# (Visual) Attention

- Perception and awareness of a visual object seems to involve attending to the object. Do we have to attend to an object to perceive it?

- Some tasks seem to proceed with little or no attention while others require concentration, effort. Why the difference? Does this imply that perception can occur without attention?

## Some Questions

Is attention required to perceive, recognize an object?

What happens to information (objects, events) that we do not attend to?

How much information can we attend to at once? How is this related to other capacities (other sensory systems, memory limits, action and motor control)?

How do we select what we will attend to? Is attention driven by our intent (top-down), driven by events in the world (bottom-up) or a mixture?

# Attention and Capacity

Many aspects of human cognition have a limited capacity.

Some of the limits are obvious. You can't see objects behind you because no light from them is reflected onto your retina (sensory limits). If you are turning a door knob with your right hand, you can not simultaneously be using that hand to wave to someone (motor limits).

The focus here is on processing limits.

# Some Basic Processing Limits

Working memory is your ability to hold onto items to mentally use them.  For example, you look up someone's phone number and then  "rehearse it" while you punch it into you phone.

What happens if someone interrupts you before you can dial the number by asking you a question?

The "forgetting" of the phone number is because working memory has a limited capacity for language materials. Similar limits exist for visual and spatial materials and for sound.

# Limits in Visual Perception

One limit in vision is our ability to resolve detail. It is good for items imaged on the fovea and poorer as we move away from the fovea. For example, in order to read (English), you move your eyes across the material, left-to-right, roughly from content word to content word. This focuses each word on the fovea.

If you are looking at a real-world scene, your gaze shifts from location to location. While you may be able to detect large objects in peripheral vision, details are not resolved.

# Eye Movements

When we read or inspect the visual world, our eyes move. These eye movements are a series of jumps (saccades) and fixations. In general, you move your eyes about 3 times per second (slightly faster during reading). The saccades take about 40 – 50 ms and the fixations are about 150 - 300 ms.

Special cameras can be used to record eye movements so we can monitor where people are looking.

# Selection

How do we select what to attend to?

1. Attention is “captured” by elements in the visual array. This is bottom-up selection.
2. Attention is directed by knowledge and learning to different areas of the scene. This is top-down selection.

## Bottom-up Selection

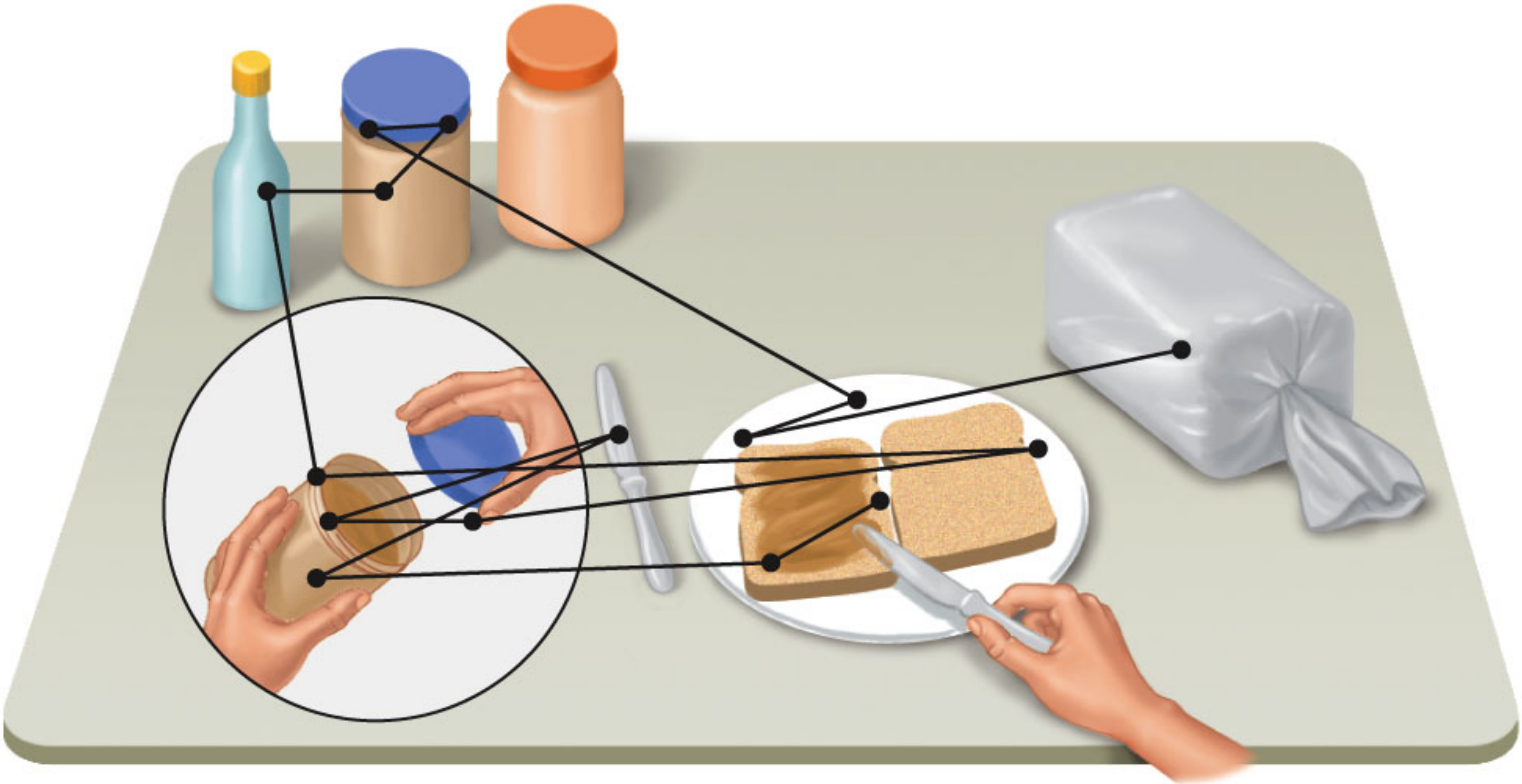
*Stimulus salience*. Parkhurst and colleagues developed the idea of a salience map. They start with features of the visual world that are coded by primary visual cortex: contrast, color, orientation. Using these qualities, they computed where changes occur in the visual image. (see Figure 6.5 in text for an example).

They also presented the same pictures to participants, one at a time and measured where the participants looked. Initial fixations on each picture were much more likely to fall onto the areas of change (high saliency).

# Top-down Selection

Prior knowledge can also alter where we look. If we are doing a visually guided motor task, we look at items as they are manipulated or needed to do the task.

The next panel shows the eye movements and fixations for an individual making a peanut butter sandwich. Note that they move from object to object as the person carries out the motor sequence to perform the task.



# Effects of Experience

Put an individual in a driving simulator. Stop signs can occur at intersections or in the middle of a block. The observers (driver) are more likely to detect the stop sign at an intersection (they often miss them in mid-block).

The stop signs in mid-block are often not fixated. Most at intersections are. The driver's experience is guiding their eye movements.

# Eye Movement Control Summary

Initial exposure to a scene may result in the first few eye fixations being driven by salience: bottom-up factors.

Experience and the intent of the observer exert strong control over eye movements, particularly in situations involving motor control or where there is detail that an observer intends to perceive (e.g. reading).

# Feature Integration Theory

Treisman proposed a theory of the role of attention in visual perception.

1. In a preattentive stage, features of the visual world are extracted. These are the primitives - the basic elements from which perception is built.

2. Using focused attention, features are combined.

3. The resulting "percept" is compared to memory. If a match is found, the object is recognized.

# The Features and a Prediction

The basic features are the attributes coded in the primary visual cortex, including color, contrast, orientation.

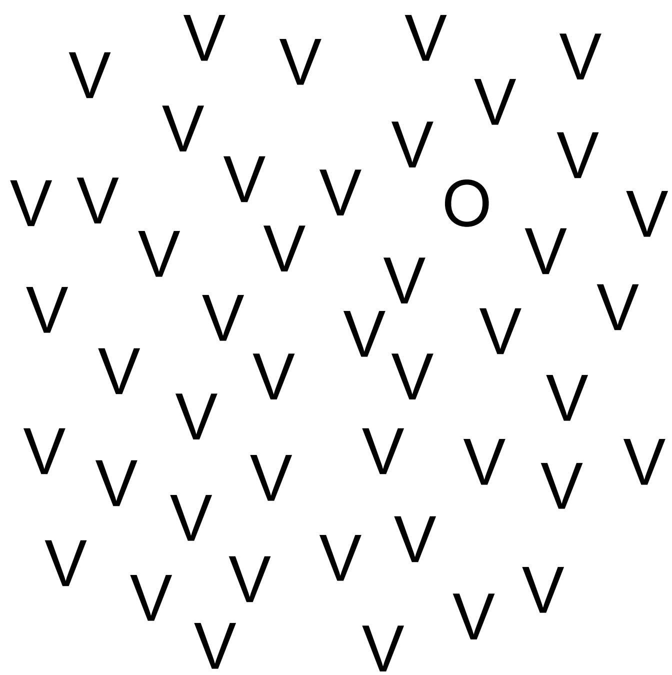
From the work on bottom-up attention, these should be what a person initially attends to.  So, if an object can be distinguished from the rest of a scene by only these attributes (without having to combine them), then a task should be easy.

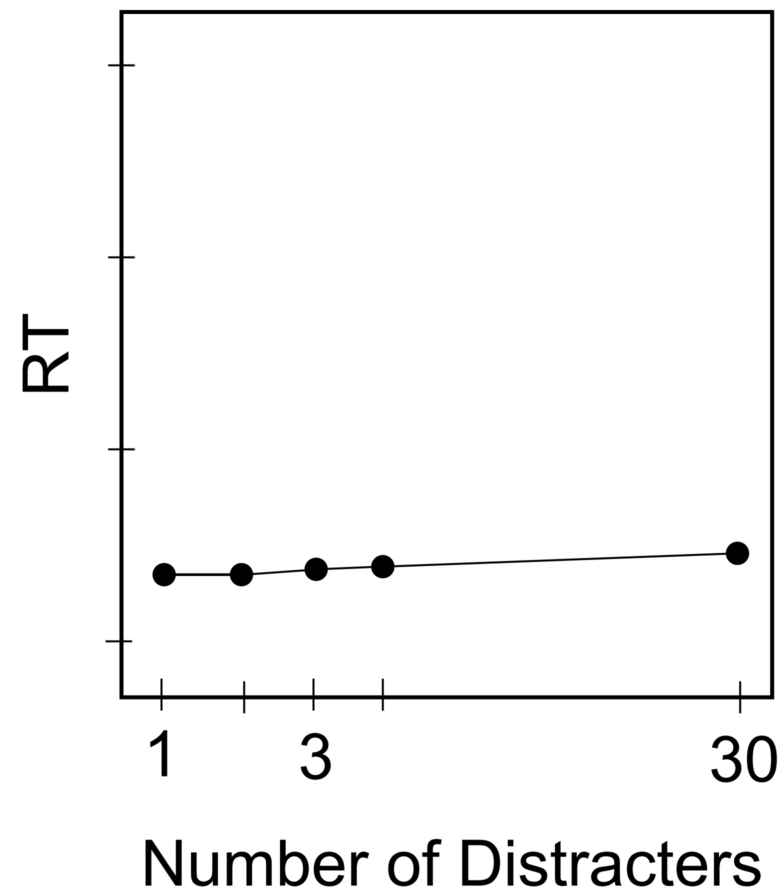
# Visual Search Task

Observers are given a target letter to look for. If it is present, they are to respond "yes" as fast as possible. If it is not, they are to respond "no" as fast as possible.

For example, they are asked to search for an O in a field of Vs. On different trials, the number of distracters (Vs) is varied. The position of the O is varied also (randomly).

# Visual Search - 2



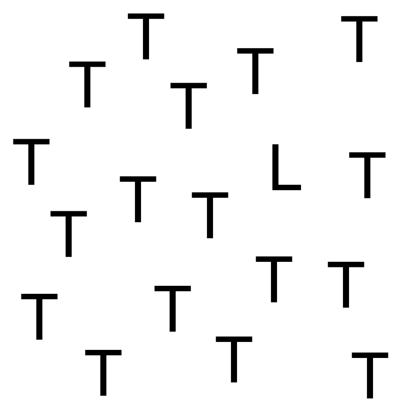


# Visual Search - 3

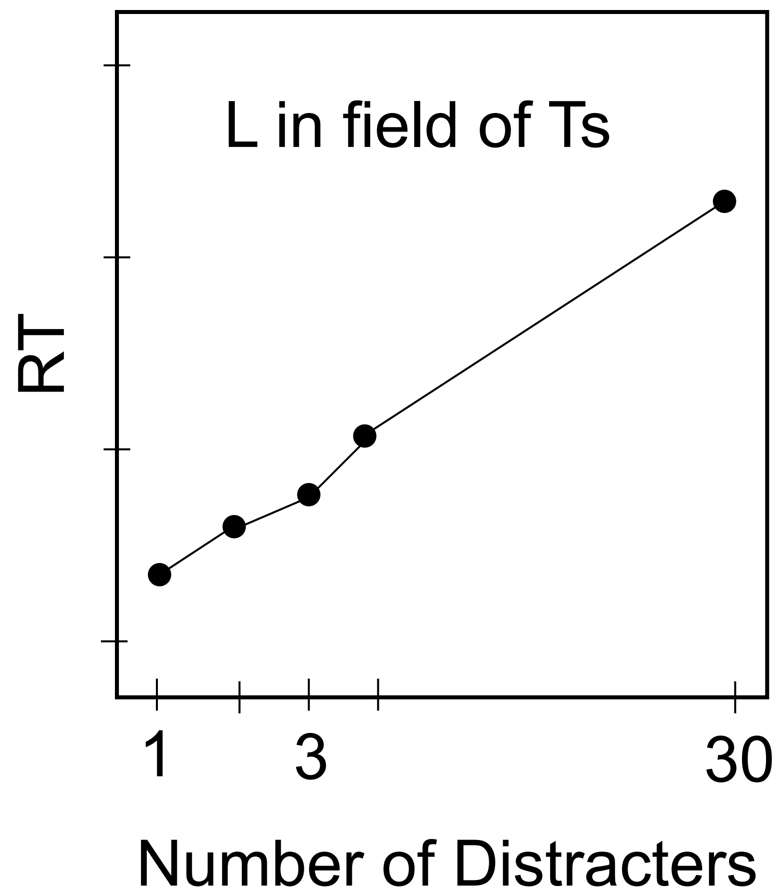
Searching for an O in a field of Vs is fast and easy.  The observer's speed is not affected by the number of distracters.  It is as if the target (O) "pops-out" of the field of distracters.

This is not the case for all searches.  For some types of objects, the more distracters, the slower the search.  The next panel shows searching for an L in a field of Ts.

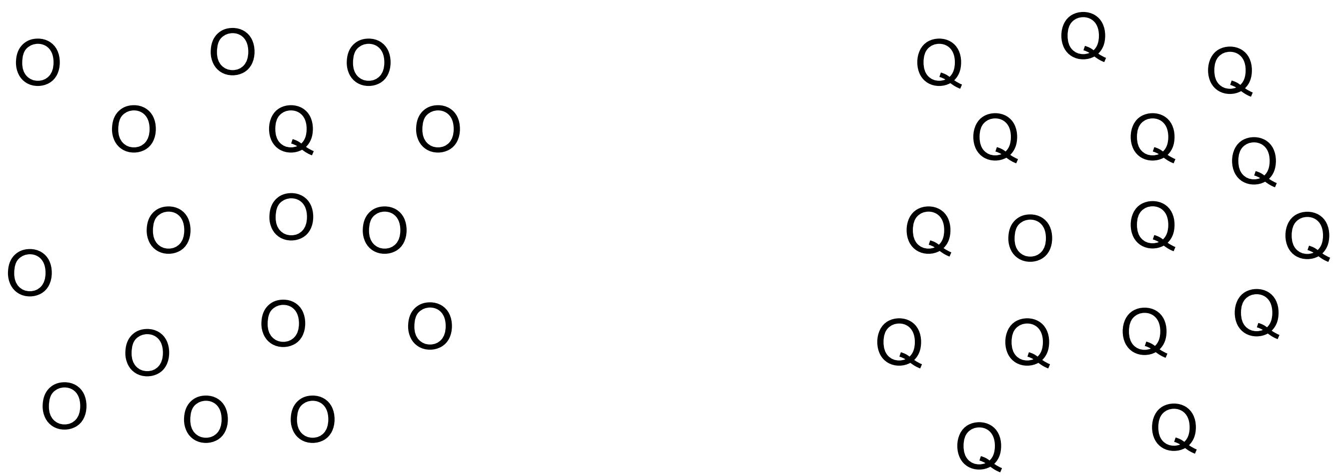
# Visual Search - 5



Here, as the number of distracters increases, the task becomes harder.



# Visual Search - 4




Searching for a Q in a field of Os is also easy. However, searching for an O in a field of Qs is hard, slow and RT increases with the number of distracters.

## The primitives

Based on observers' data, the basic features are:

1. Color
2. Orientation
3. Curvature
4. Contrast and Brightness
5. Closed areas
6. Motion

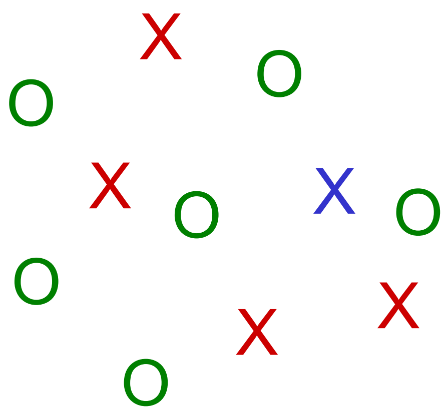
# Focused Attention

Following feature extraction, the visual system has a set of maps of feature values.  For example, it would have a color map, an orientation map, a location map, and a motion map.

The next stage uses focused attention to assemble features into objects.  The features at a particular location are combined.

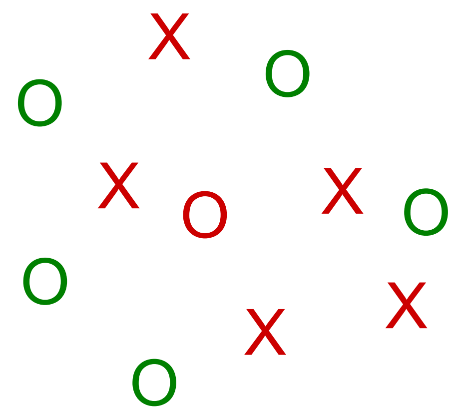
If this stage requires attention, then any target identification task that requires features to be combined to identify the target should be (more) difficult as the number of distractors increases.

# Focused Attention - 2




Find the blue X
- easy

When the target is only identifiable after combining features, the task is hard (slow and more errors are made).




Find the red O
- hard

# Focused Attention - 3

Why is focused attention necessary? Perception typically feels effortless.

Attention may be the glue that is used to bind features together to create our perceptual experience. Attending to one area seems to precludes attending to other areas at the same time (at least in Treisman's studies).

# The Results of NOT Attending

What happens to information that is not attended?

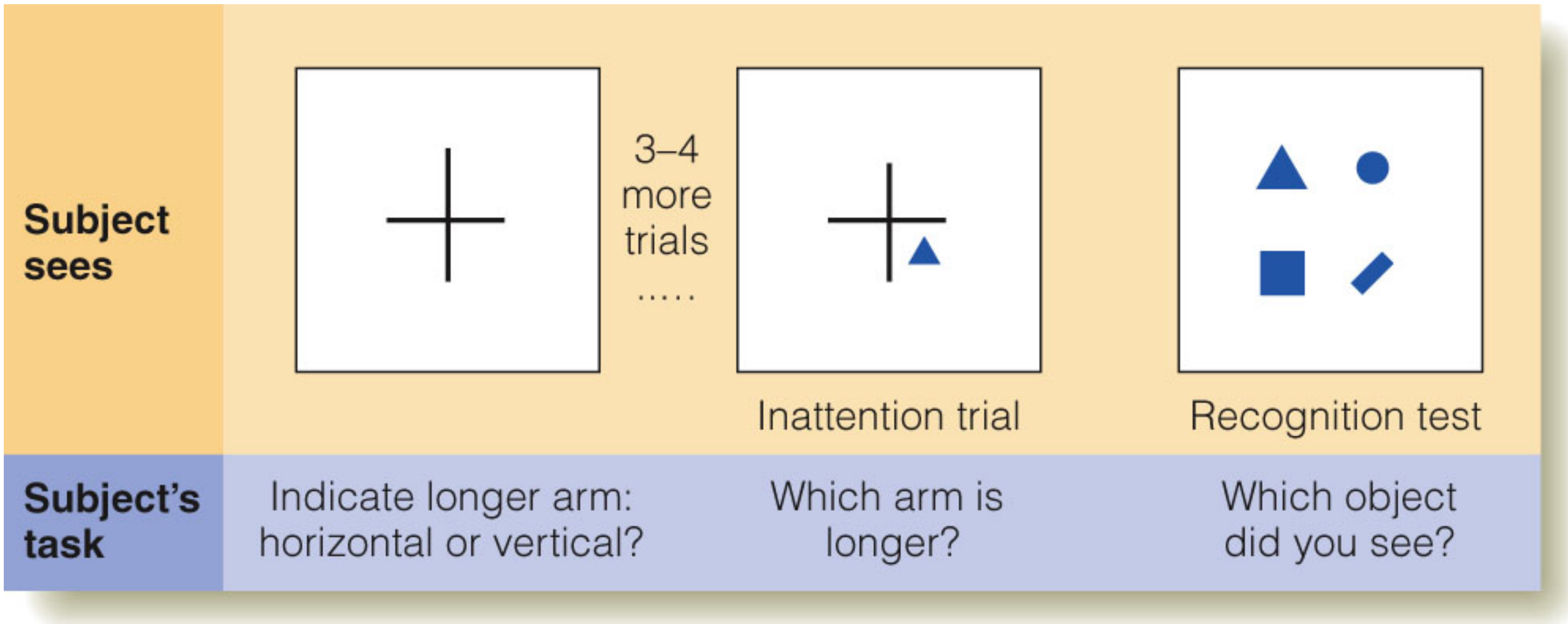
We'll look at two phenomena that suggest that without attention, information is not perceived:

1. Inattentional blindness.
2. Change detection.

# Inattentional Blindness

The observer is given a task. A cross is going to be projected briefly on the screen in front of them. They are to make a judgment about whether the vertical arm of the cross is longer or the horizontal arm is longer.

On some trials, a small geometric shape is also present, near the cross, in the display. On these trials, after making their response to the cross, observer's are given a forced choice for what geometric figure was also presented.


Subject sees
3–4 more trials .....
Inattention trial
Recognition test
Subject's task
Indicate longer arm: horizontal or vertical?
Which arm is longer?
Which object did you see?

# Inattentional Blindness

Observer performance on indicating what shape was present is no better than chance.

In another study, observers are asked to watch a video of a game. It is like a three–on–three basketball game. Observers are to count the number of passes of the ball. About 45 seconds into the game, a person in a gorilla suit walks thru the game (this takes about 5 seconds).

At the end of the video, when asked if anything unusual happened, about half of the observers say no.

# Change Detection

The observer sees a picture, then a blank field, then the picture is repeated, but with one of the objects from the original deleted. (see Figures 6.11 and 6.12 in text for examples). Observers typically fail to notice that an object has been deleted. The pictures need to be alternated back and forth multiple times for observers to notice the difference.

# Change Blindness

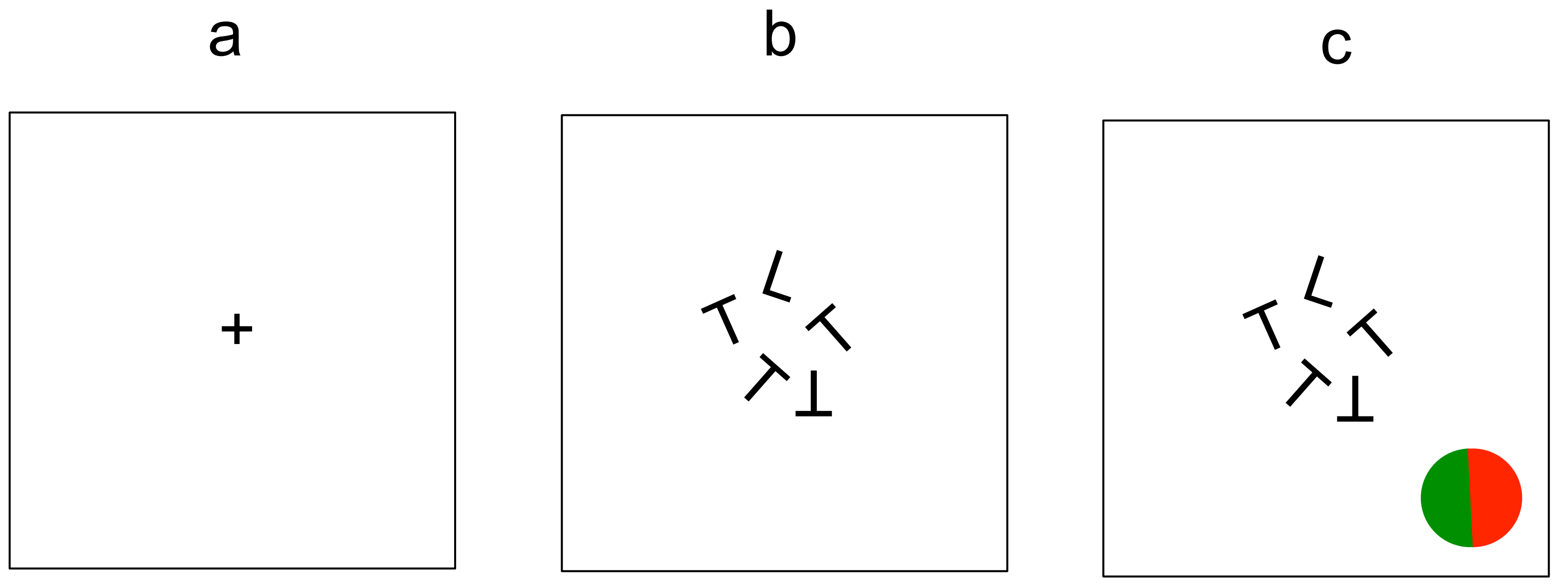
This is called *change blindness*. It can also be demonstrated with a video clip. In it, two people sitting at a table are talking. After a few seconds, the scarf of one of the people is removed. A couple seconds later, it is back on. Even when viewers are warned that changes may occur, the change is noticed on only about 25% of the trials.

If observers are asked in advance about whether they think that they will notice the change, most (over 80%) say that then will. Clearly, our intuitions do not match actual performance.

# Is Attention Required?

Inattentional blindness and change blindness seem to indicate that Treisman was right.  Attention is required to perceive an object.

Are there examples of perception without attention?




Observers are asked to fixate on the + in panel a. A group of letters is presented (b), then an object in one of the four corners of the display is presented (c). Finally, a mask replaces the object.

The object in panel c could be a face (male or female) or a circle (red on left, green on right or the reverse). The object is present for only 150 ms (followed by a masking display).

In the display of letters, either all of the letters are the same or one is different.

Observers are asked to do one of three tasks: Detect if the letters are the same, report if the face was male or female or the disk was red-green or green-red, or do both.

Central task performance. When asked to do the letter task (and ignore the object), observers are 80 – 90% correct.

Peripheral task. When asked to ignore the letters are do the object task, observers are 80-90% correct.

Dual task. When asked to do the letter task then the object, observers are good at the letters. Face recognition is also good, but the circle color performance is poor (near chance).

Observers *can not* do both the letter task and report the circle correctly.

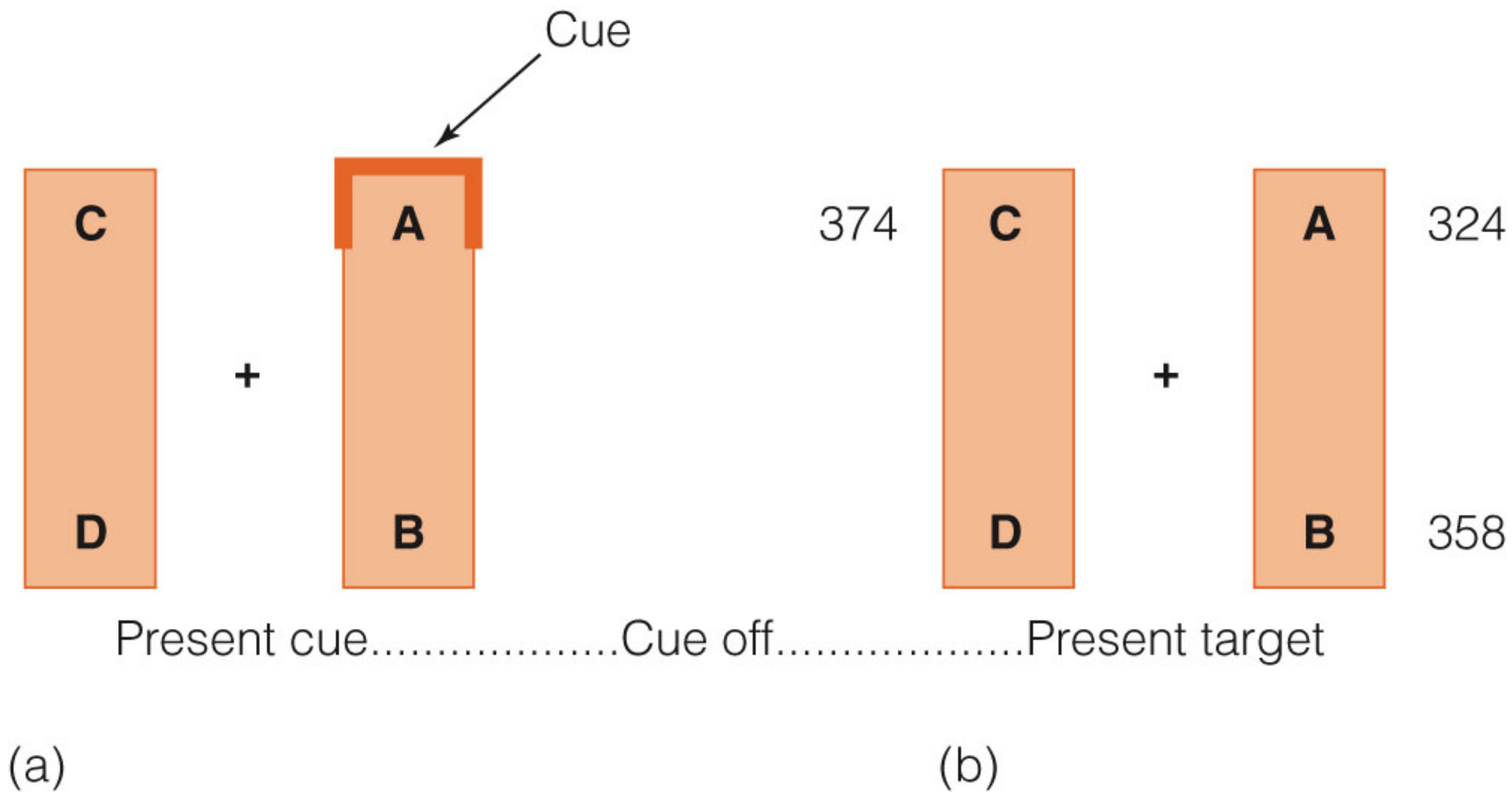
Observers can do the letter task and correctly report the face (male or female).

This does not imply that faces can be perceived without focused attention. Rather, it would appear that observers have the capability of handling face perception and letter recognition simultaneously (or that they do not demand the same attentional resources).

# Does Attention Make Perception "Better"

If we cue the observer to the likely position of a target, they are faster at identifying the target. So, speed of processing is influenced.

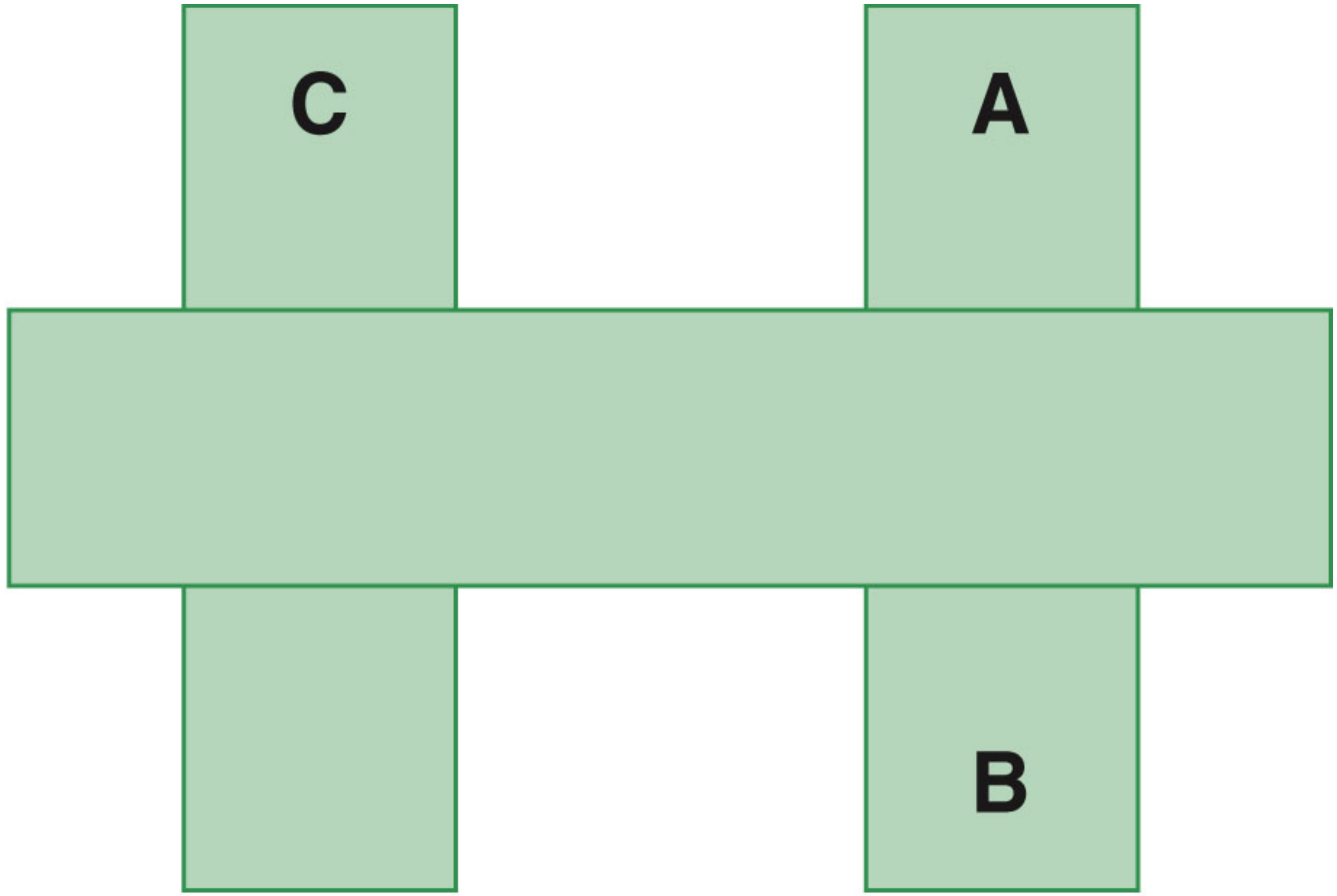
This influence is not "just" for the cued location.


Cue
C
A
D
B
+
374
324
358
Present cue..................Cue off..................Present target
(a)
(b)

Here, observer is faster to respond to target at cued position.  If target is in same “region” (location B), observer is not as fast as cued location, but faster than locations in non-cued region such as C.

Note that C and B are equally far from A.

The advantage for location B is found even if another region is overlaid over the display that obscures part of the cued region.




(a)

# Does Attention Make Perception "Better" - 2

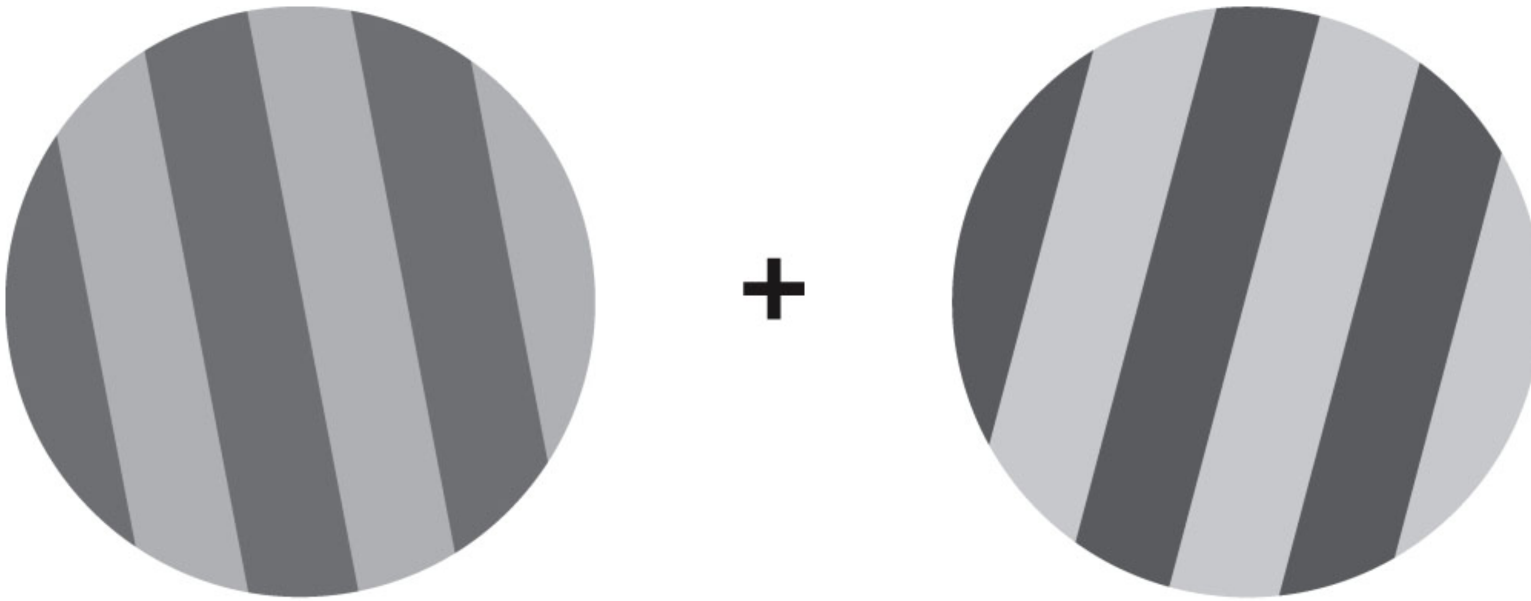
So, perception is faster with focused attention and the influence of attention extends to other regions that are grouped with the focused region. That is, the distribution of attention is influenced by the Gestalt Laws.

But, is perception "enhanced". Is it clearer, more vivid or otherwise different?

Carrasco et al. presented pairs of gratings (alternating light and dark regions) to observers.

Observers were asked to fixate on a center cross. Then, the grating pair was displayed (shown in next panel).

Their task was to report the orientation (pointing more left or right) of the grating with the higher contrast.



(b)

In this pair, observers report that the grating points right (grating on right has higher contrast).

The crucial trials had a small dot flash briefly on the left or right side of the fixation before the grating pair. This would draw the observers attention to that side (even though they continue to look at the cross).

When one of the two gratings was clearly higher in contrast, the presence of a flash had no influence on performance.

When the two gratings were the same, the one on the side that observer's attention was drawn to (by the flashed dot) was the one that observers reported the orientation for.

Since observers were instructed to report the orientation of the higher contrast grating, this implies that the grating that received focused attention was seen as having higher contrast.

So, it appears that focusing attention can enhance our perception of an object.

# Is Attention the Glue of Perception?

Returning to Treisman's Feature Integration Theory, the role of attention is to "glue" the disparate elements (features) together into an object.

Interestingly, there is some evidence for this both in experiments with observers and in a form of neurological damage.

# Illusory Conjunctions

A display is briefly flashed that contains red Os and blue Ts. The observer is asked to press a key, as fast as possible, if there is a blue O. Observers false alarm (report illusory conjunction) on 10 – 20% of the trials.

Treisman and colleagues have reported a number of results showing evidence for illusory conjunctions.

## Balint's Syndrome

In this disorder, there is damage to the parietal lobe (which underlies our "where" map of location information in perception). Feature Integration Theory indicates that individuals with this disorder should have difficulty in combining features because of the inability to focus attention at a location in visual space. Patients with Balint's Syndrome produce illusory conjunctions, even in situations where they are allowed to inspect a display for 10 seconds.

# Binding and Attention

Is attention part of the solution to the binding problem?

Maybe.  The synchrony of neural firing that was previously discussed is enhanced by attention. However, relatively little is known about the relation between synchrony and attention.  Further, it is not clear that synchrony is the solution to the binding problem.

## Attention and Neural Coding

We previously showed evidence that neurons in V1 in a monkey show enhanced firing when the object in their visual field is task relevant. This suggests that attention operates to increase neural response.

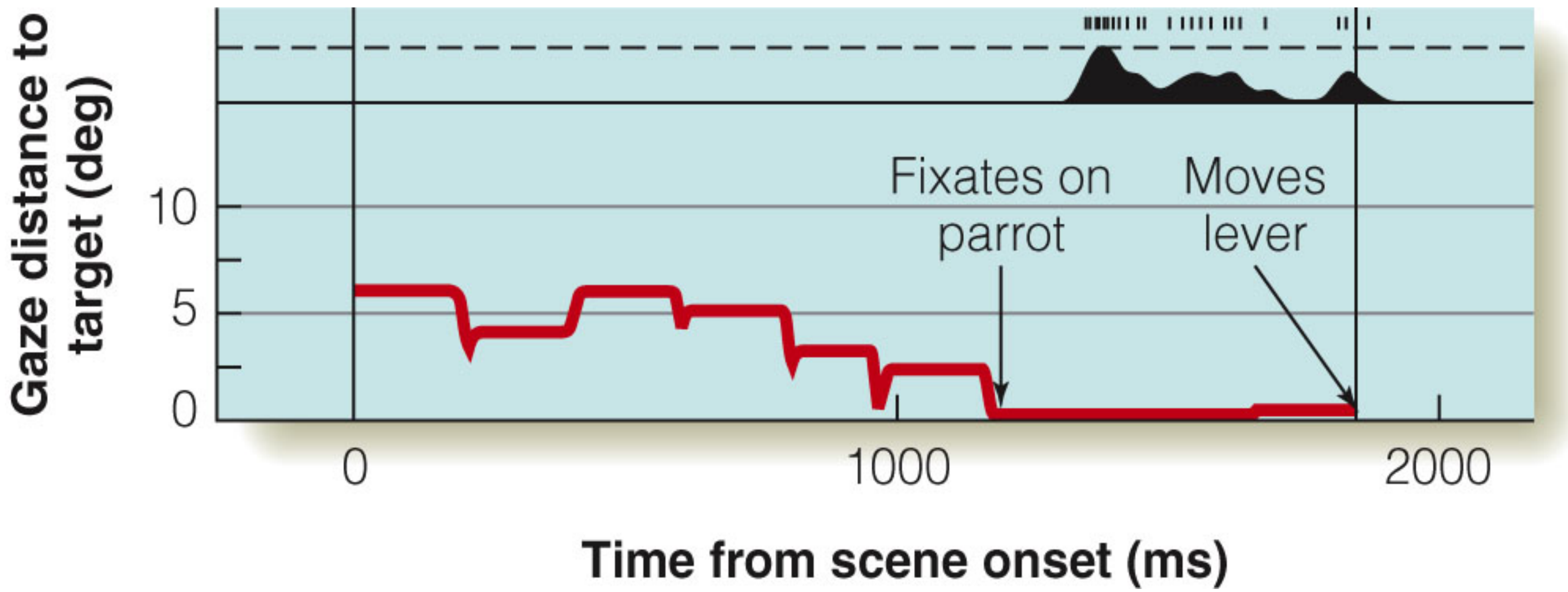
This influence is found in all visual areas. Further, the effect is larger in MT and V4 than in V1. Put another way, the further into the brain we go (higher in the visual system), the greater the influence of attention on neural firing.

# Attention and Neural Coding Example

A monkey is trained to recognize certain objects (such as a parrot) and pull a lever when it sees one (for a food reward).

An electrode in area IT shows that when the monkey sees a picture of a parrot (in isolation), the neuron fires.

Next, we give the monkey a complex scene with a parrot in it and record the monkeys eye movements and the response of the cell in area IT.




The data show that the neuron in area IT does not fire until the monkey fixates its gaze on the parrot. This implies that attention was necessary for this neuron to code the presence of the parrot.

# Neural Coding and Attention Summary

So, even though a visual image may contain an object, the coding of the object, including the firing of neurons that are sensitive to the object properties, may require attention. Perception is both the coding of the image on the retina by the early neural mechanisms and the relevance of the information to the task and situation of the observer.

# Beyond Visual Attention

Our focus has been the role of attention in visual perception.

Attention is also relevant to memory (basically, if something is not attended, it is not remembered).

There is also the issue of how one switches the focus of attention and just how much information can be attended at once (capacity).

## Beyond Visual Attention - 2

Finally, the limits of attention are influenced by experience.

A novice typist is easily distracted and can not do a second task while typing (without a loss in typing speed and/or accuracy).

A skilled typist can type material (from a handwritten copy) while simultaneously carrying on a conversation. Experience typing has reduced the attentional demands of the task.

# An Example of Failure of Selective Attention

Participants (you) will see two lists. In each list, the items will be in one of four ink colors: red, blue, green, or yellow. Your task is to name the ink color for each item in the list. You are to do the entire list and do it as fast as you can.

XXXX
XXXX
XXXX
XXXX
XXXX
XXXX
XXXX
XXXX
XXXX
XXXX
XXXX
XXXX

List 1. Rows of XXXXs in different ink colors.

RED
RED
GREEN
YELLOW
BLUE
YELLOW
GREEN
YELLOW
BLUE
GREEN
RED
BLUE

List 2. Rows of color words in different ink colors.

## Stroop Task Results

Participants take much longer to name the ink colors for the color words than for the XXXs. They also make more errors. The word (the name of a color) interferes with the naming of the ink color. This is termed Stroop interference.

Stroop interference is a *failure* of selective attention. You are trying to name the ink color, but the word that carries the color is being read. Reading is a highly practiced skill. Apparently, you can not attend only to the ink color without also reading the word.

## Stroop Interference Implications

If information is occurring where you have focused your attention and is a part of what you are attending to, it would seem that you can not "filter" it out.

Alternatively, you have sufficient attentional capacity to read each word at the same time that you recognize the ink color.  Because the capacity is available, it is used.

Both of these emphasize that the ability to focus attention and selectively process only certain information is limited.