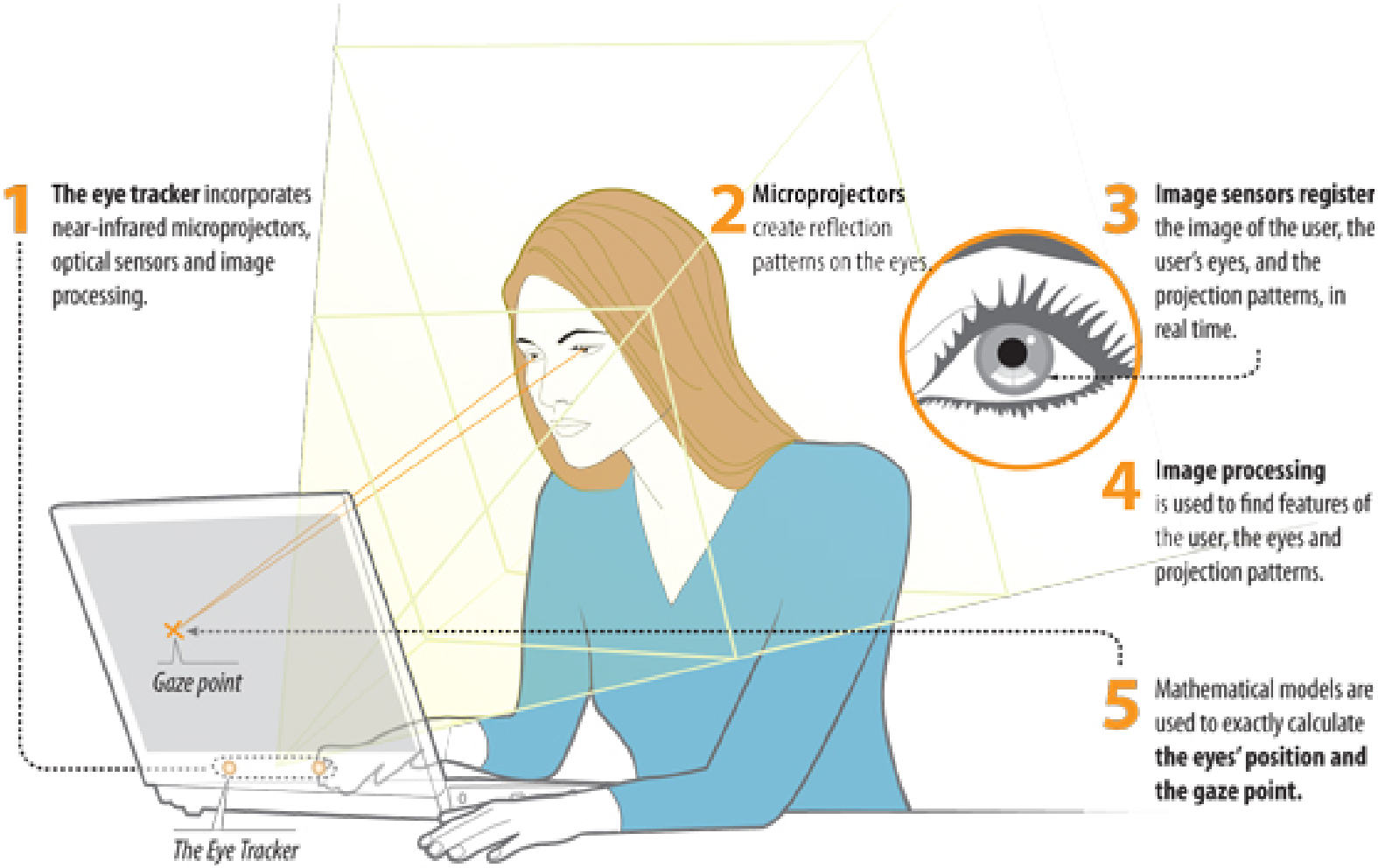
Embodied dimensionality reduction

peripheral vision is blurred

Need to scan the scene: gaze!!



Eye Trackers



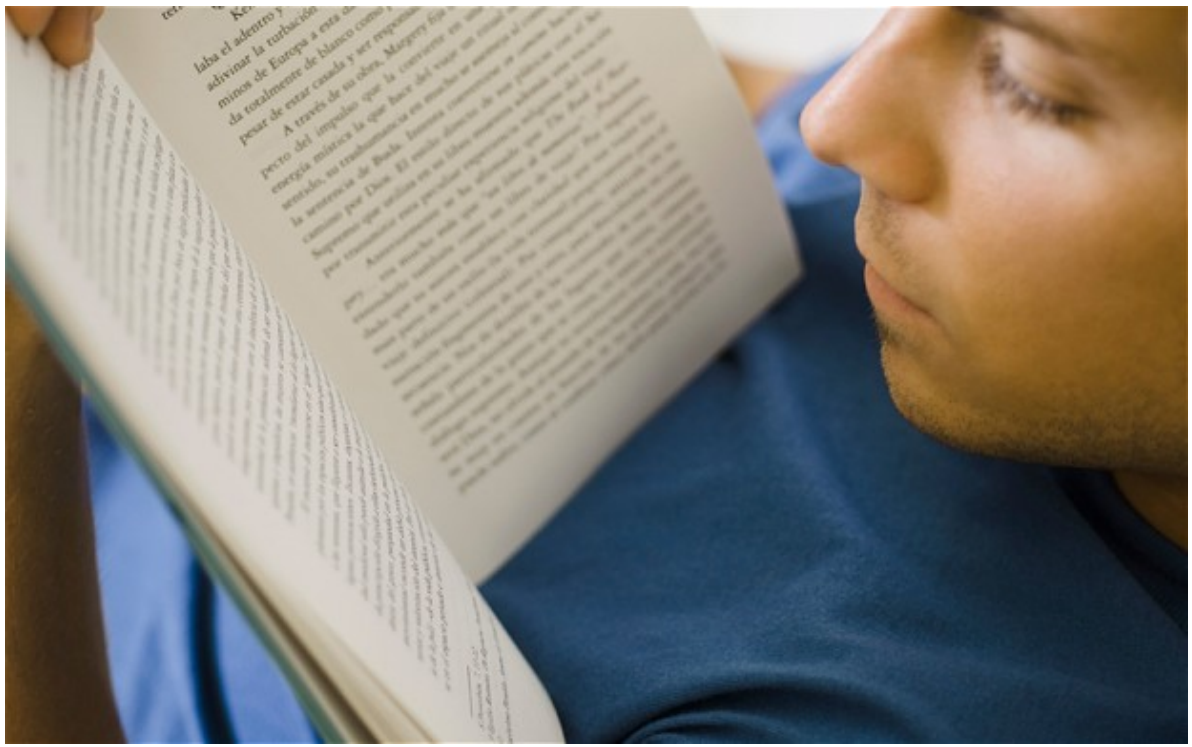
Eye Trackers

- very useful for visual attention research
- essential for research with small babies

But Attention is a Mental Process

Not only Gaze!! Examples:

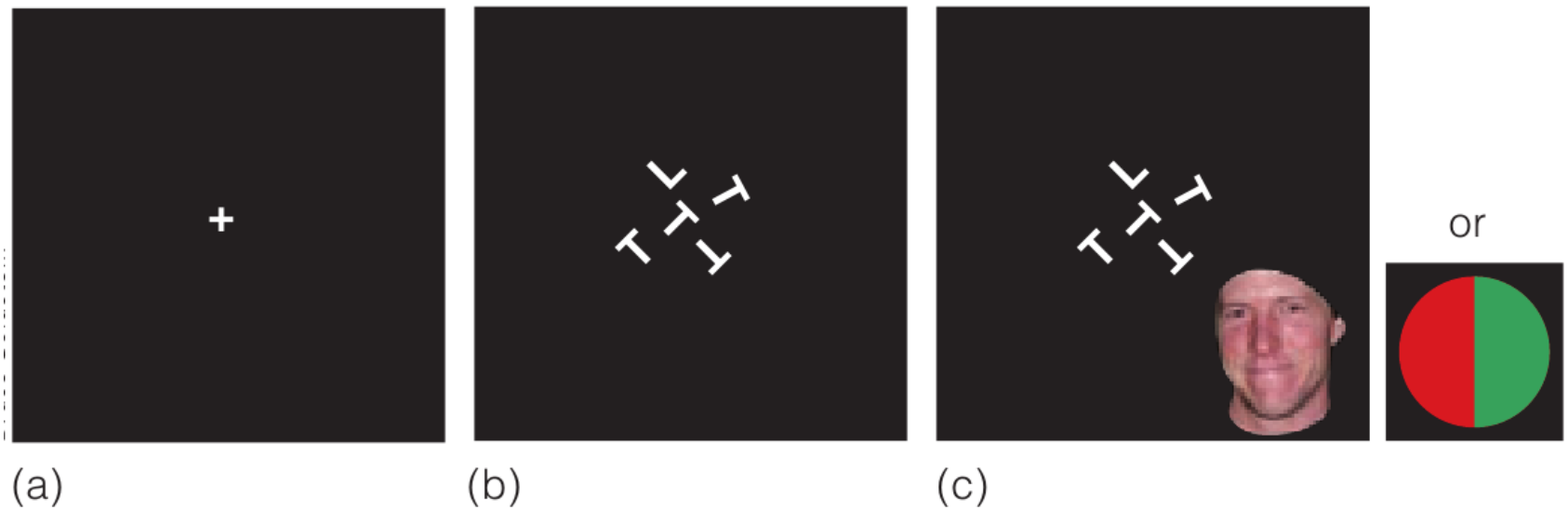
- reading without paying attention
- we can pay attention to peripheral vision



Overt vs Covert Attention

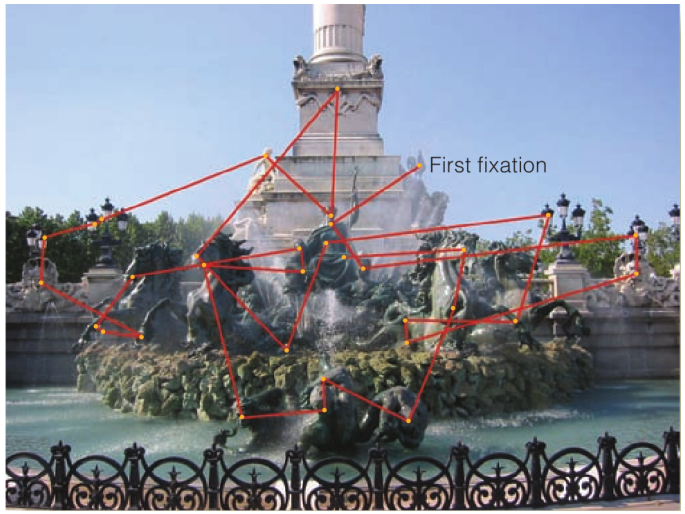
Overt looking directly at the attended object

Covert attention without looking



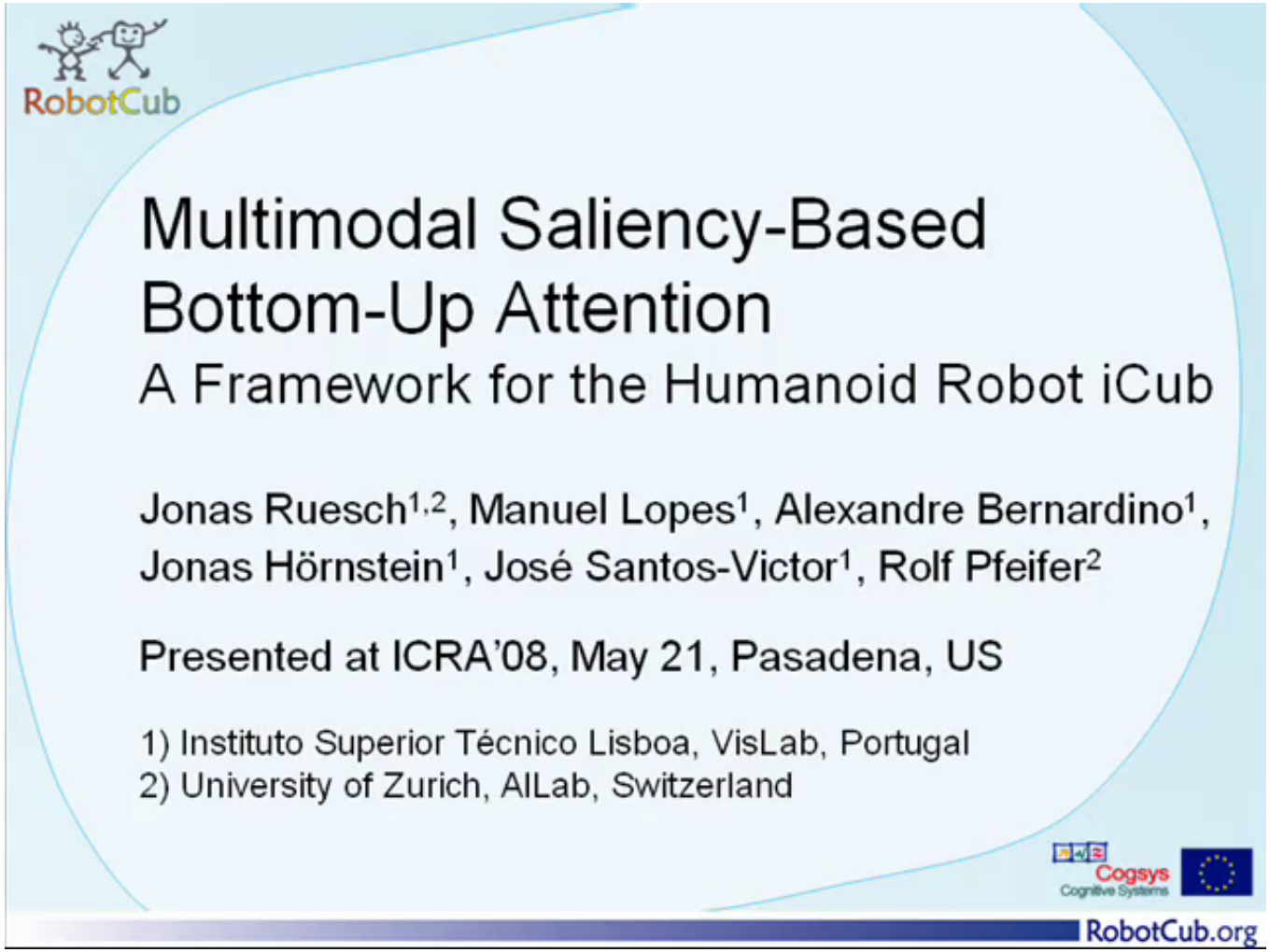
Li et al (2002)

What Determines How We Scan A Scene?



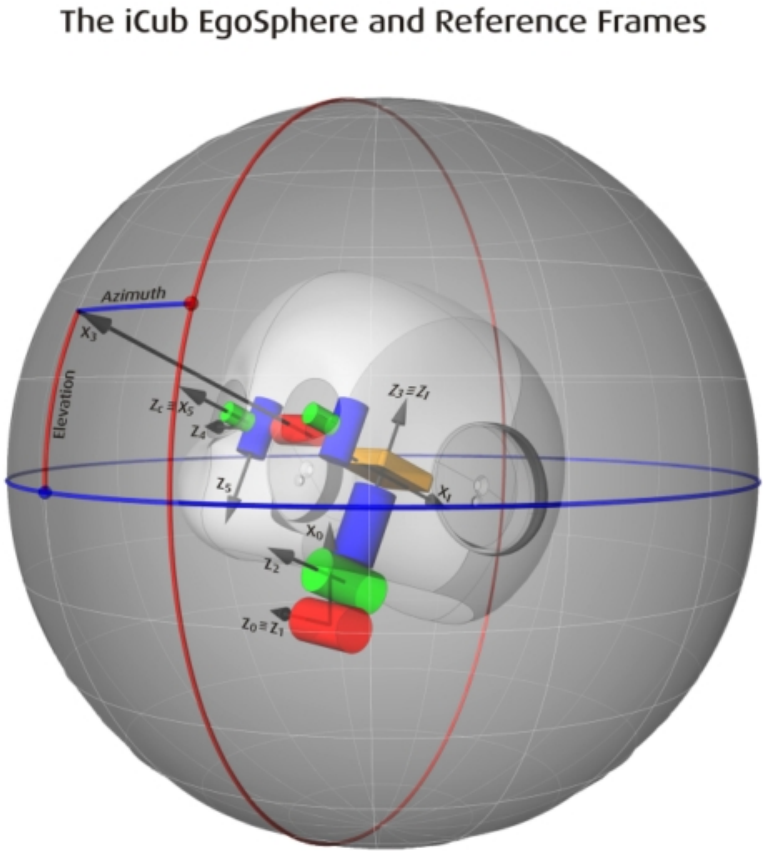
- Stimulus salience (bottom-up)
- Knowledge about the scene (top-down)
- Nature of the observer's task
- Learning from past experience

Stimulus Salience (Bottom-Up)



http://youtu.be/Z7y-7VX6-Qw

Egosphere [1]

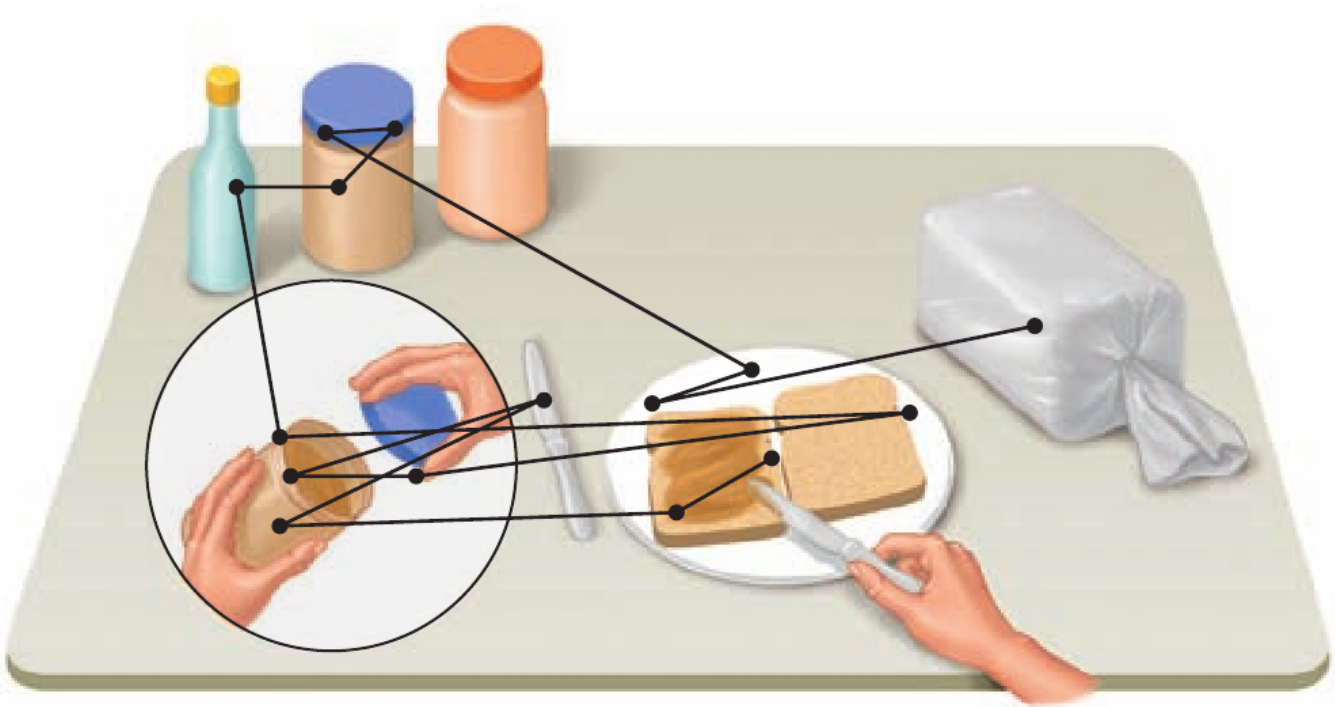


- based on: colour, movement, sound, faces

[1] J. Ruesch, M. Lopez, A. Bernardino, and J. Hornstein. "Multimodal saliency-based bottom-up attention a framework for the humanoid robot iCub". In: *IEEE ICRA*. Pasadena, CA, 2008, pp. 962–967

Knowledge-Based Attention

Task-Oriented Attention

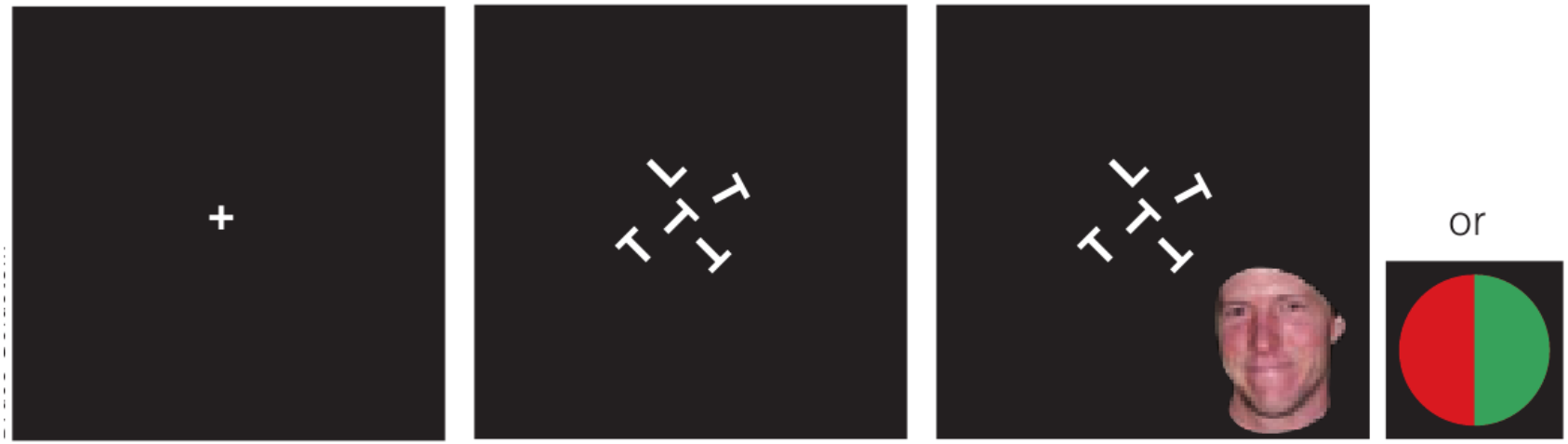


Attention vs Perception

Perception can occur without attention, but:

1. lack of attention can impair perception
   - inattention blindness
   - change blindness
2. focused attention can enhance perception
   - more vivid perception
   - binding features into coherent perception

Perception Without Focused Attention

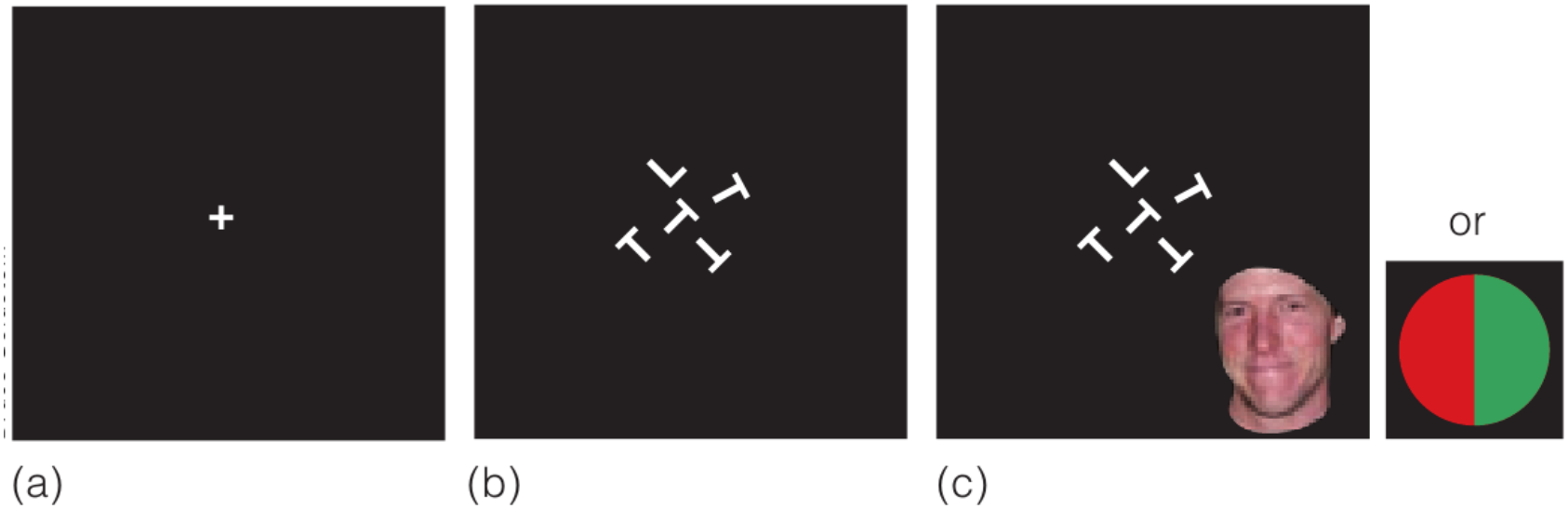


Always look at the centre of the screen

1. are all the letters equal?
2. is the face male or female? Is the disk green-red or red-green?
3. both conditions simultaneously

Li et al. (2007)

Perception Without Focused Attention

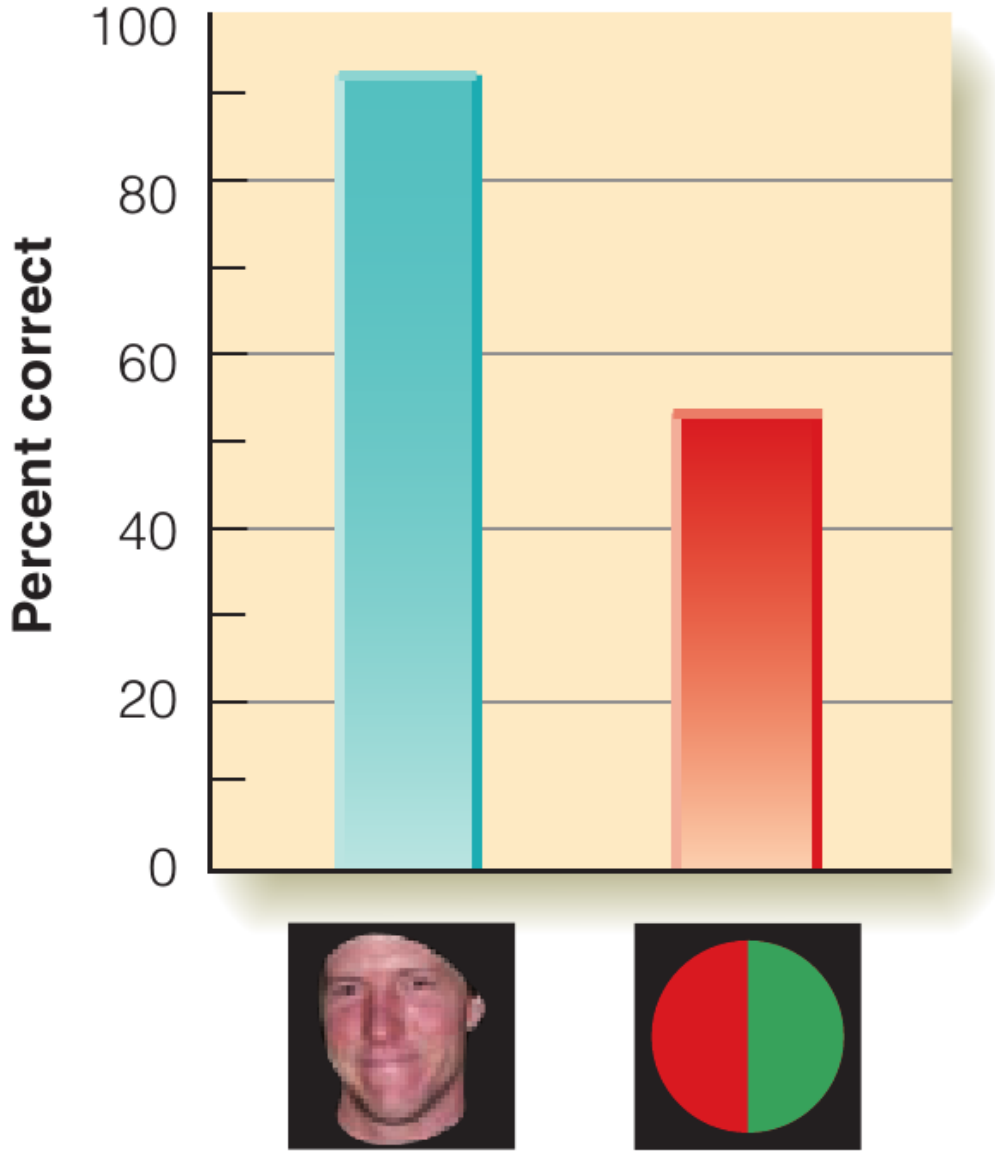


Always look at the centre of the screen

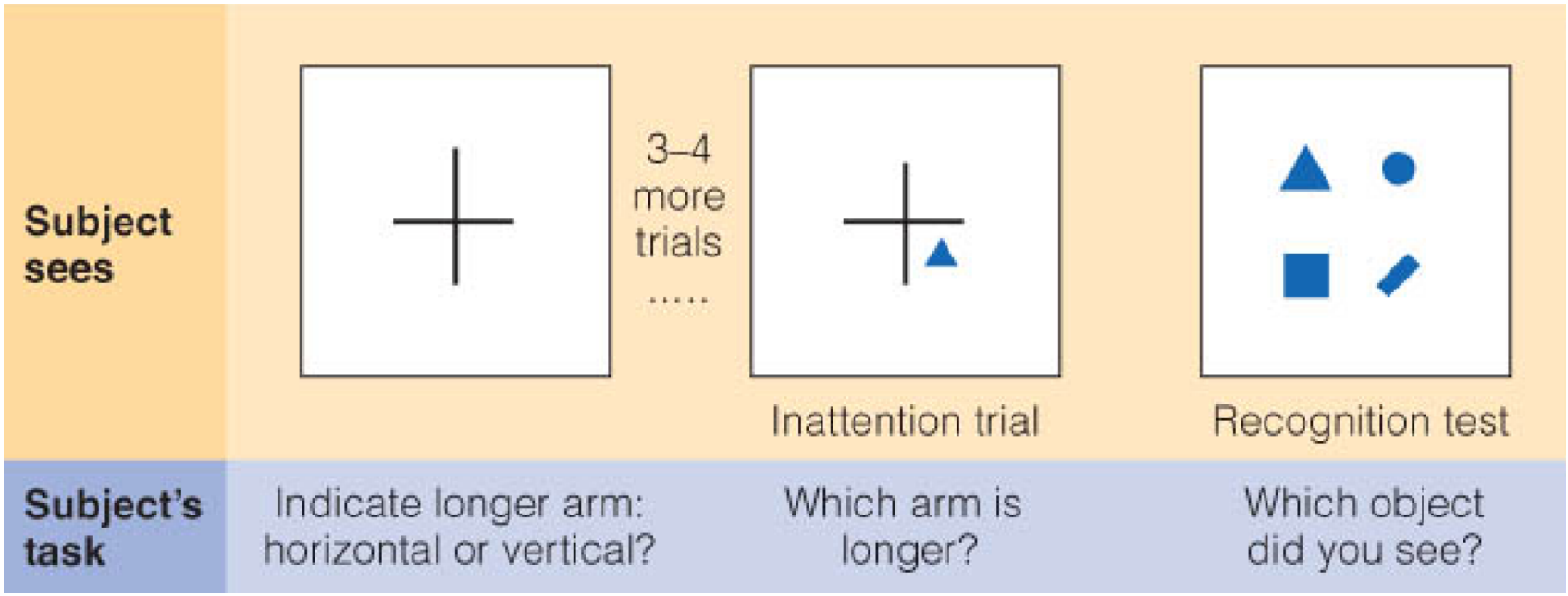
1. are all the letters equal? 80-90% correct
2. is the face male or female? Is the disk green-red or red-green? 80-90% correct
3. both conditions simultaneously see next slide

Li et al. (2007)

Perception Without Focused Attention

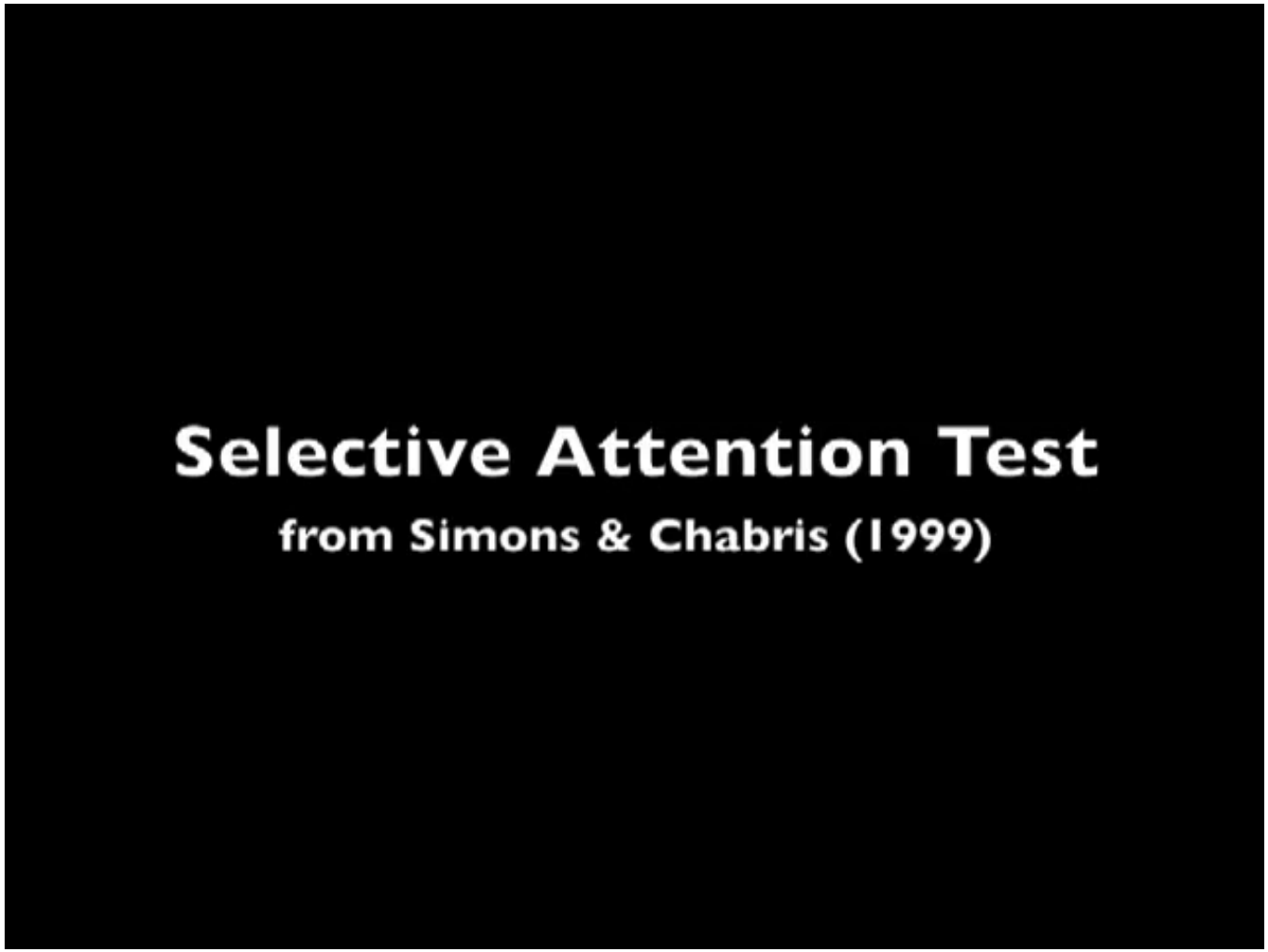


Inattention Blindness



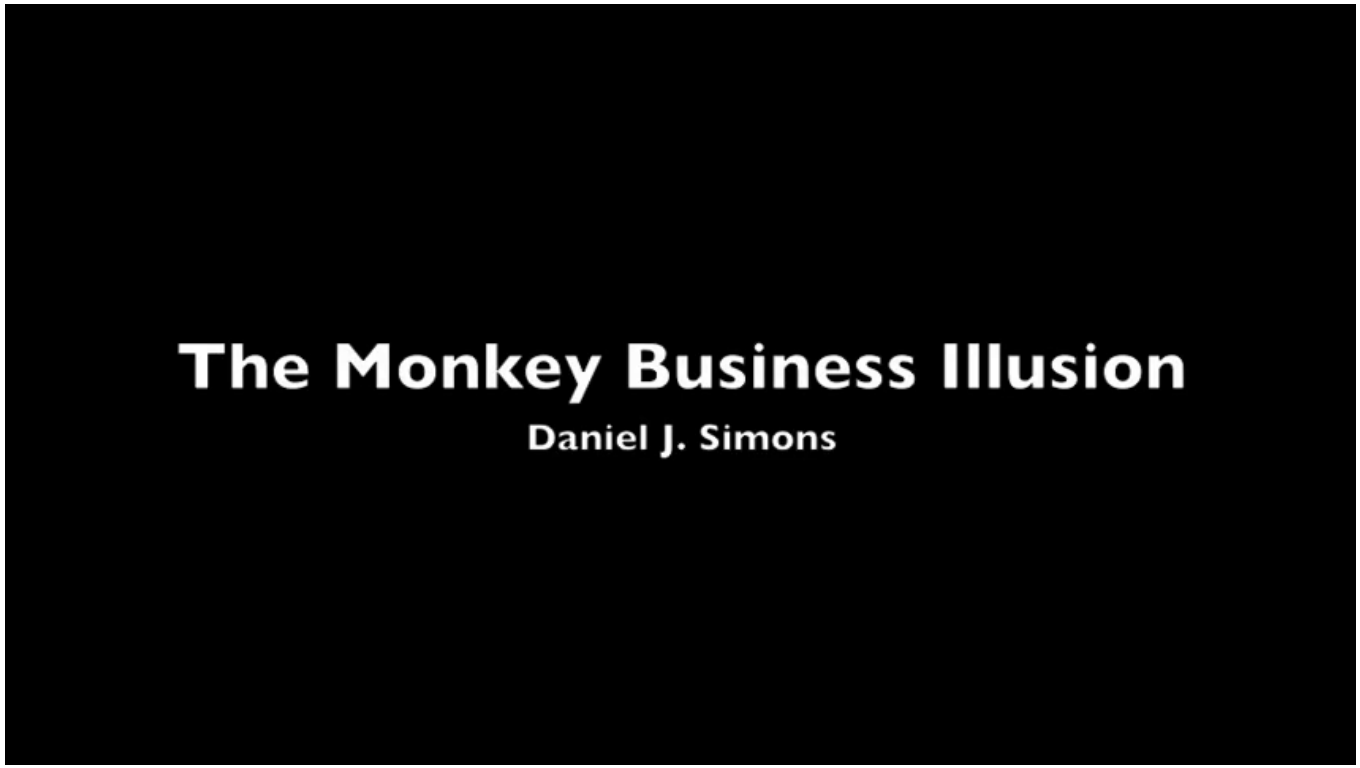
Mack and Rock (1998)

Inattention Blindness: Demonstration



http://youtu.be/vJG698U2Mvo

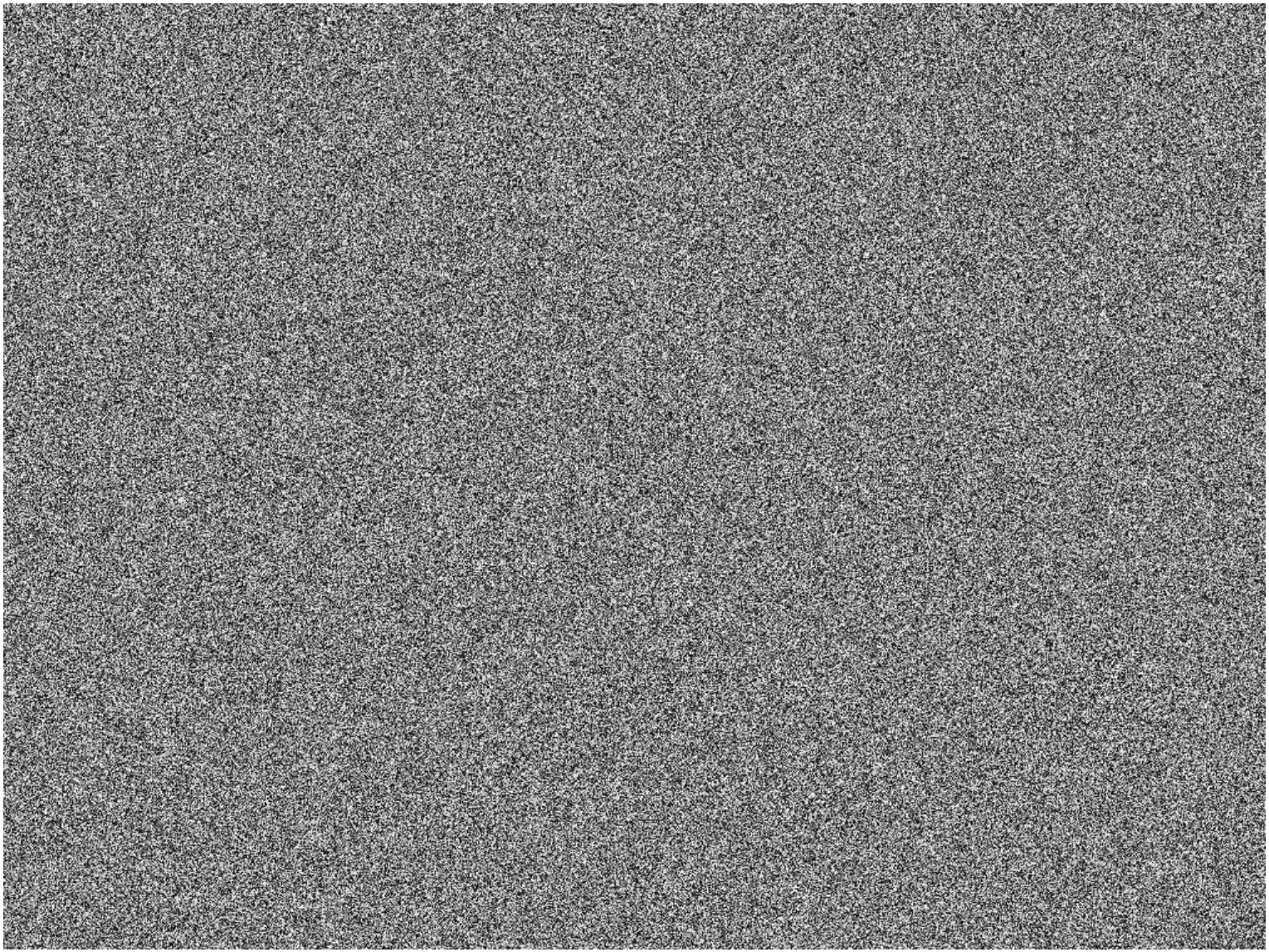
Inattention Blindness: Demonstration 2



http://youtu.be/IGQmdoK_ZfY

Change Blindness: Demo 1

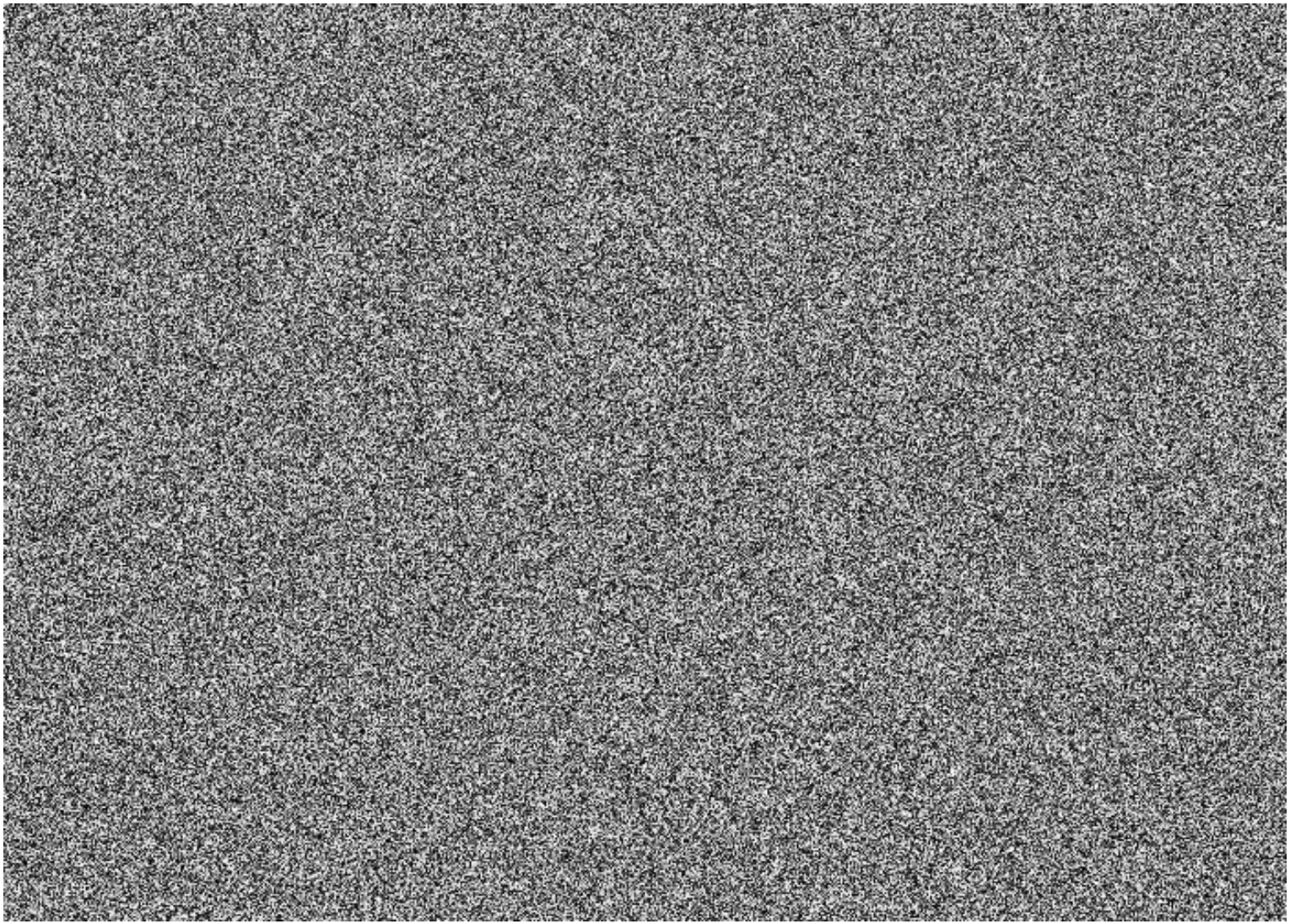
Change Blindness: Demo 1



Change Blindness: Demo 1

Change Blindness: Demo 2

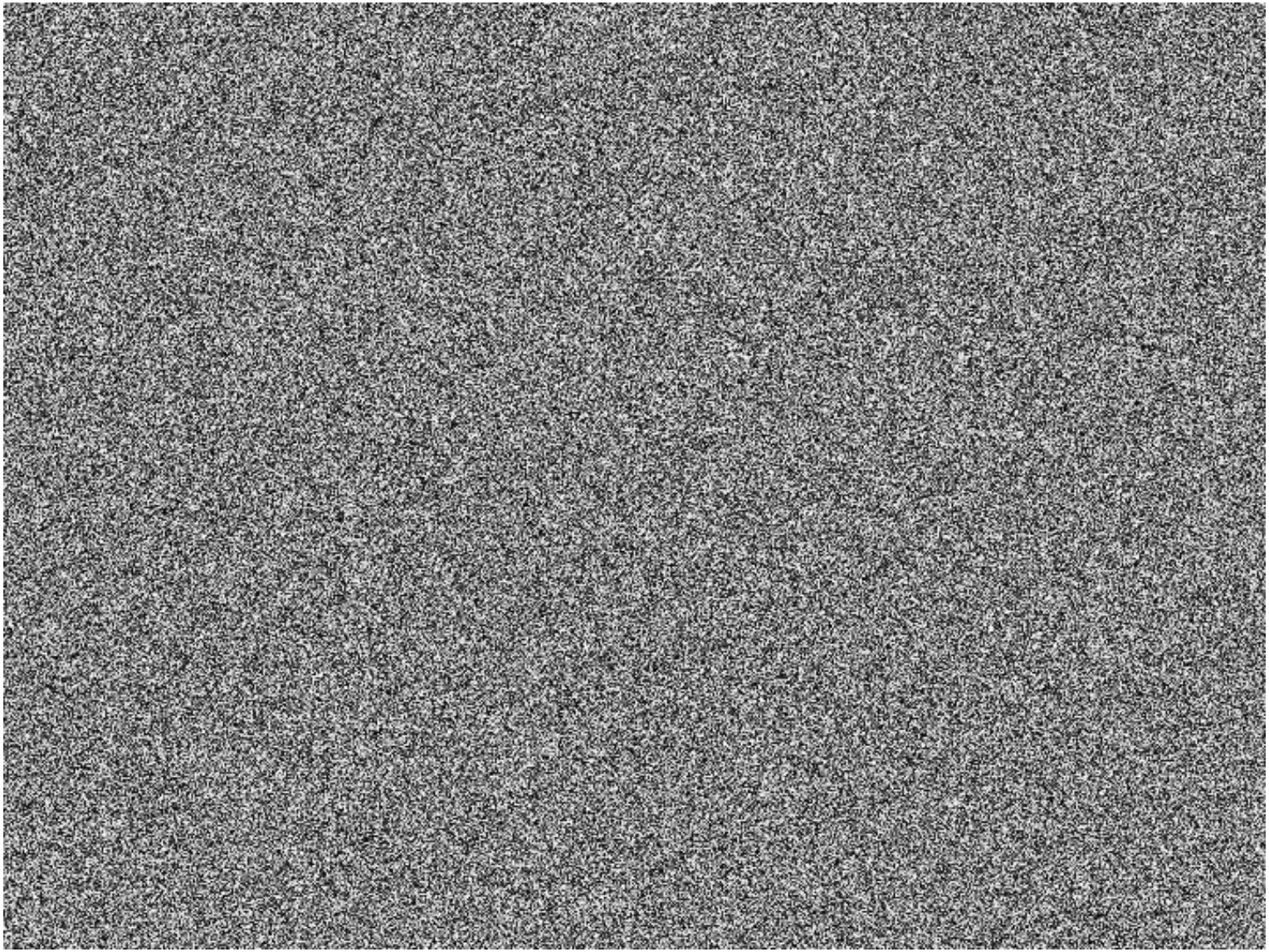
Change Blindness: Demo 2



Change Blindness: Demo 2

Change Blindness: Demo 3

Change Blindness: Demo 3



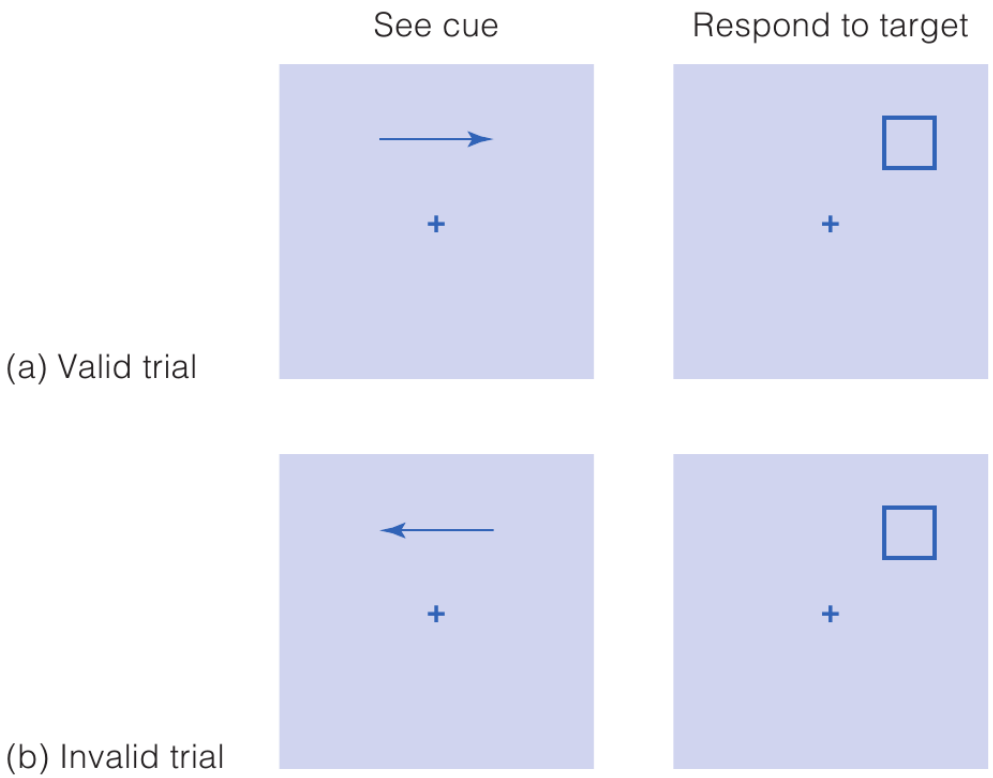
Change Blindness: Demo 3

Continuity Errors

Levin and Simons (1997)

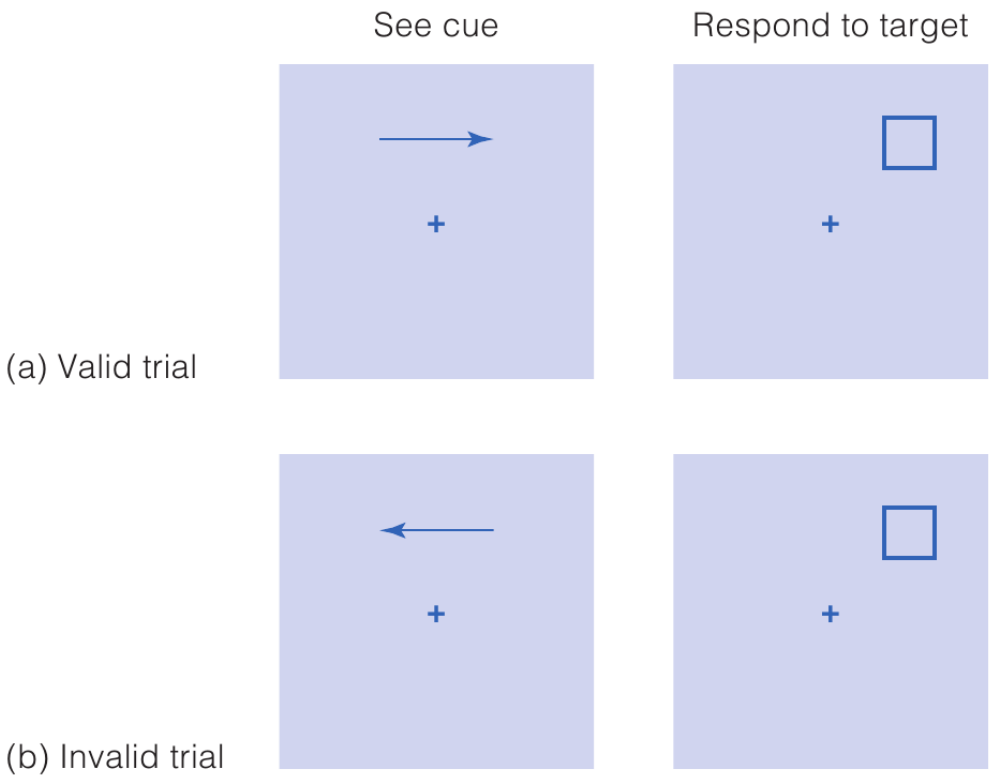
Does Attention Enhance Perception?

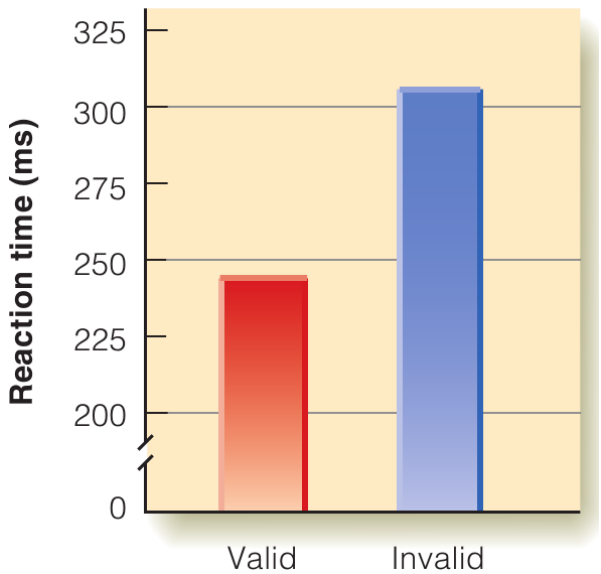
Measuring reaction times:



Does Attention Enhance Perception?

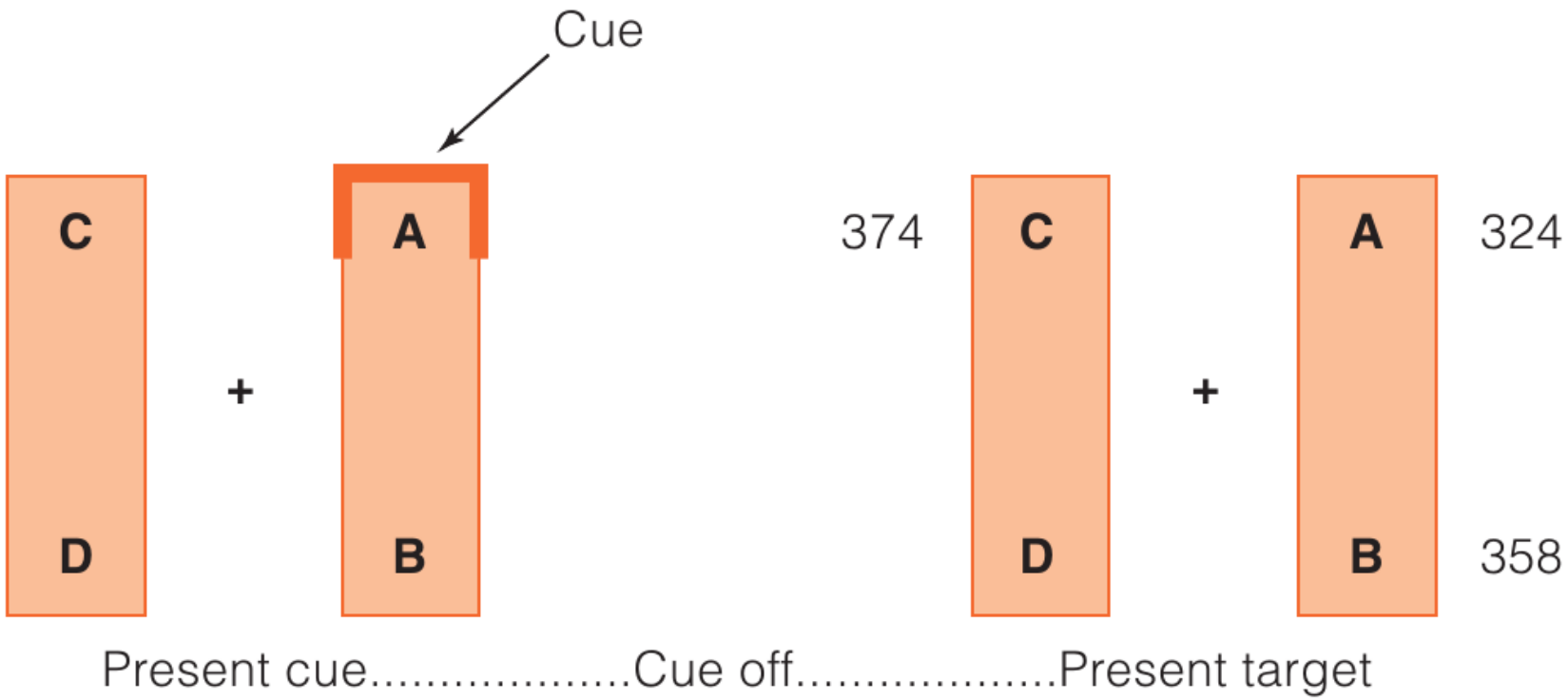
Measuring reaction times:





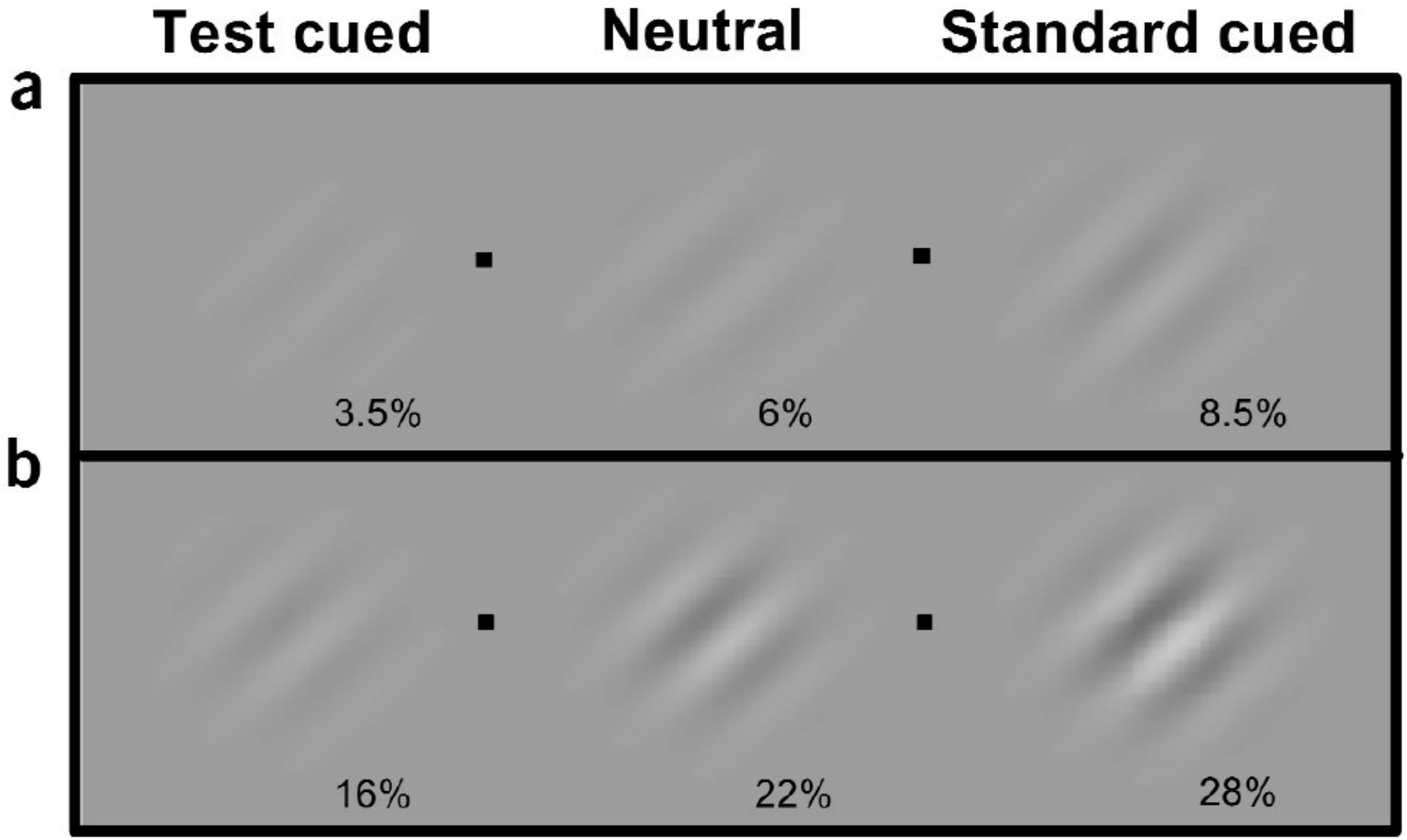
NOTE: fixed gaze!

Cues Affect Objects



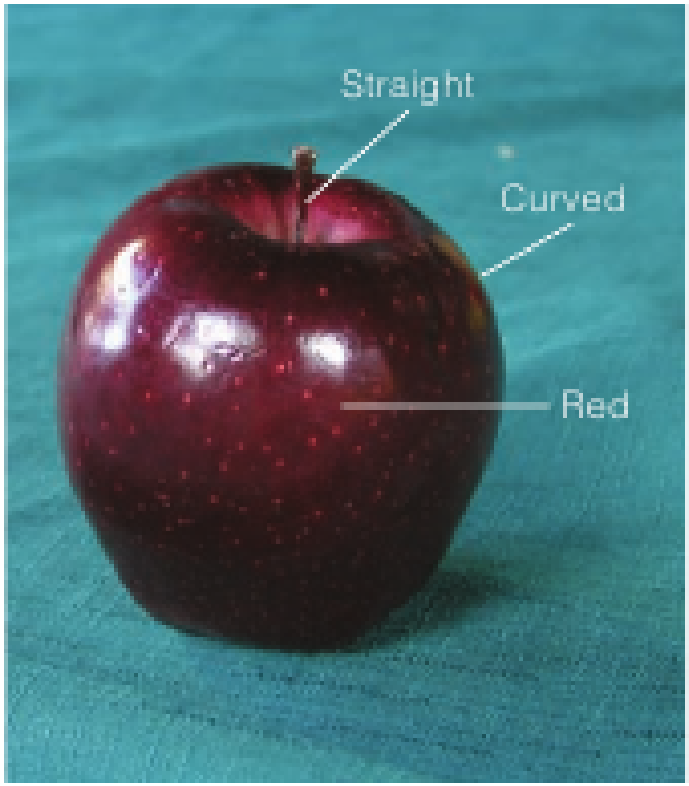
Egly et al. (1994)

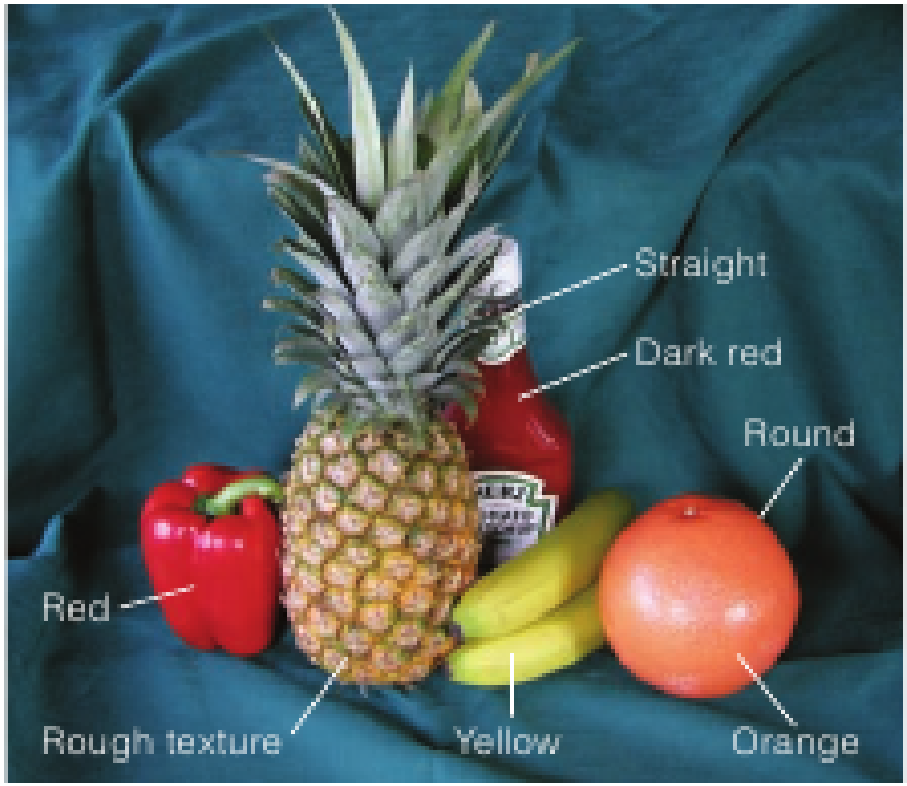
Are Attended Objects More Vivid?



Carrasco Et Al. (2004)

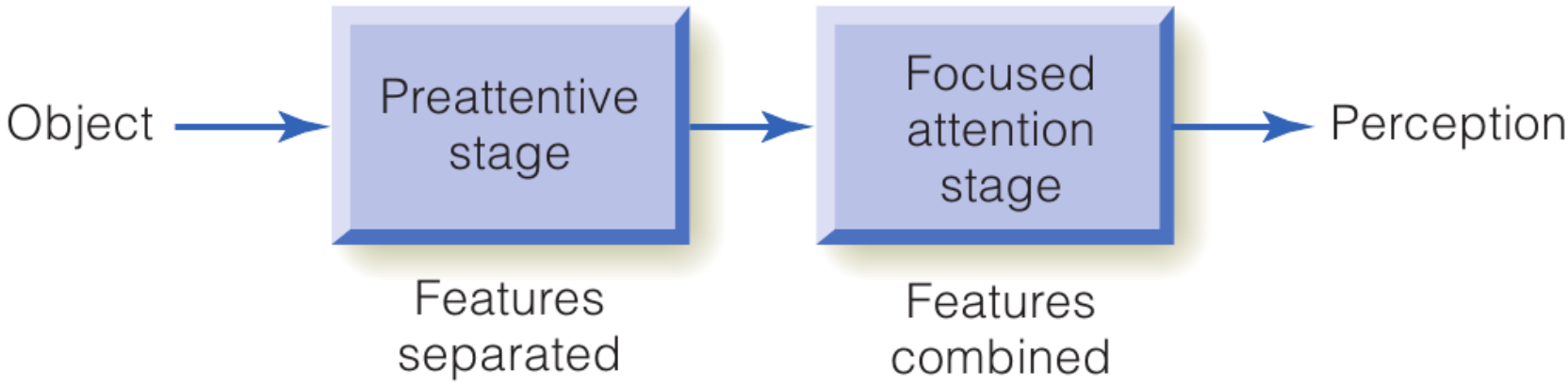
Binding Problem





shape, movement, colour, . . . put together

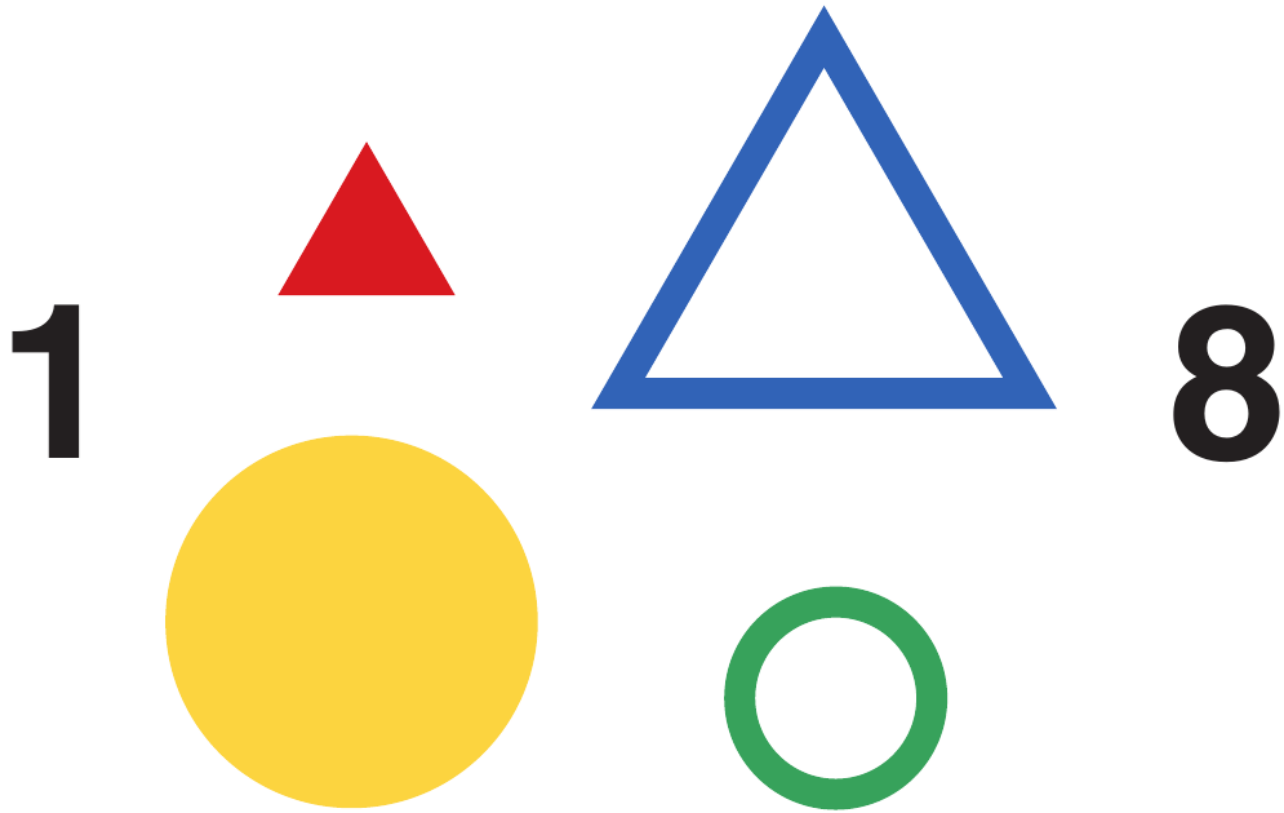
Feature Integration Theory



We need to focus our attention on each object in turn

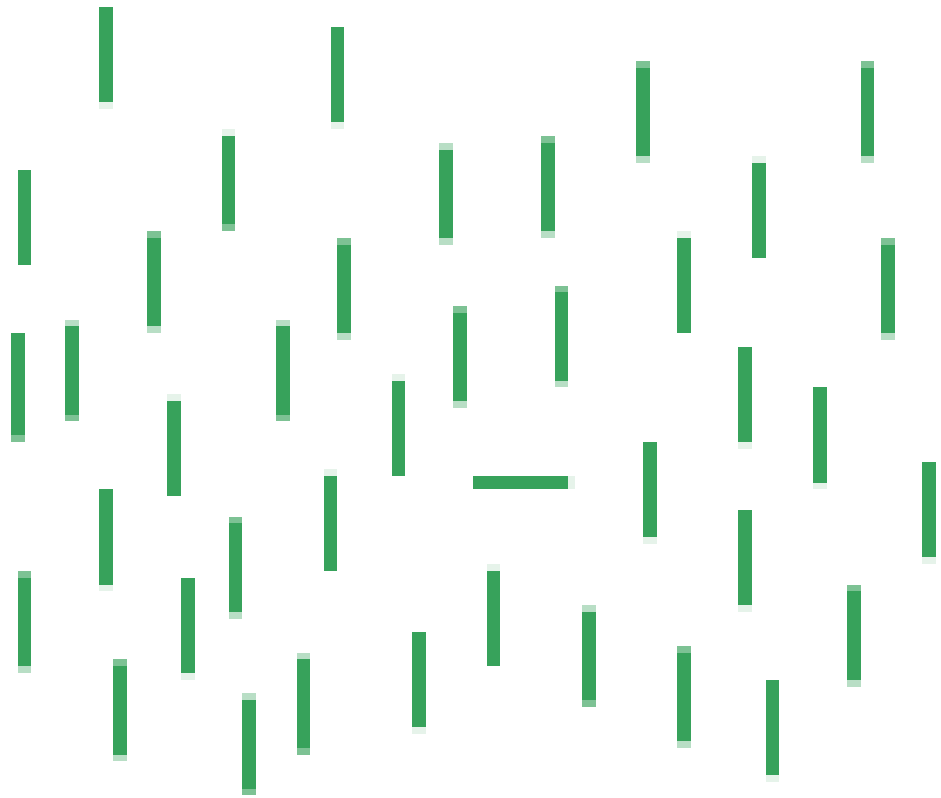
Illusory Conjunctions

associate features with wrong object

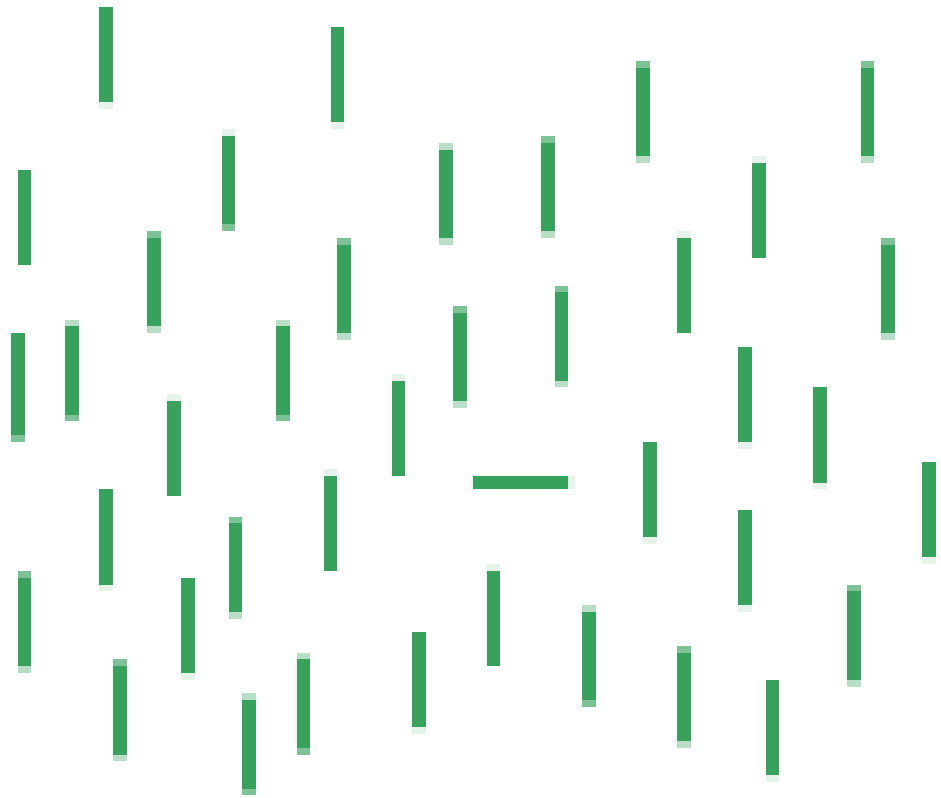


Treisman and Schmidt (1982)

Visual Search: No Conjunction

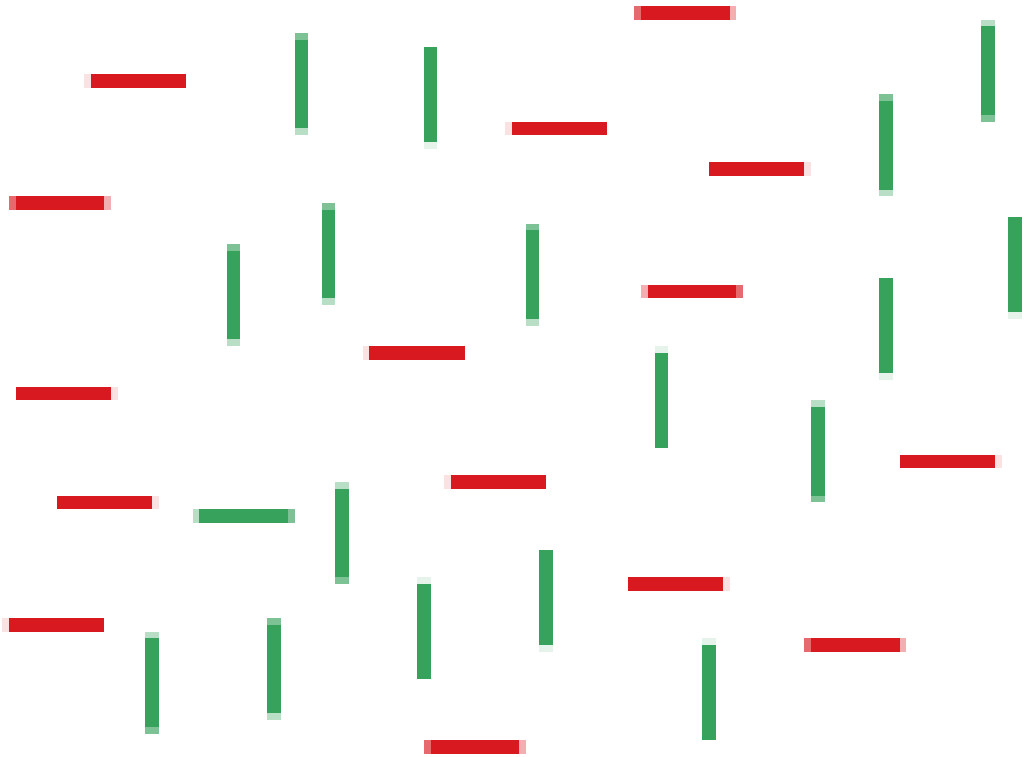


Visual Search: No Conjunction

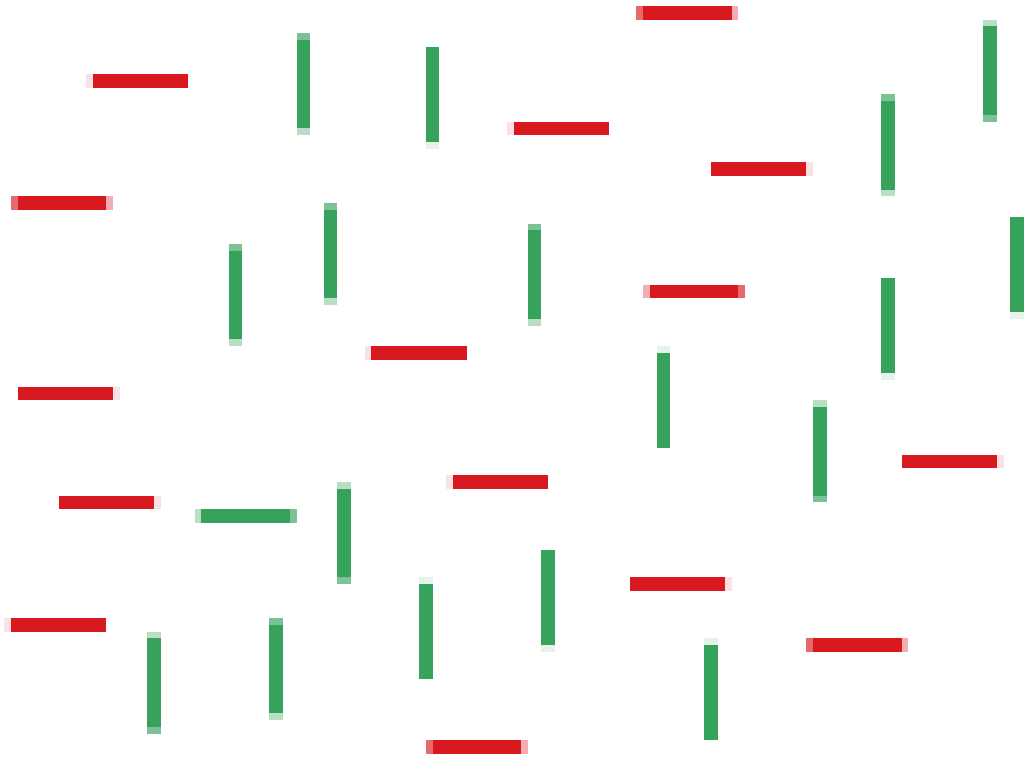


solution evident without focusing on a location

Visual Search: With Conjunction



Visual Search: With Conjunction

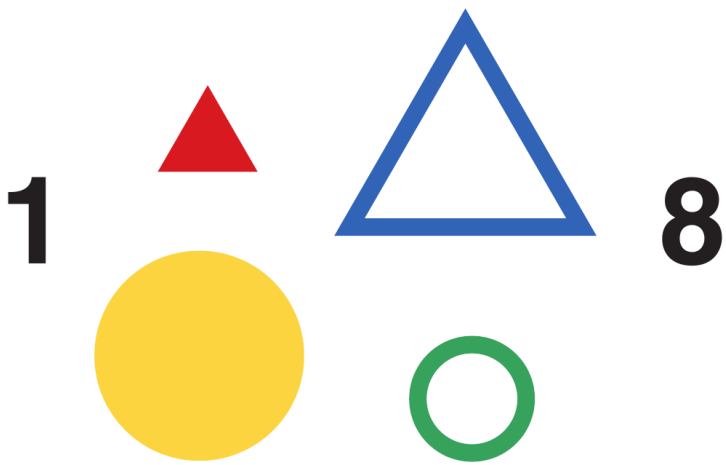


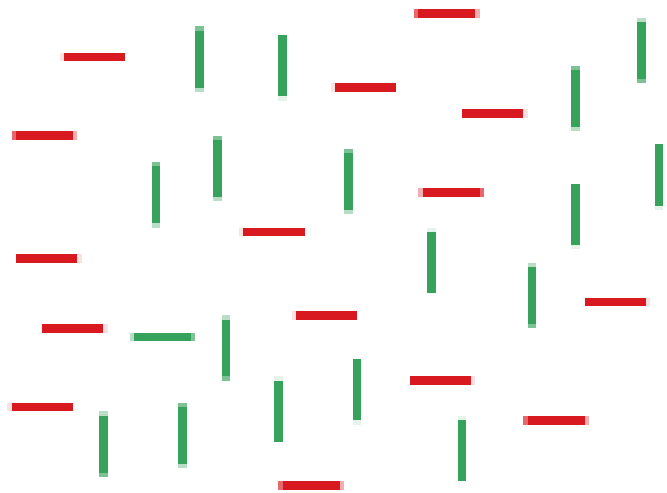
focusing on the location is necessary

Case of R.M.

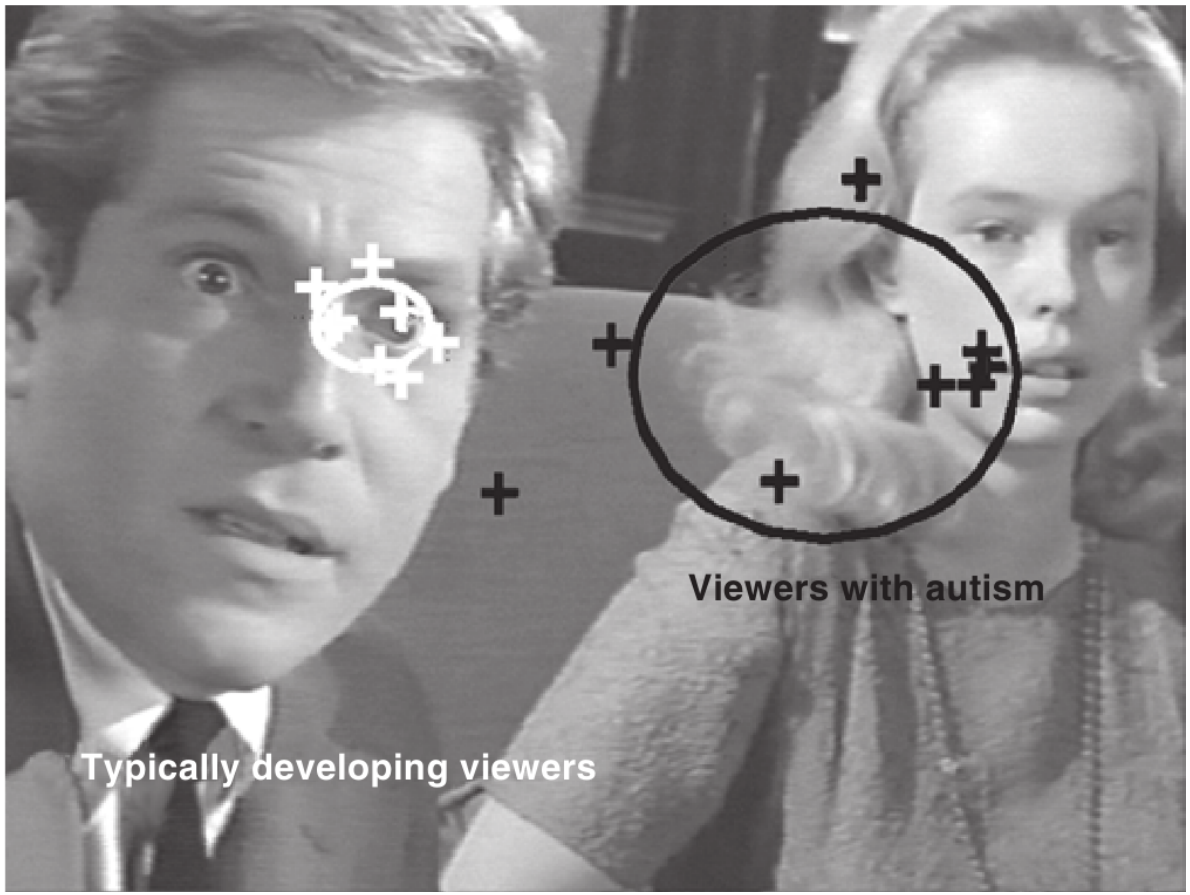
Patient with Balint's syndrome (inability to focus attention on individual objects)

- can not perform Treisman and Schmidt's task
- can not perform conjunction search





Attention and Autism



"Who is afraid of Virginia Woolf?" (1966)

Klin et al. (2003)

Attention and Autism

"Who is afraid of Virginia Woolf?" (1966)

Klin et al. (2003)

Outline

Perception outside the laboratory

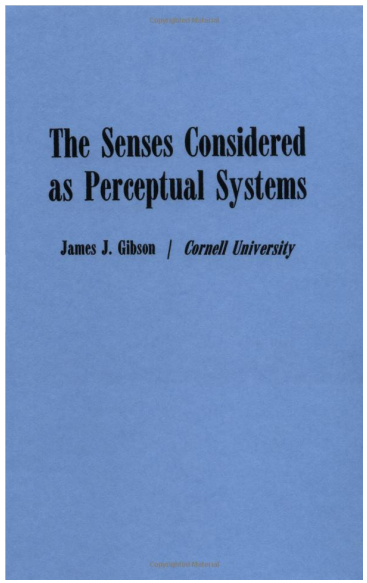
in real life:

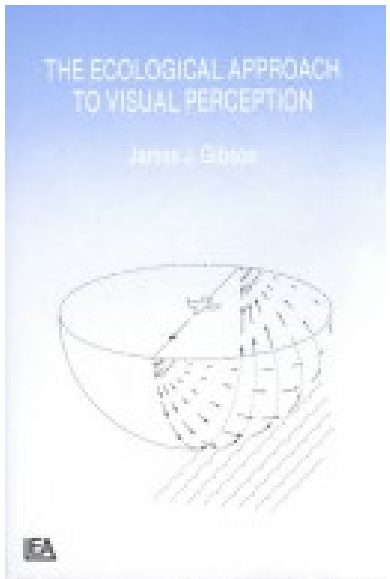
- perception and action tightly connected
- all senses jointly contribute to perception
- the observer's goal drives perception

Covered here, mainly:

1. navigation
2. reaching and grasping

James J. Gibson





Ecological Theory of Perception:

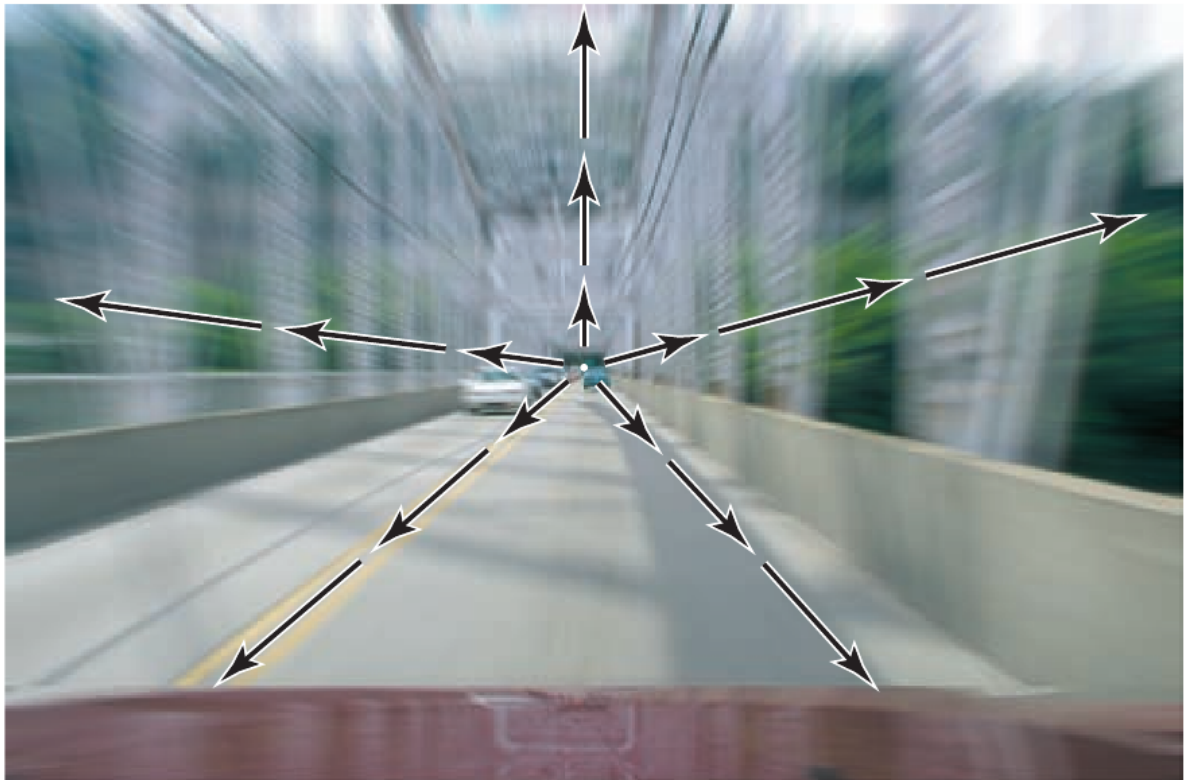
- moving observer
- look for information in the environment (out there)

Why is it important?

1. we evolved in a fast moving environment (hunting, being hunted)
2. the amount of information is overwhelming

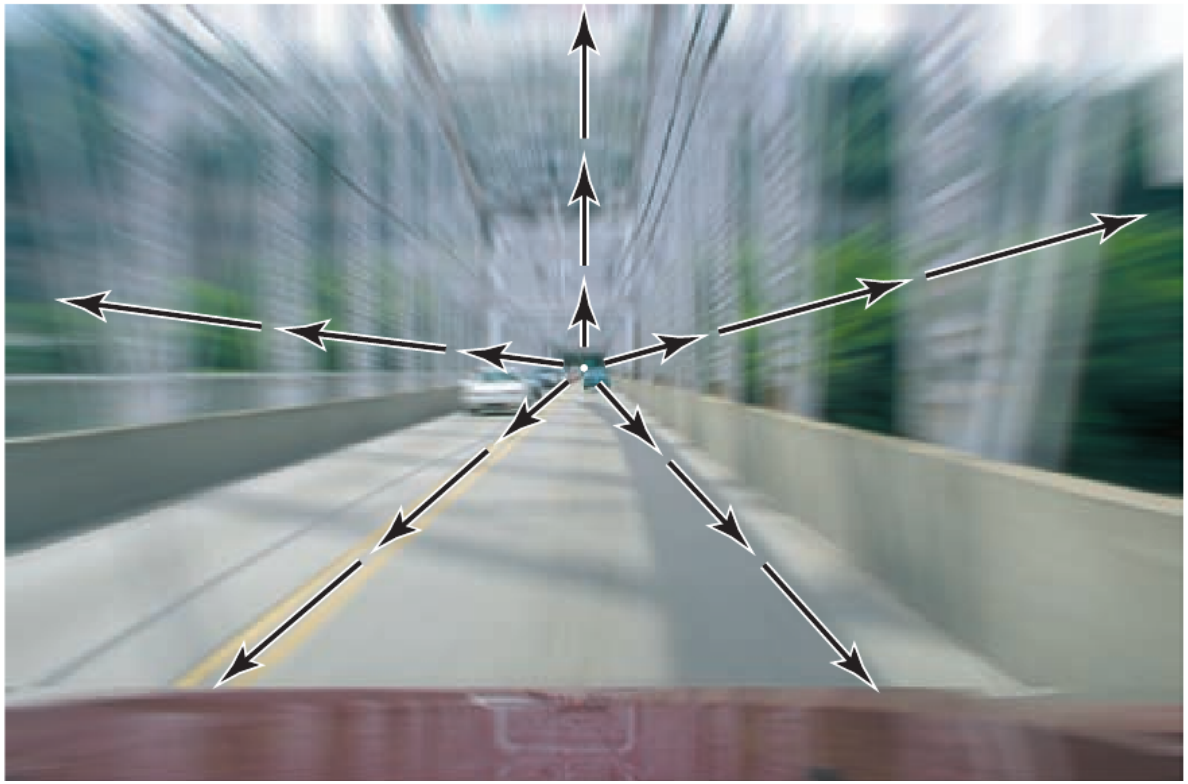
Imperative to find concise, relevant and invariant representations of the world

Optic Array and Optic Flow



Optic Array: structure in surfaces, textures and contours
Optic Flow: movement of elements in the Optic Array

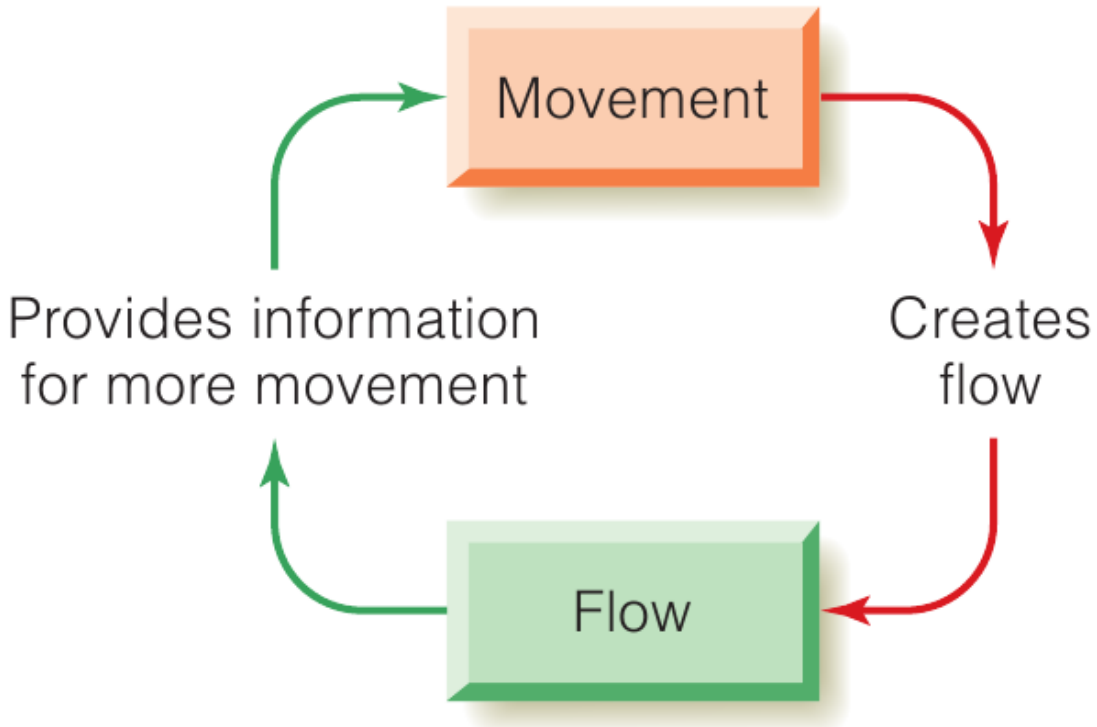
Optic Flow



- more rapid near the observer (gradient of flow, info on speed)
- no flow at the destination: focus of expansion
- invariant information

Self Produced Information

"we need to perceive to move and we need to move to perceive"



Example: Vertical Parallax

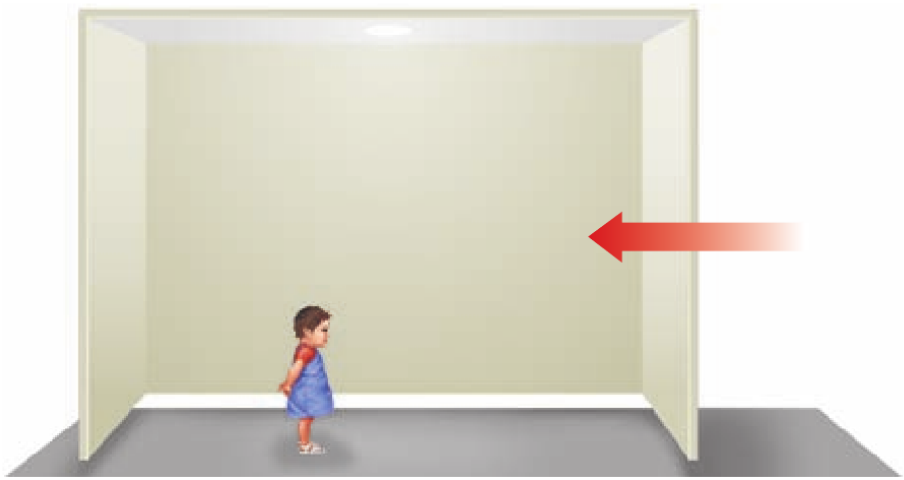
Walking produces an ondulatory vertical motion of the head

This helps perceive depth

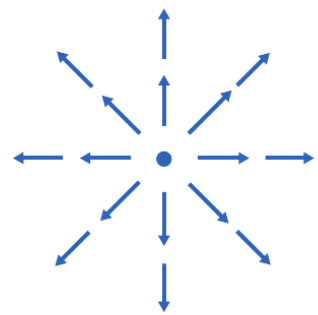
Example: grasshopper

The Senses do not Work in Isolation

balance example



(a) Room swings toward person.

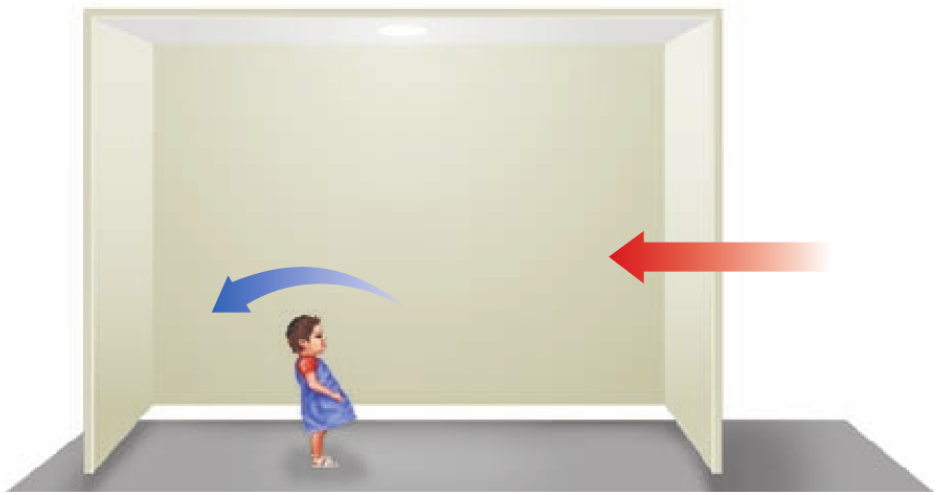


Flow when wall is moving toward person.

Lee and Aronson (1974)

The Senses do not Work in Isolation

balance example

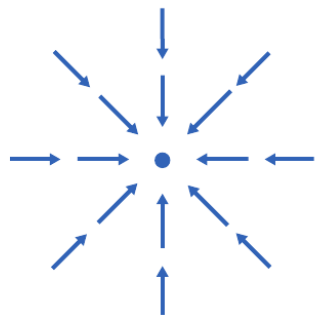


(b) Person sways back to compensate.

Lee and Aronson (1974)

The Senses do not Work in Isolation

balance example



Flow when wall is moving away from person.

(c) When room swings away, person sways forward to compensate.

Lee and Aronson (1974)

The Senses do not Work in Isolation

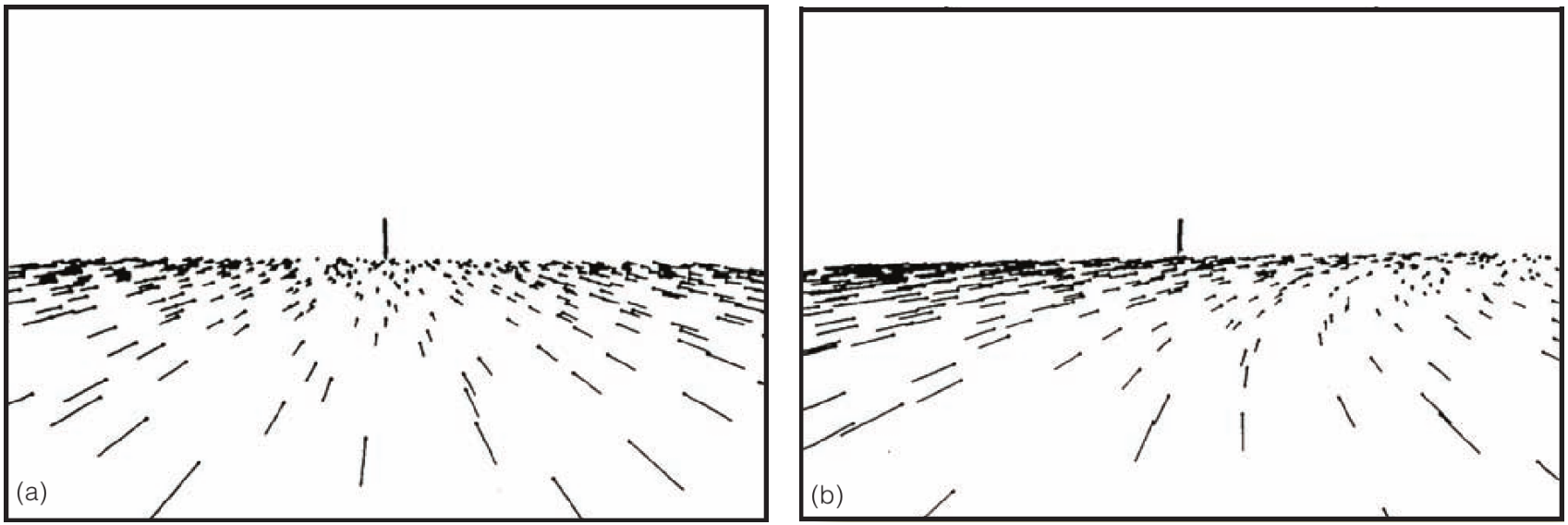
sensory integration:

- vestibular system
- sensors at joints and muscles
- vision

Vision is powerful:

- toddlers: 26% swayed, 23% staggered, 33% fell down
- adults: swayed with just 6mm movement
- starts early: 4 months

Optic Flow and Navigation

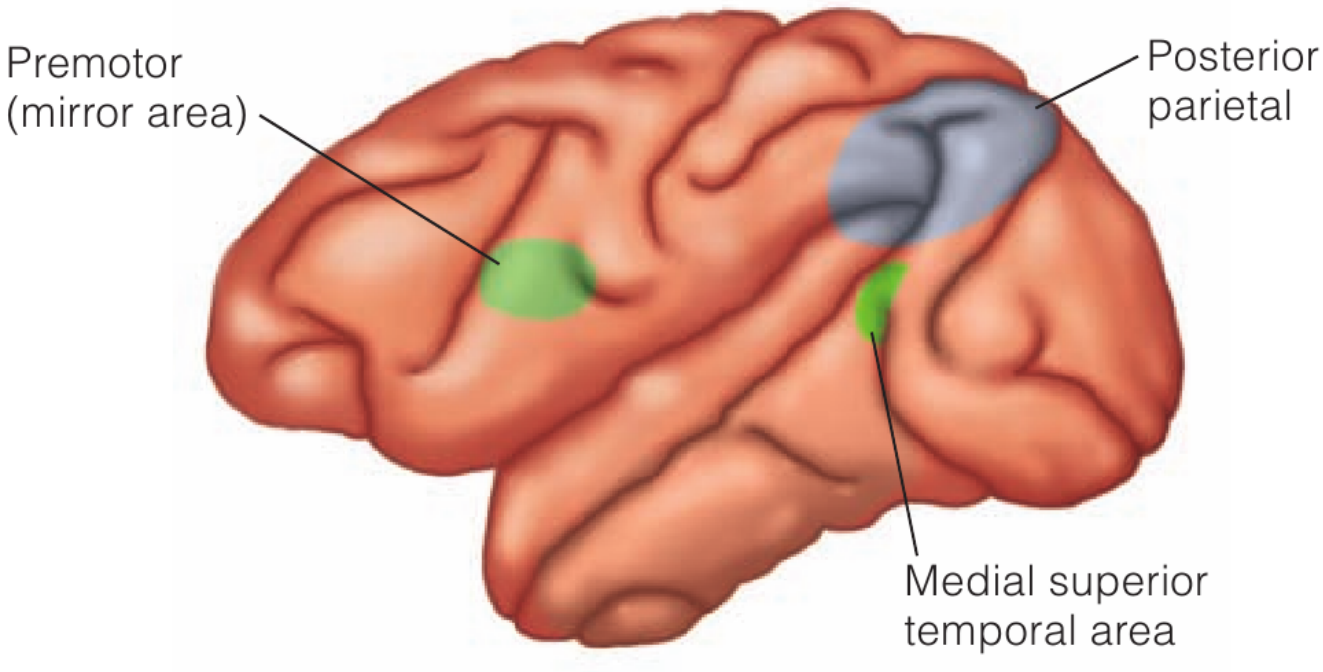


subjects guess the heading within 0.5-1.0 degrees

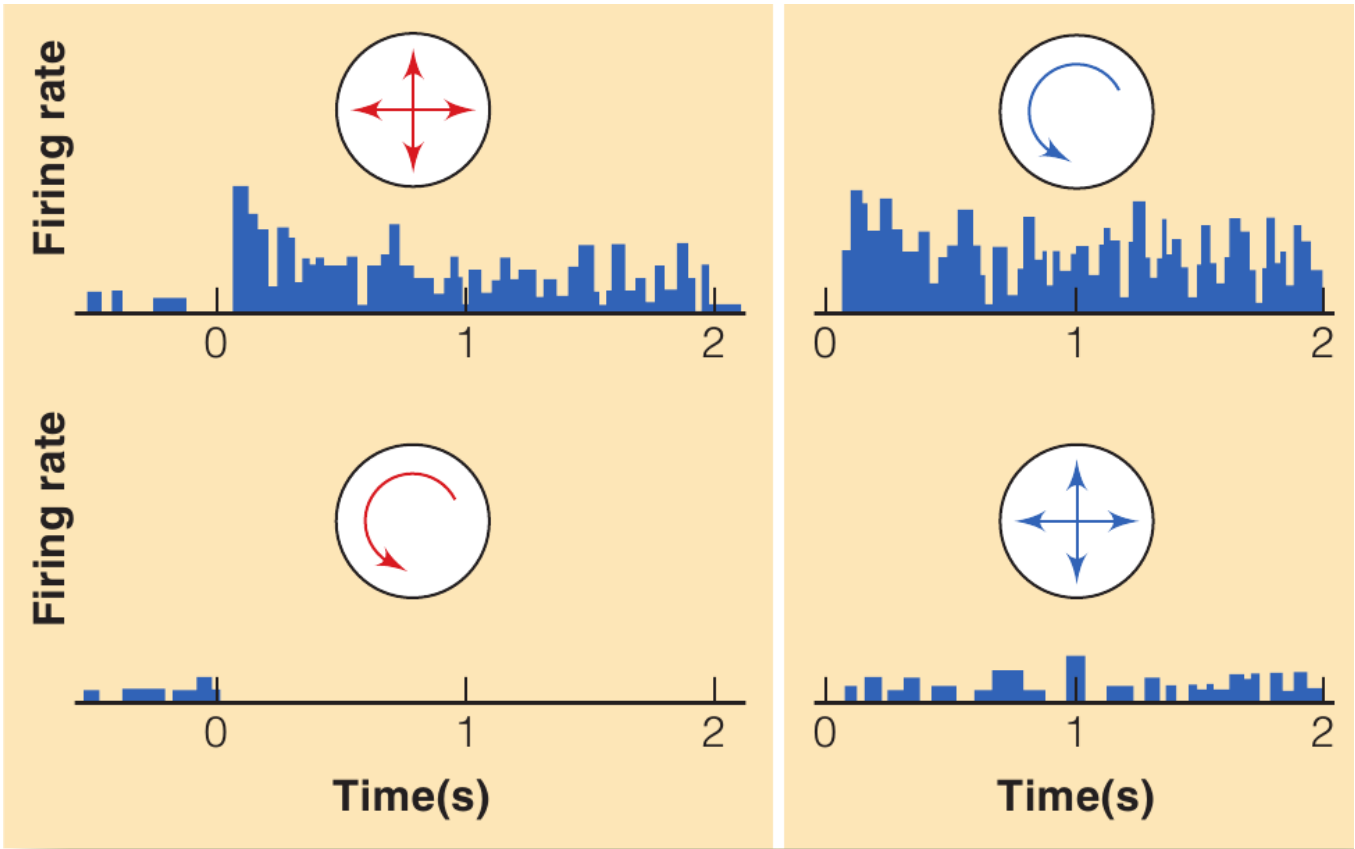
Warren (1995)

Physiology: Optic Flow Neurons

medial superior temporal area (monkey)



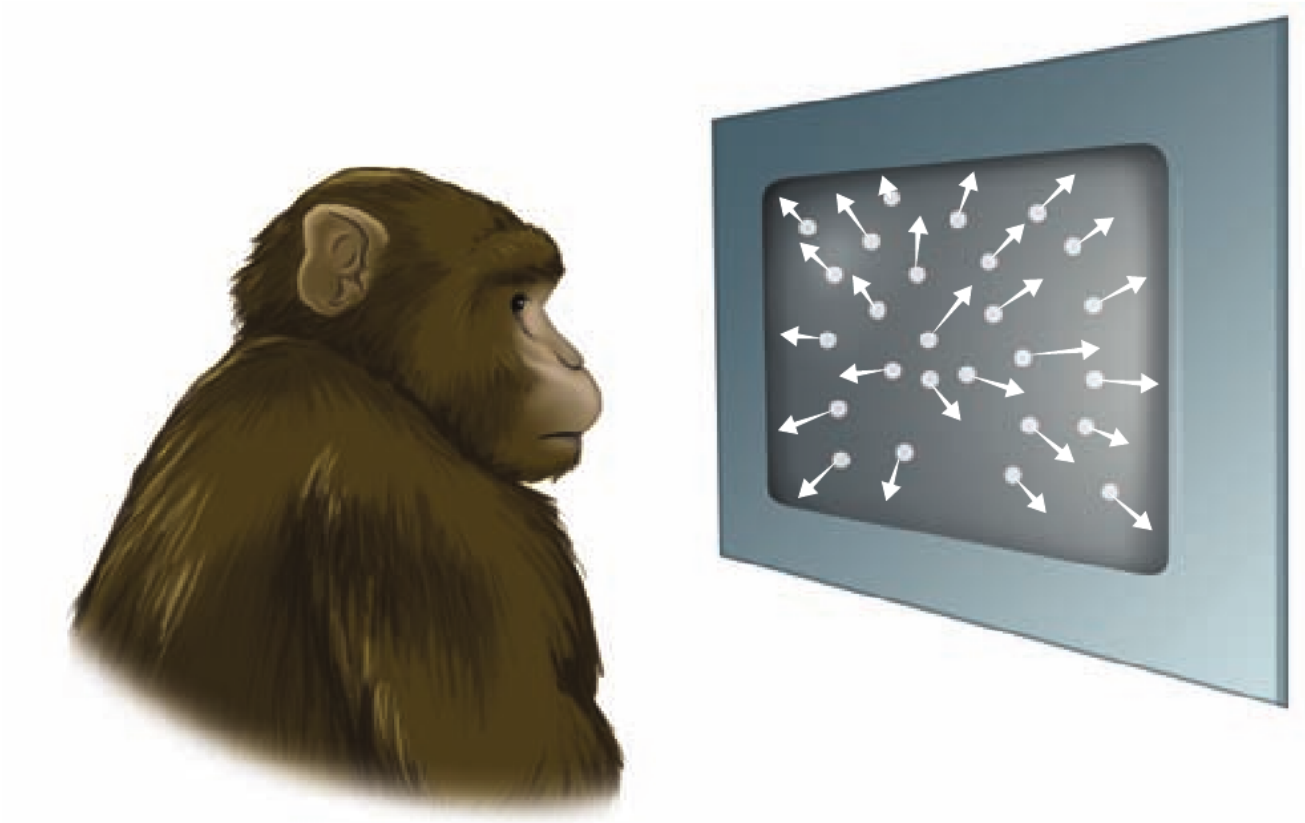
Experiments with Monkeys



Graziano et al. (1994)

How does this affect perception?

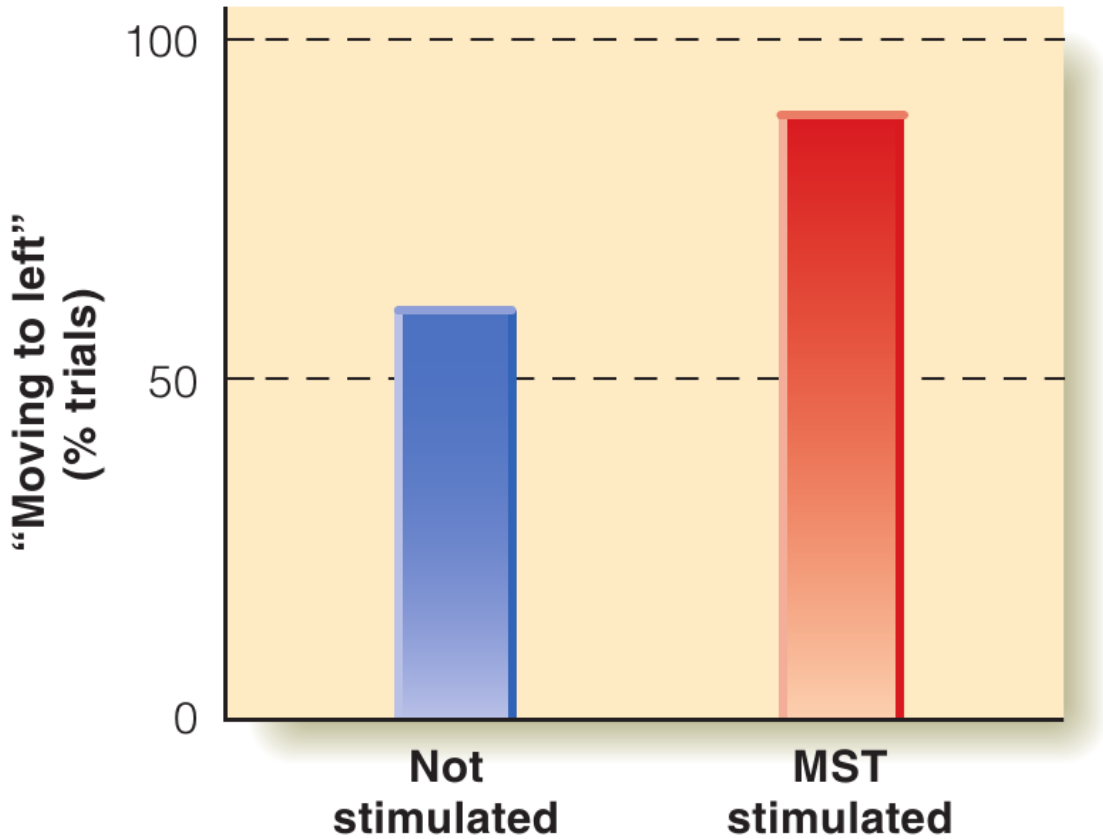
Stimulating neurons in Medial Superior Temporal (MST) area



Britten and van Wezel (2002)

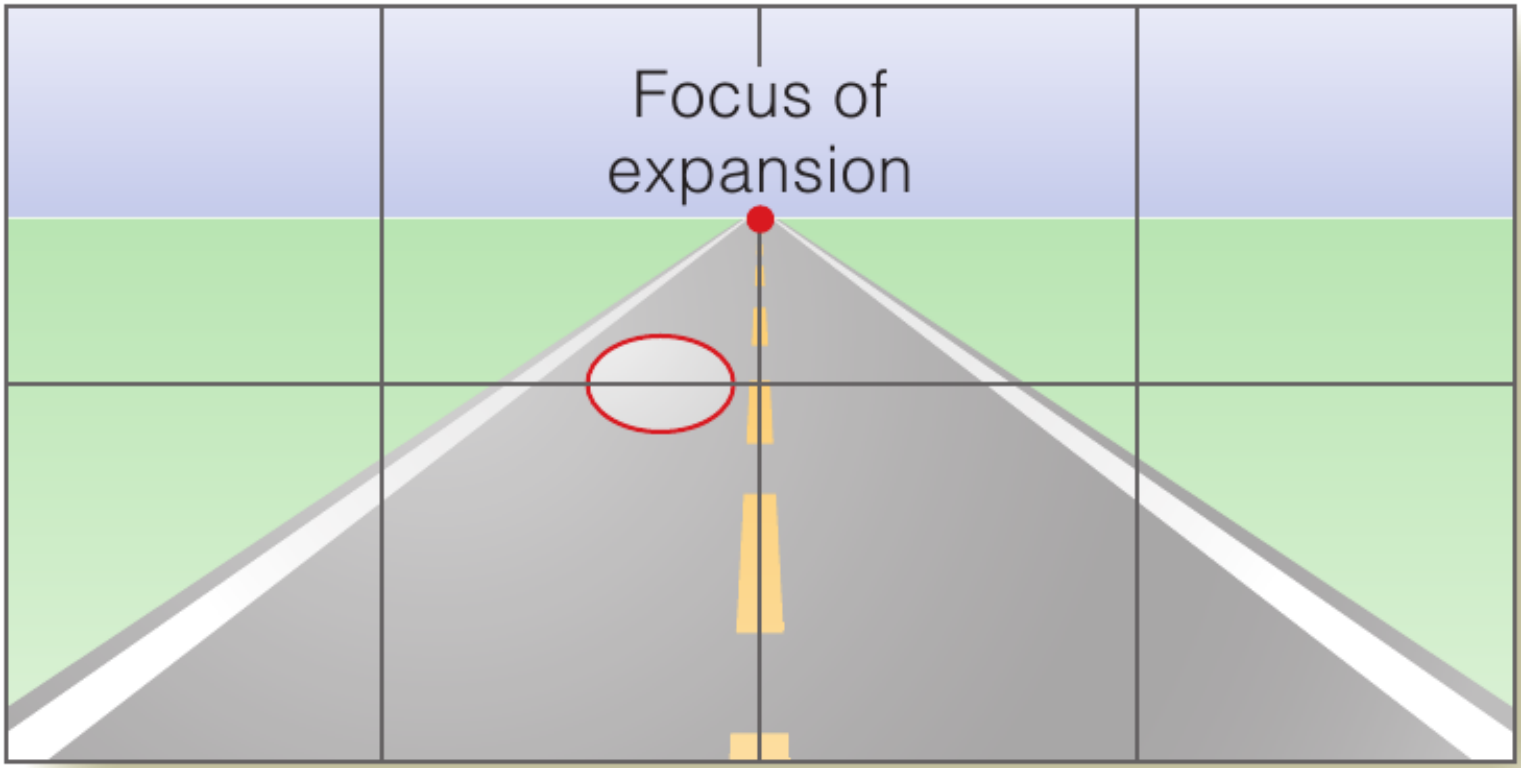
How does this affect perception?

Stimulating neurons in Medial Superior Temporal (MST) area



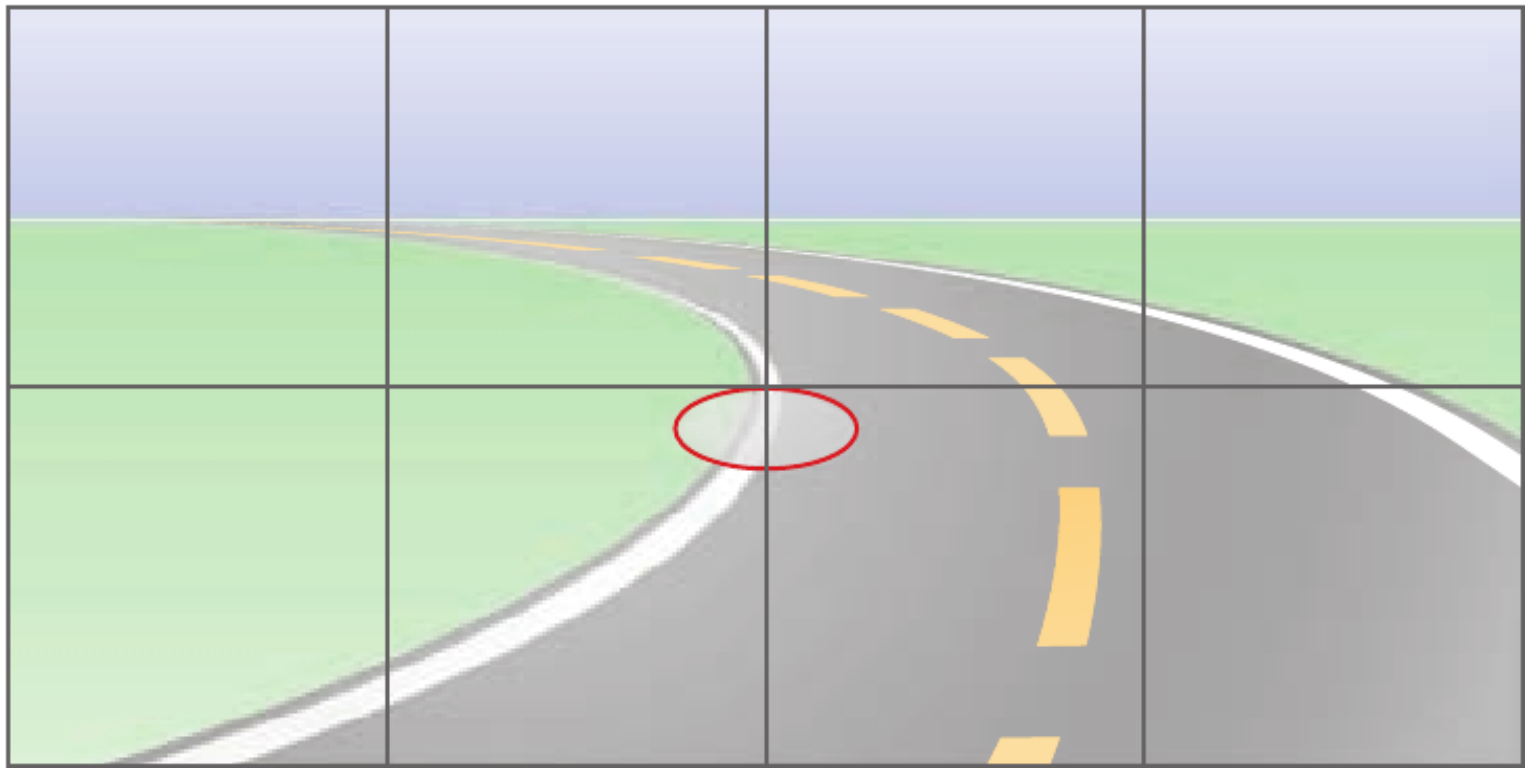
Britten and van Wezel (2002)

Not only Optic Flow: Gaze in Driving



Sinai et al. (1998)

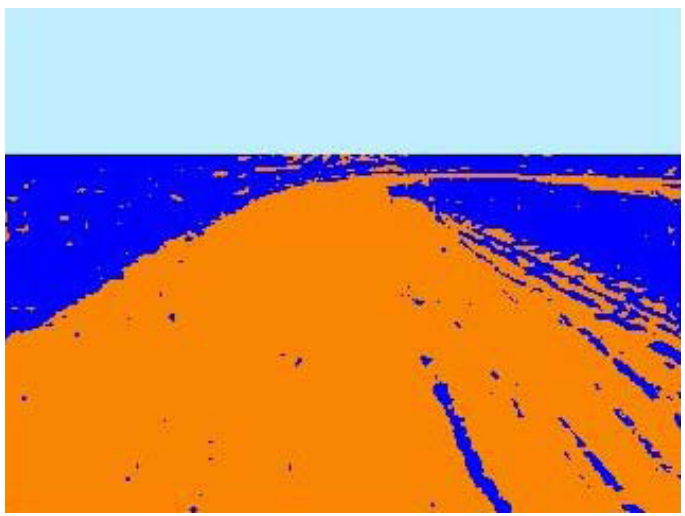
Not only Optic Flow: Gaze in Driving



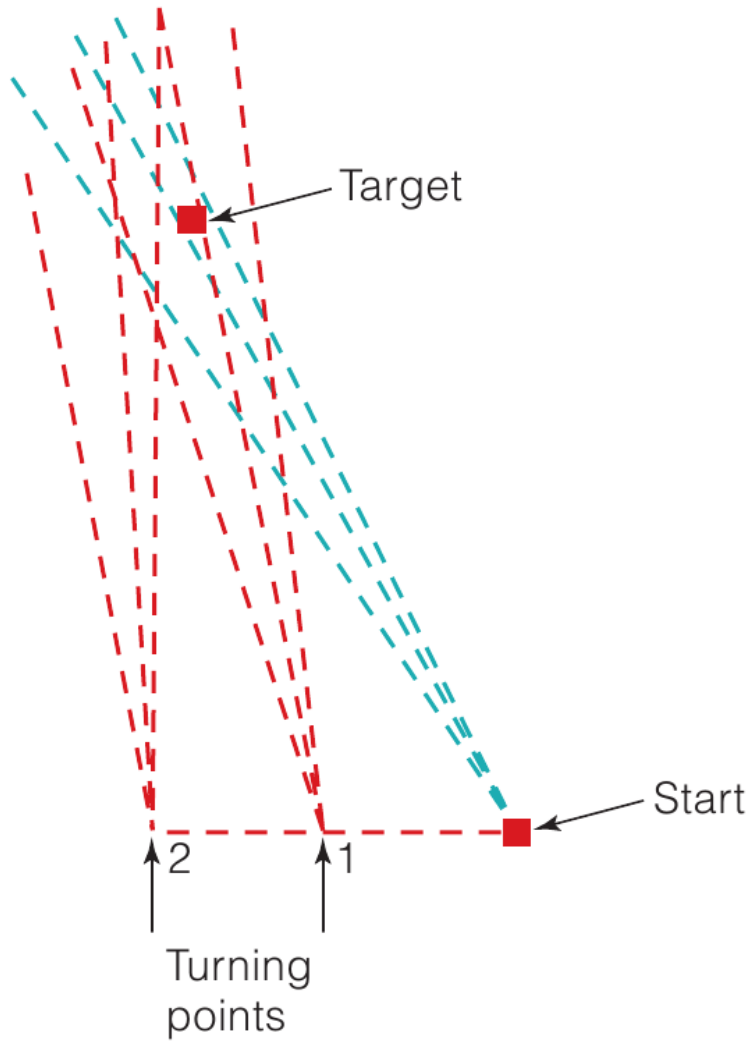
Sinai et al. (1998)

Example: autonomous vehicles

- DARPA Grand Challenge
- vehicles equipped with a number of sensors
- estimate road from camera input
- using colours too complex (lighting, different environments)

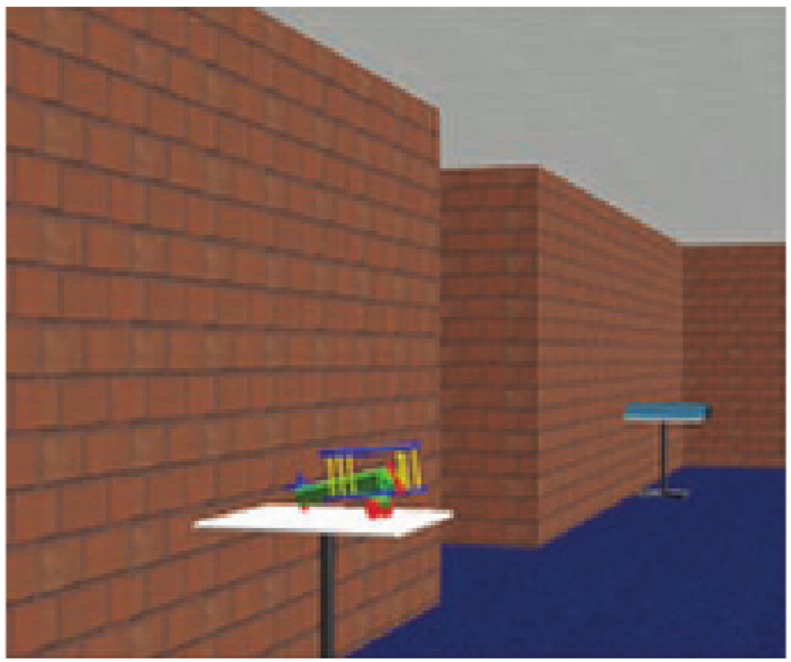


Not only Optic Flow: Blinded Walking

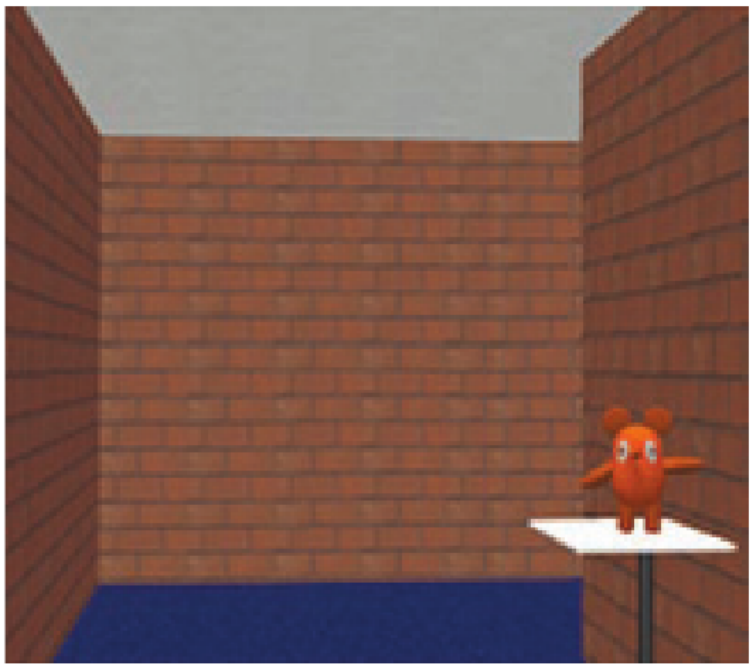


Philbeck et al. (1997)

More Navigation: Landmarks



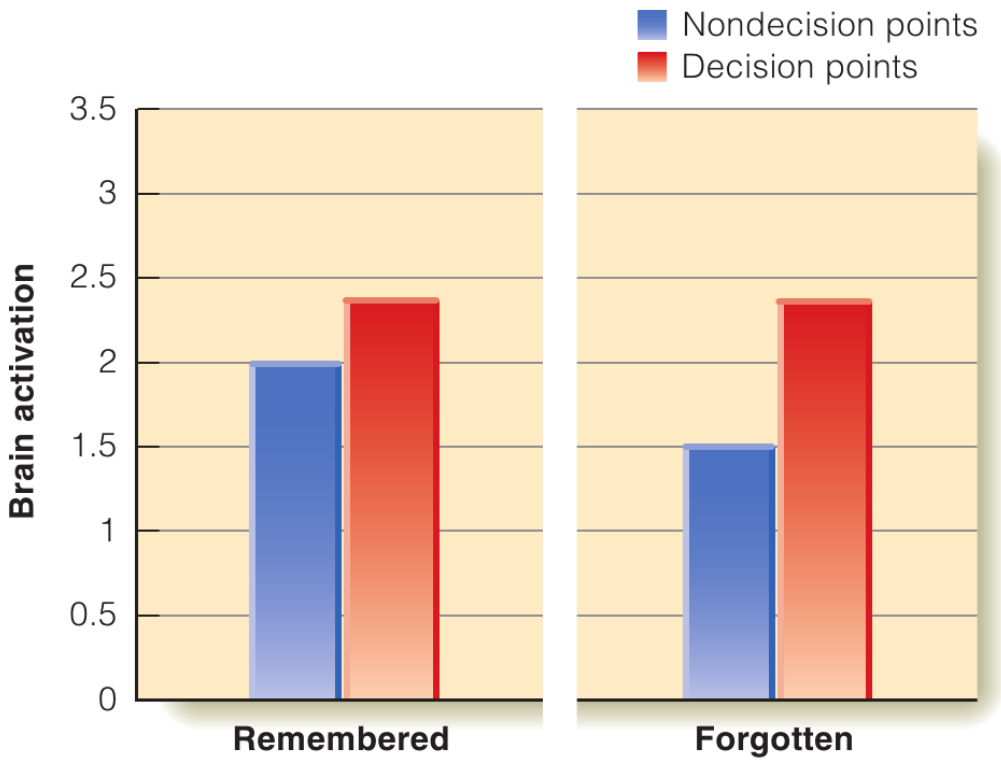
(a) Toy at decision point



(b) Toy at nondecision point

Jansen and van Turennout (2004)

More Navigation: Landmarks



Jansen and van Turennout (2004)

Acting on Objects: Affordances

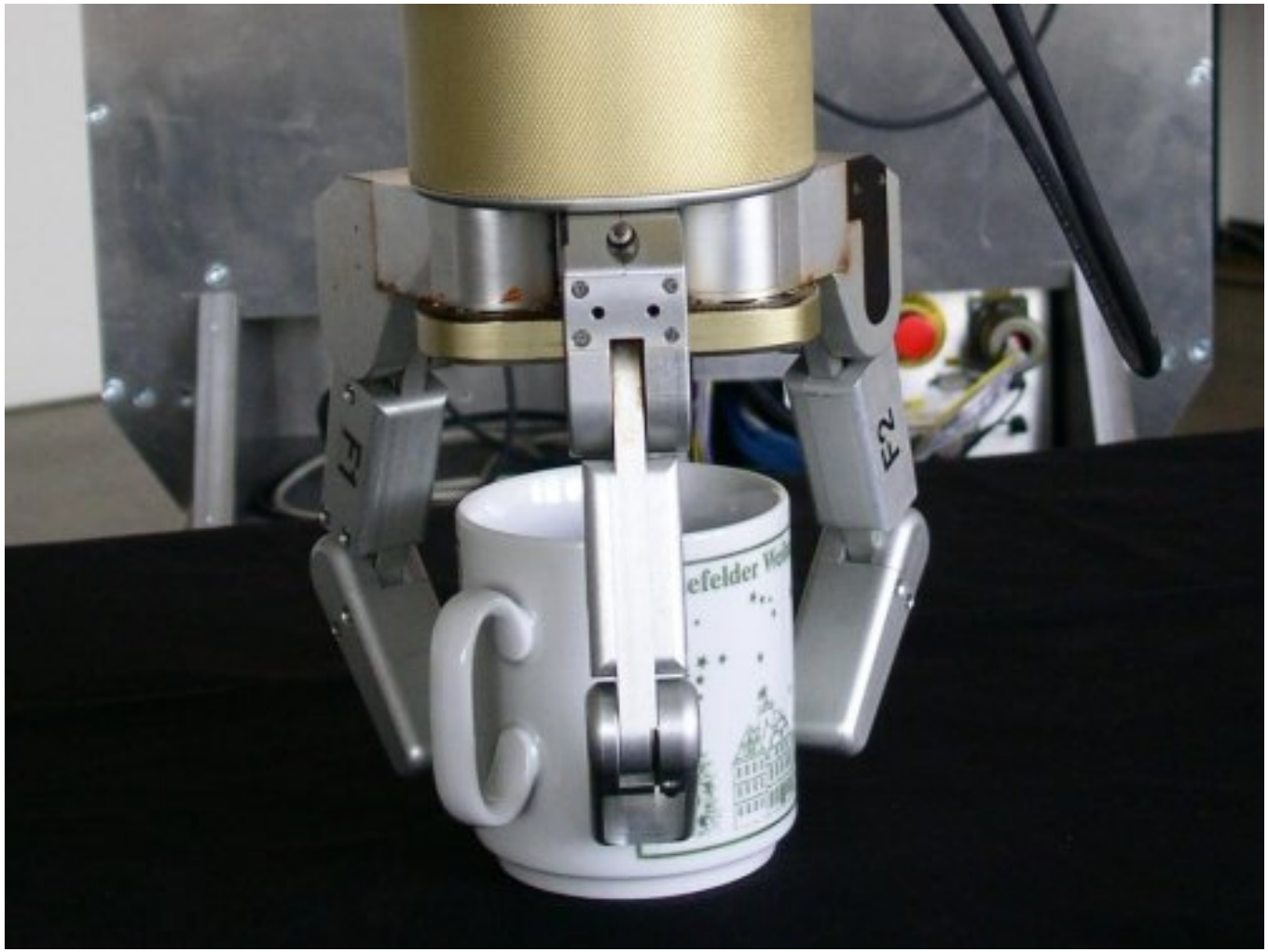
Acting on Objects: Affordances

They afford sitting

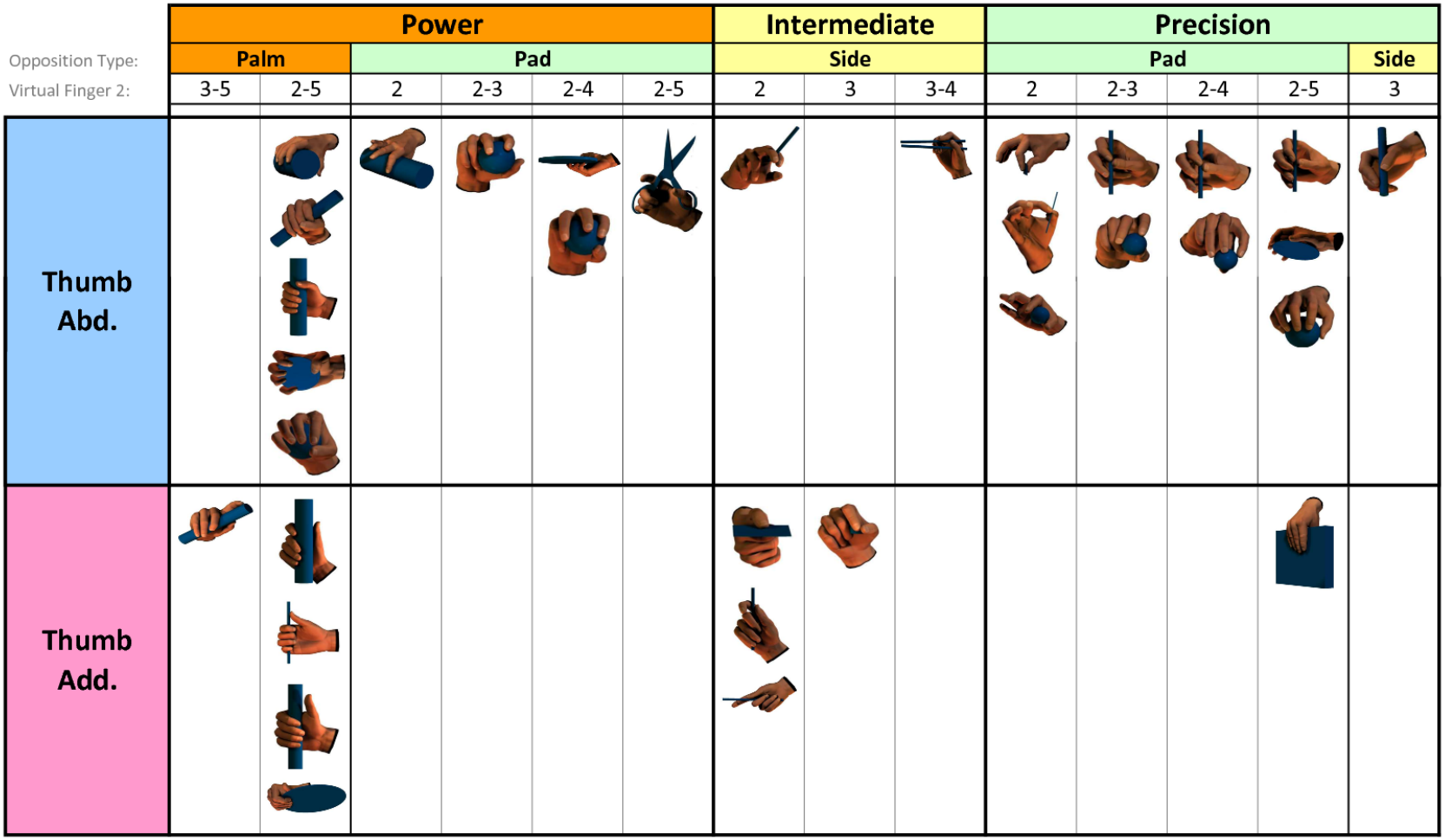
Reaching and Grasping

Reaching and Grasping

A huge research question in robotics

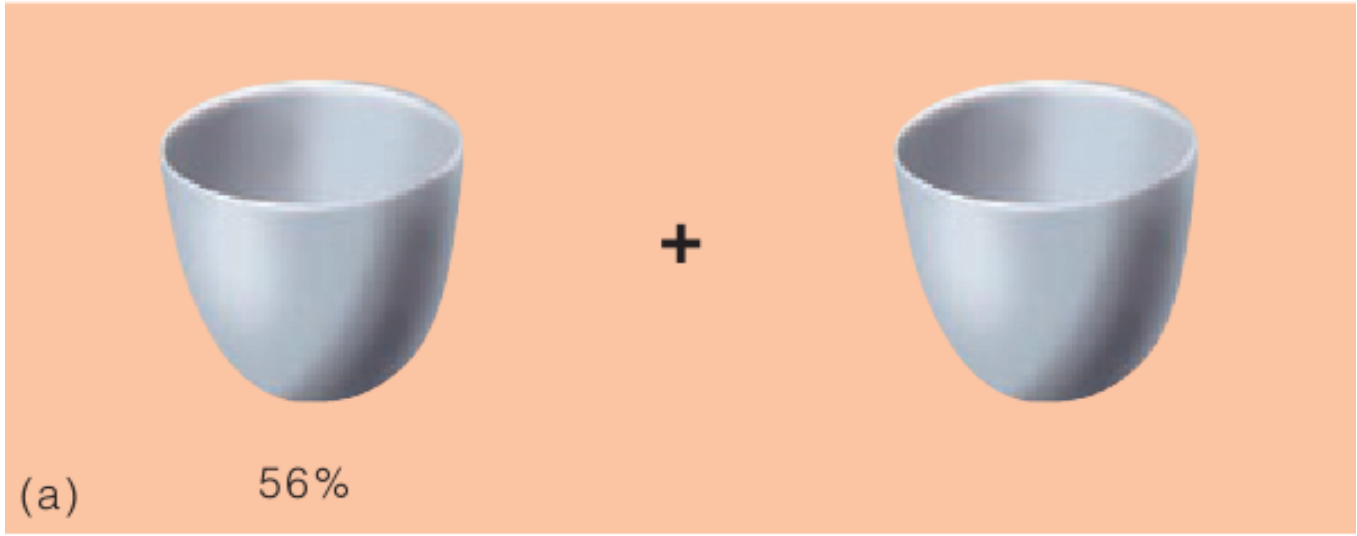


Grasp Taxonomy



Affordances in Reaching and Grasping

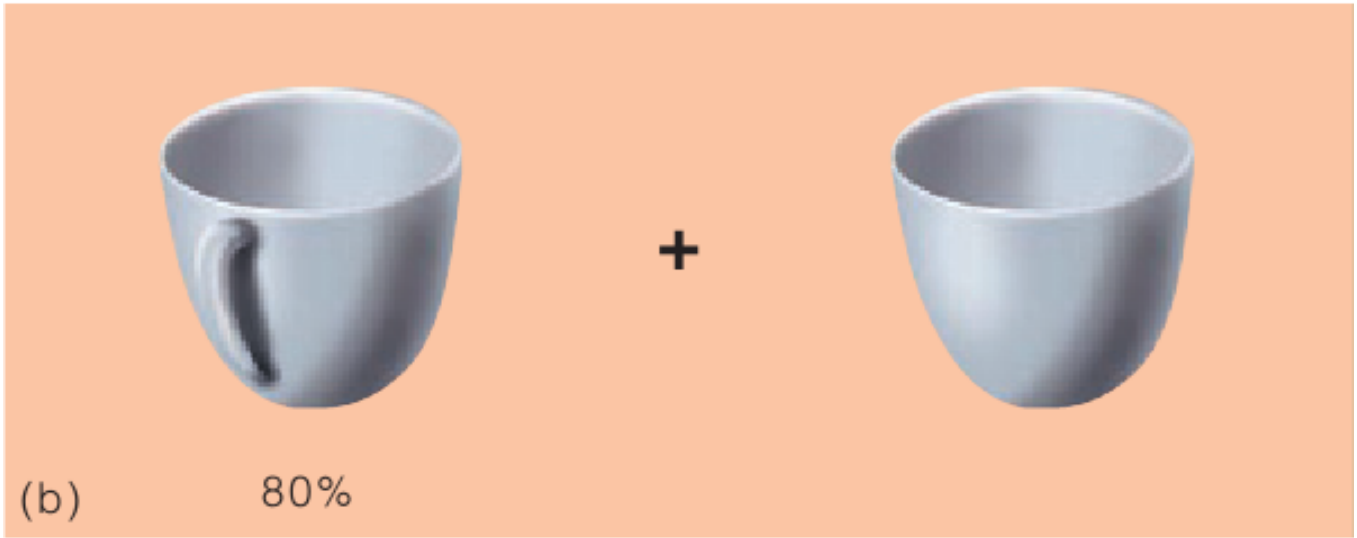
Affordance test



Patient with Extinction (only right object detected)

Di Pellegrino et al. (2005)

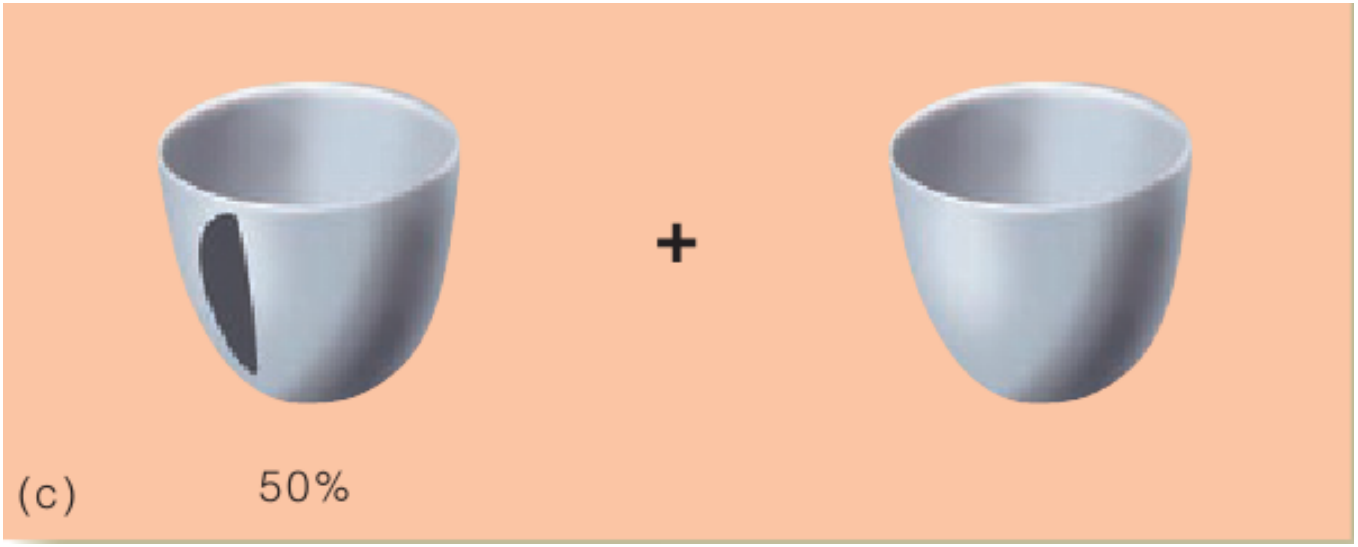
Affordance test



Patient with Extinction (only right object detected)

Di Pellegrino et al. (2005)

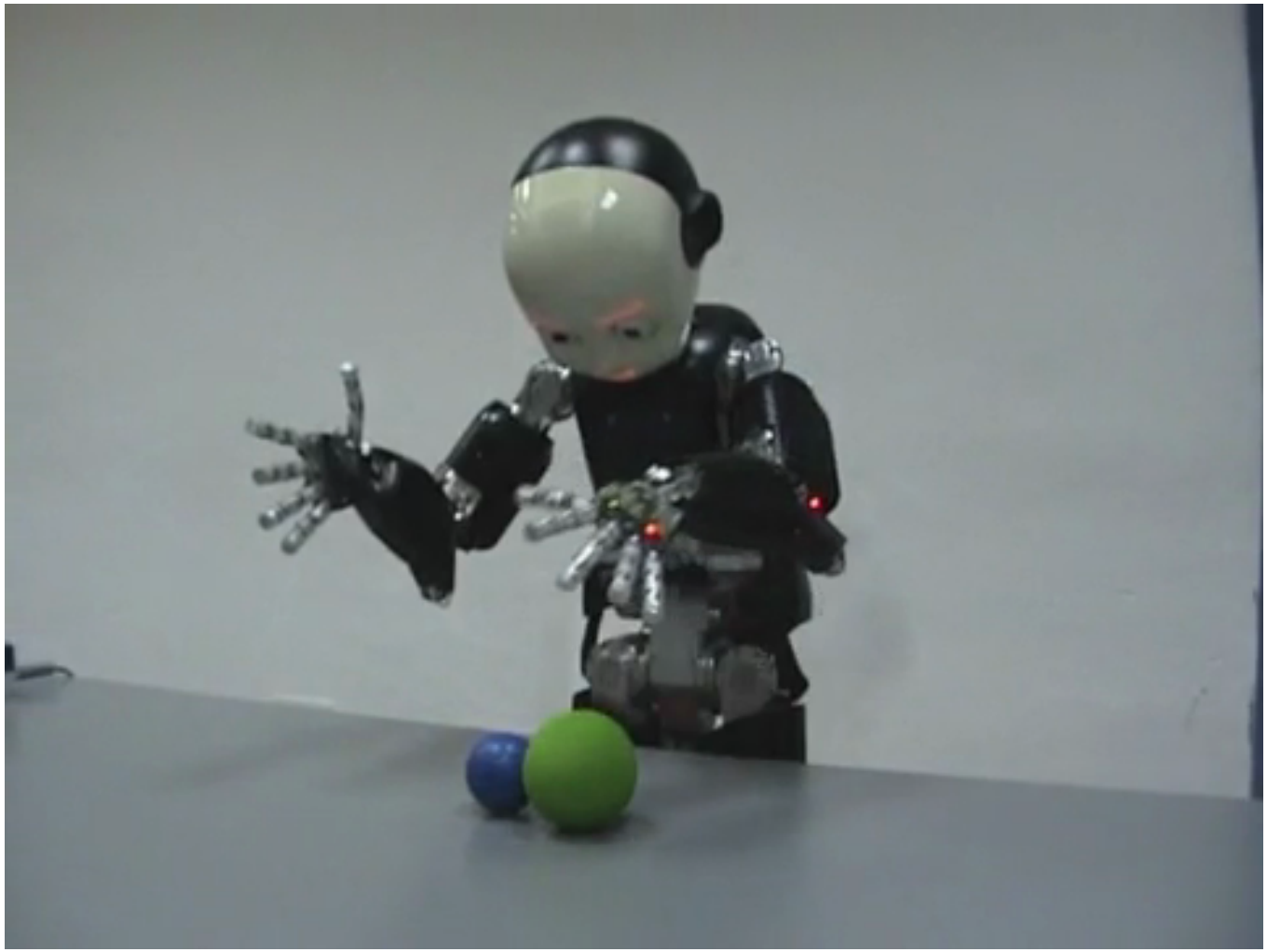
Affordance test



Patient with Extinction (only right object detected)

Di Pellegrino et al. (2005)

Example: words and affordances [2]

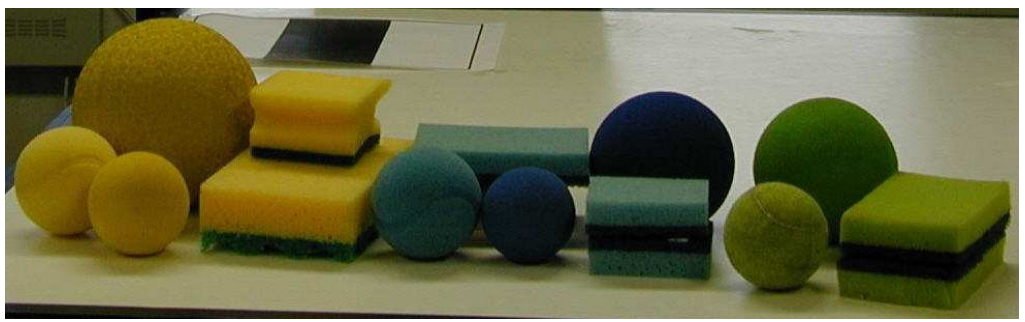


[2] G. Salvi, L. Montesano, A. Bernardino, and J. Santos-Victor. "Language bootstrapping: Learning word meanings from perception-action association". In: *IEEE Trans. Syst., Man, Cybern. B* 42.3 (June 2012), pp. 660–671

Example: words and affordances

Find associations between:

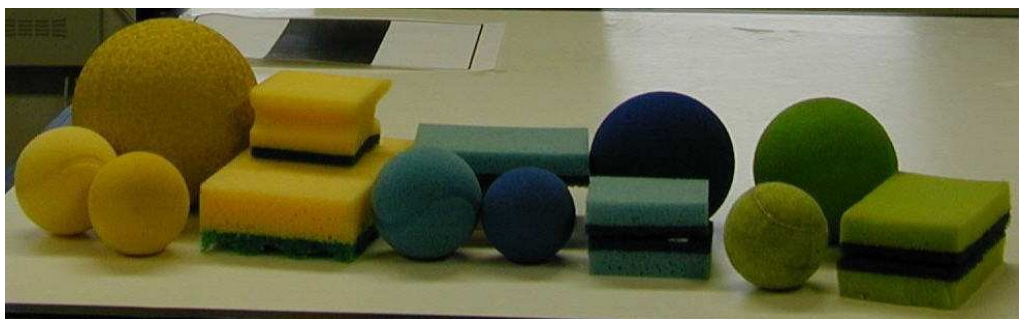
- actions (tap, touch, grasp)
- object properties (shape, size, color)
- effects (obj vel, obj-hand vel, contact...)
- spoken words (*"the robot grasps the ball, but the ball falls"*)



Example: words and affordances

Find associations between:

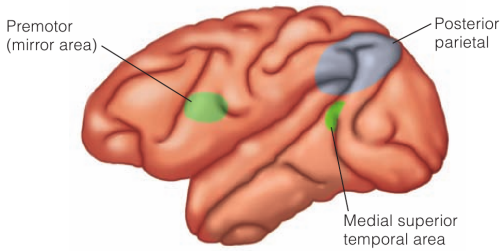
- actions (tap, touch, grasp)
- object properties (shape, size, color)
- effects (obj vel, obj-hand vel, contact...)
- spoken words (*"the robot grasps the ball, but the ball falls"*)



the meaning of words is grounded into the robots action/perception world

Physiology of Reaching and Grasping

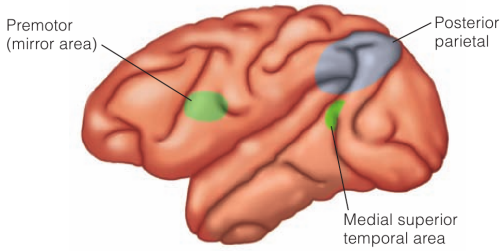
Neurons in the parietal cortex of the monkey respond to goal-directed reaching



- they respond if the monkey is reaching to achieve a goal (obtain food)
- they do not respond for same movement without goal
- they respond even before the monkey reaches for the object

Physiology of Reaching and Grasping

Neurons in the parietal cortex of the monkey respond to goal-directed reaching

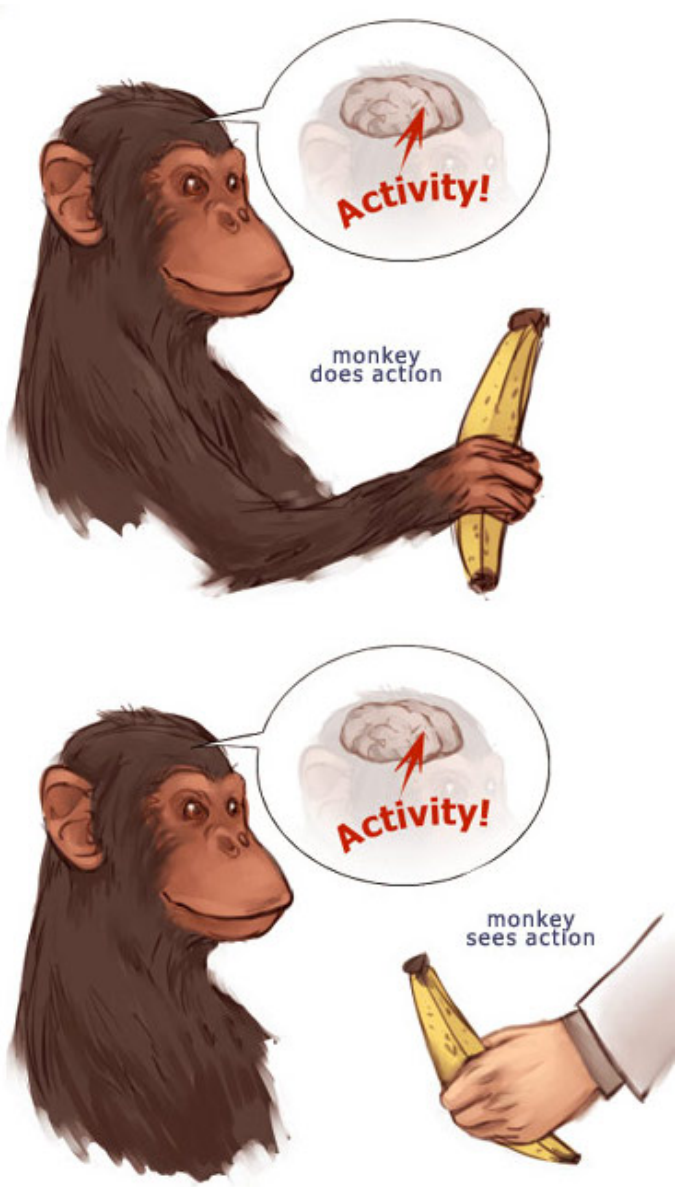


- they respond if the monkey is reaching to achieve a goal (obtain food)
- they do not respond for same movement without goal
- they respond even before the monkey reaches for the object
- they respond when observeing others achieving the goal

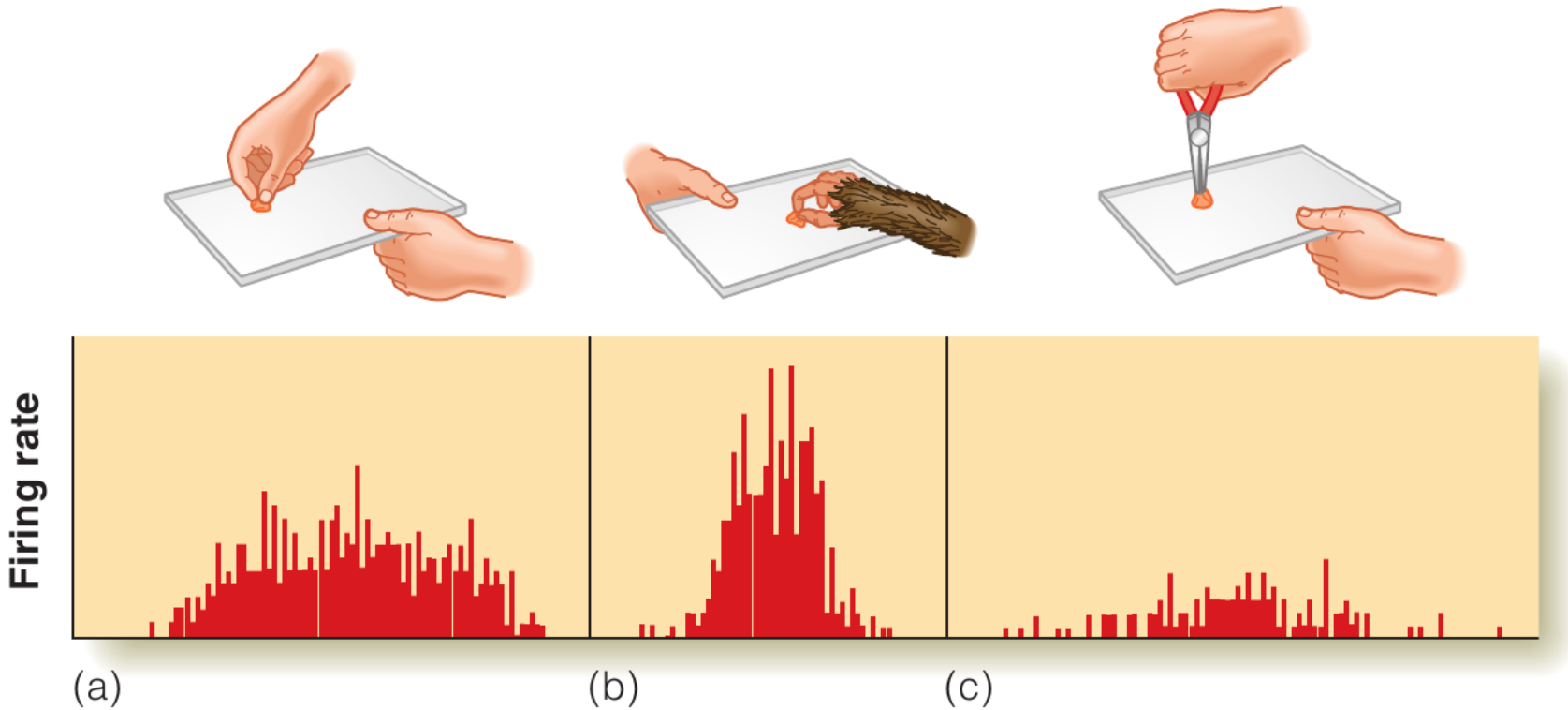
Anecdote

Rizzolatti's laboratory in Ferrara, Italy

a researcher took a lunch break...

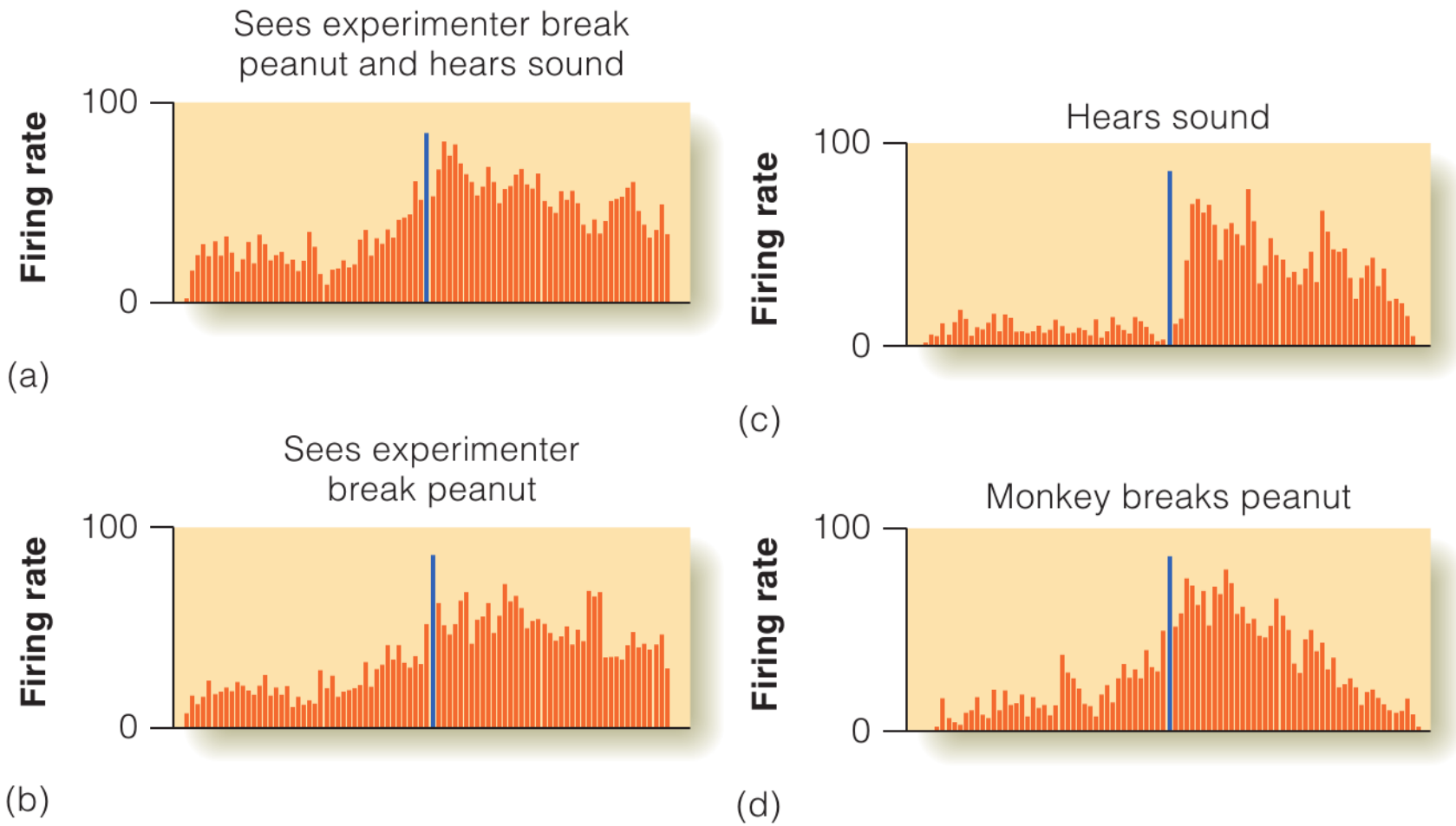


Observing Other People's Actions: Mirror Neurons



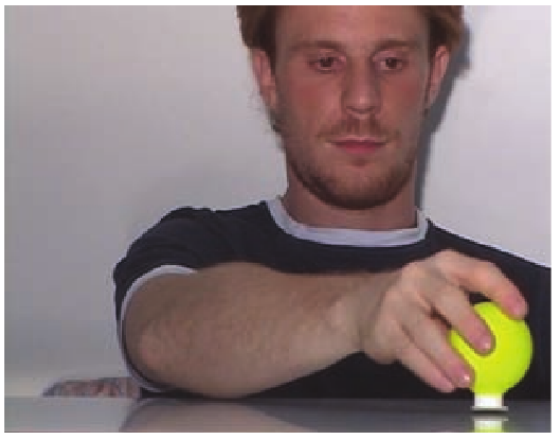
Rizzolatti et al. (2000)

Audiovisual Mirror Nowrons

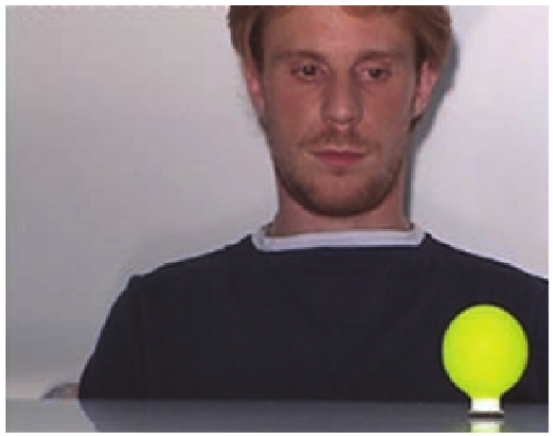


Kohler et al. (2002)

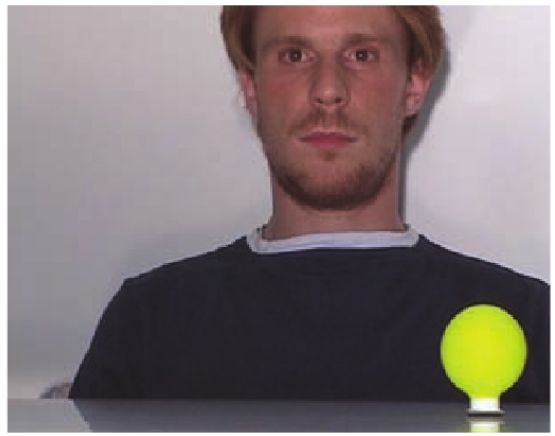
Predicting People's Intentions



(a) Grasp



(b) Gaze

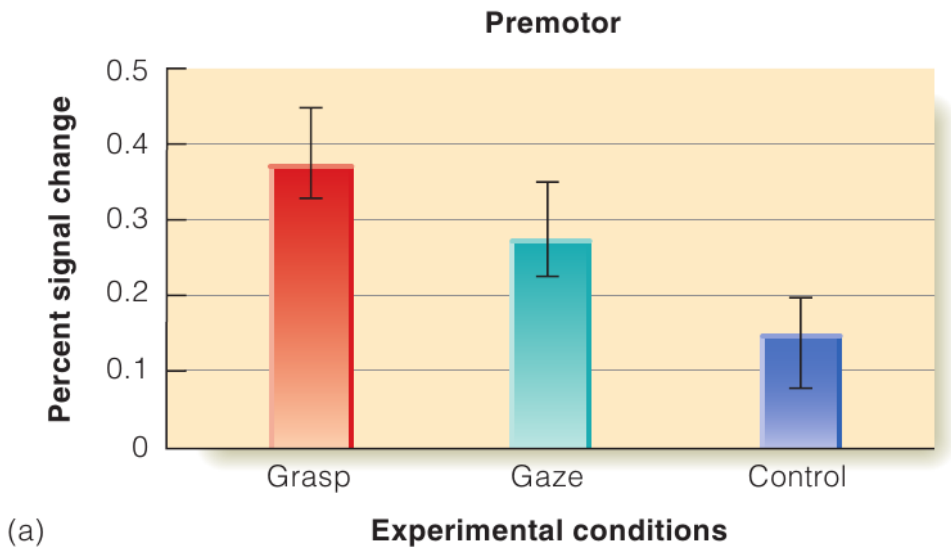


(c) Control

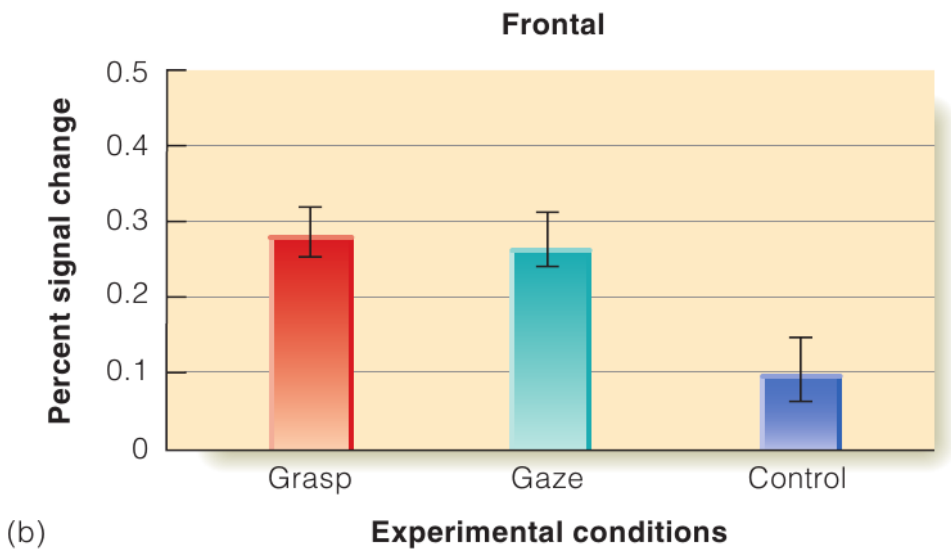
Record activity in the action observation system in the brain (mirror neurons, premotor cortex. . .)

Pierno et al. (2006)

Predicting People's Intentions

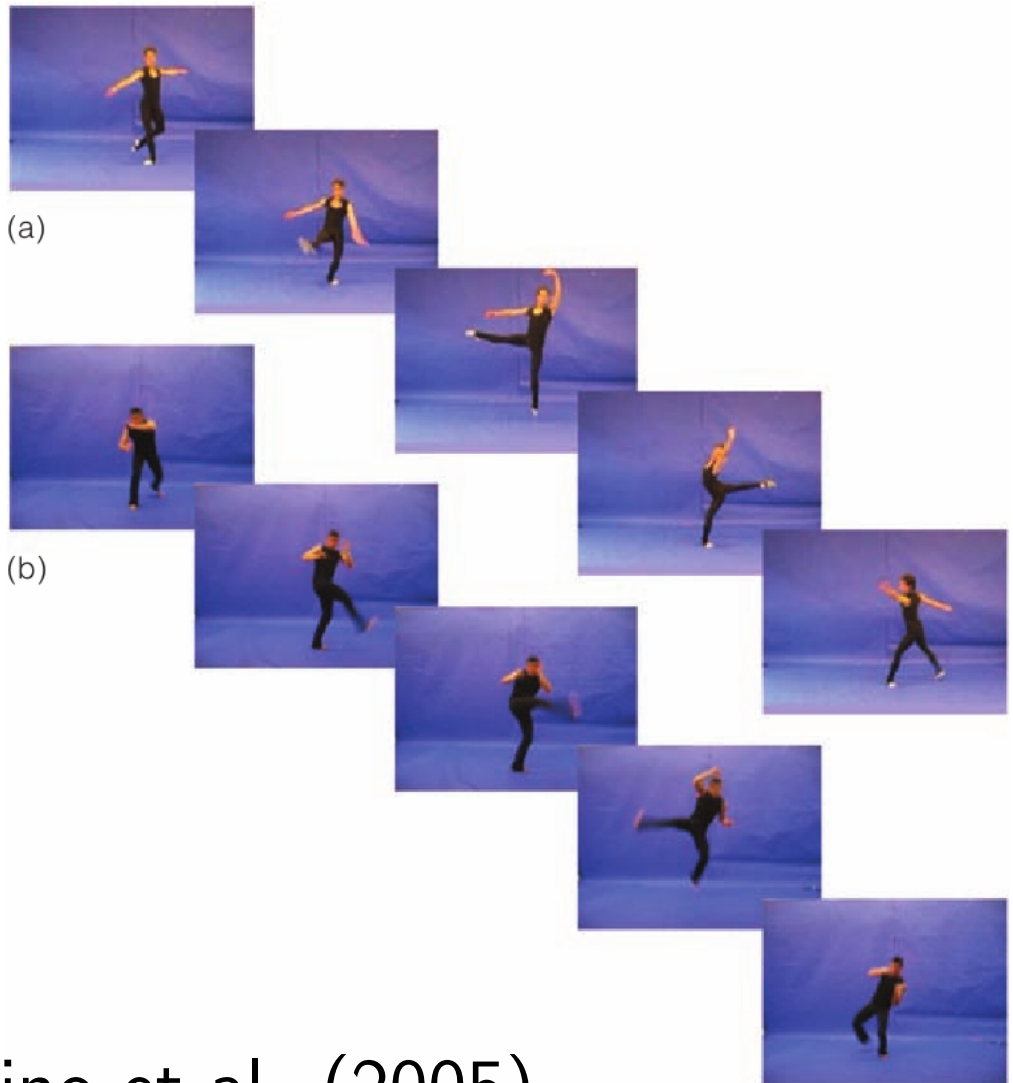


(a)



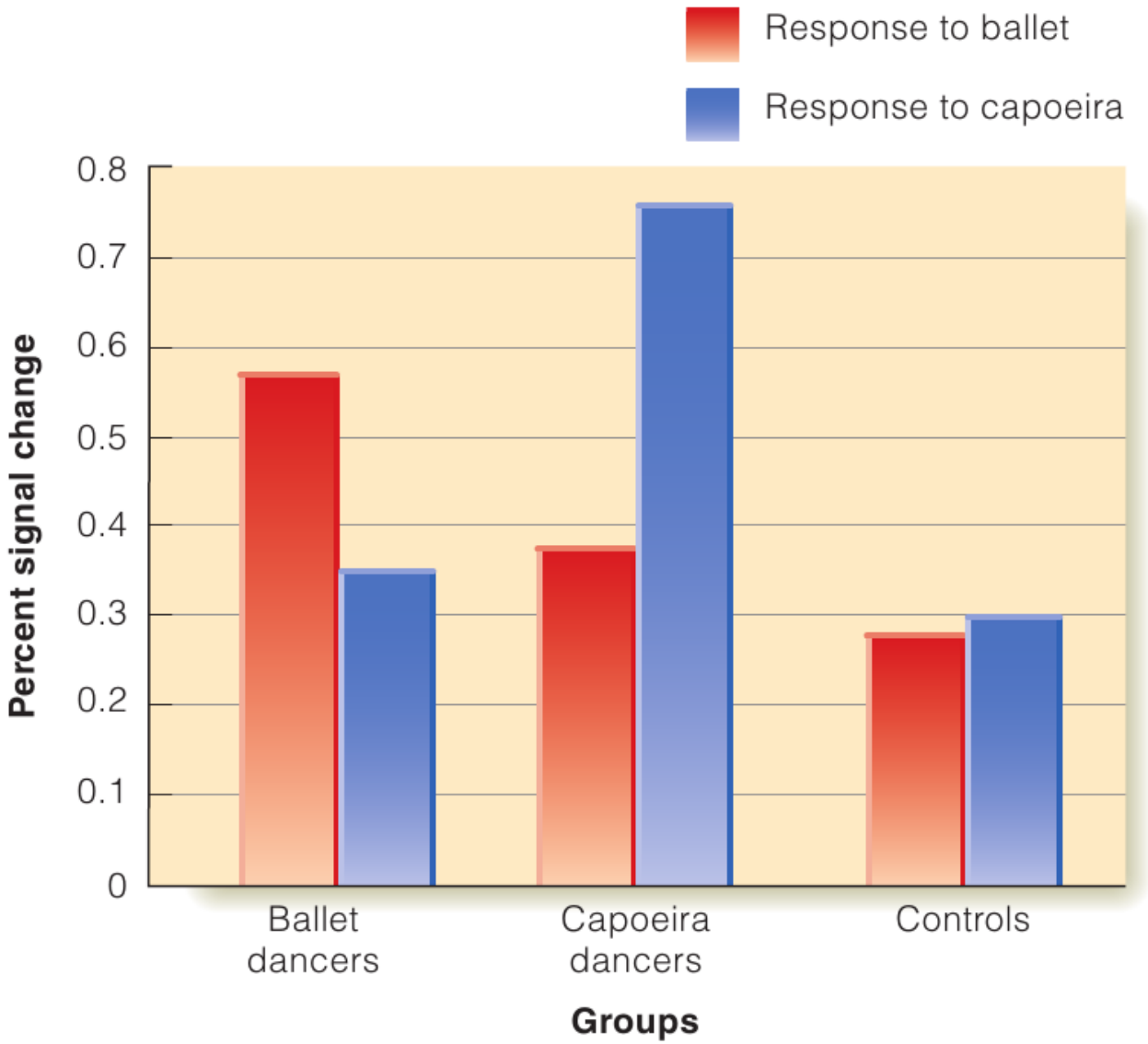
(b)

Mirror Neurons and Experience



Calvo-Merino et al. (2005)

Mirror Neurons and Experience



Controlling Movements With the Mind

http://youtu.be/77KsE--Adp8