Attention

Alex O. Holcombe

Updated on Thu Aug 19 14:13:18 2021

Contents

List of Figures

Preface

 Hi, I'm Alex. You can contact me (*he/him*) via email (alex.holcombe@sydney.edu.au) or find me on twitter.

This mini-textbook is to accompany my lectures on attention in the University of Sydney's PSYC2016 class. It will be updated before each lecture as I am continuing to add to some sections.

You won't be able to fully understand some of the points made from this alone; you must also watch the accompanying lectures.

You can read this mini-text here on the web, or as a PDF file, or as an e-book, which can be imported into your Kindle or other e-book reader. However, the web version is the only one you should rely on - some features (e.g. movies, some kinds of images) may be missing from the other formats.

Chapter 1

Background

Some aspects of attention are common-sense enough that they were known already at the very beginning of the scientific study of psychology. As William James put it in 1890,

> Every one knows what attention is. It is the taking possession by the mind, in a 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 with others

Some aspects of attention are embedded in everyday speech:

- Please "pay attention!"
  - Implies one has to choose something ("attend" to it) to fully process it.
- "Sorry, I wasn't listening"
  - There's a difference between hearing something and listening to it. Listening means we are attending to it. If we do not pay attention, we are less likely to retain something someone says, and possibly we won't even comprehend what they said.
- "Sorry, I missed that."
  - People sometimes say this after someone else says something and the first person realizes that they didn't understand what was said. In particular, people say it when they don't think the problem was that the statement was not loud enough for them to hear. Instead, the problem is often something with attention.

- "I didn't notice that." What might people mean differently when they say that rather than choosing to instead say "I didn't see that."?

Failures to notice things explain a large number of accidents. But **why** do we have to pay attention to comprehend or retain some information? That is the subject of the next two chapters 3.

If you took PSYC1 at the University of Sydney, you already heard Caleb Owens' "Cognitive Processes 2" lectures on visual attention, which were related. They had the following learning outcomes:

- Understand and be able to give examples of situations where focused visual or auditory attention leads to limited processing of other stimuli.
- Be able to define, distinguish and give examples of focused attention, divided attention, diffused attention, inattentional blindness, and change blindness.
- Understand and be able to describe the difference between an early, late or flexible locus of selection.
- Be able to both give and interpret examples which demonstrate an early, late or flexible locus of selection.
- Be able to define and distinguish between endogenous and exogenous attention
- Understand the role of attention according to Treisman's FIT (feature integration theory) and the visual search evidence (for both feature and conjunction targets) which supports it.

You are not responsible for these! This year has its own list of learning outcomes (2).

We will both review material from Caleb's lecture and bring new ideas to bear, as well as add more material. In the parts that review what Caleb already taught, I will go quickly, so you may want to review his material from last year - I have put his slides on Canvas.

Chapter 2

Learning outcomes

You should be able to:

1. Explain why we need attentional selection.
2. Explain what limited capacity means.
3. Explain the reasons why people don't notice changes in change blindness videos.
4. Know the three kinds of attentional selection.
5. Describe how the kinds of selection connect to visual search performance and Treisman's FIT (feature integration theory).
6. Know two bottlenecks in human information processing and describe the roles they play in multiple phenomena.
7. Define, and distinguish between, top-down (endogenous) and bottom-up (exogenous) attention.
8. Explain why people miss important information in magic tricks and in the "inattentional blindness" videos.
9. Explain the role of bottom-up attention, top-down attention, expectation, task, and limited capacity in each of the phenomena discussed.
10. Understand how sensory signals can distract and the negative effects on human performance

These are high-level outcomes. You'll see more detailed points in each chapter, often with questions at the end to check your understanding of individual points.

Chapter 3

Bottlenecks

3.1 Computers

In conventional computers, such as a laptop or desktop, most of the calculations are done in the CPU - the central processing unit. Computers also have memory, called RAM. For example, the laptop that I am writing this text on had 16 GB of RAM. That is enough to store a lot of informationm- 2857 copies of the complete works of William Shakepeare.

Unfortunately, the CPU of a conventional computer, such as my laptop, can only operate on a tiny amount of memory at any one time. Thus, there is a bottleneck between the memory and the CPU. Instead of the CPU doing calculations and transformations of all the memory simultaneously, it only does so bit-by-bit.

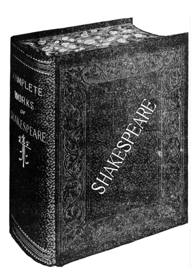


Figure 3.1: The complete works of William Shakespeare

Imagine you had the complete works of William Shakespeare in your computer's RAM (just one copy), and wanted to change each lower-case letter in it to upper-case, and each upper-case letter to lower-case. It's a pretty simple task, but the CPU can only do it for one letter at a time. We want to get all of Shakespeare's works processed by the CPU, but the CPU is a bottleneck - the data has to sit in memory waiting for the CPU to get to it.

```
## PhantomJS not found. You can install it with webshot::install_phantomjs(). If it is
```

There is a bottleneck between memory and the CPU. This is inherent to the architecture of conventional computers

3.2 Brains

After the image of an object forms on your retinas, or sounds hit your eardrum, neurons carry information about these signals to your brain.

Some of that information you will then perceive. For example, you might perceive that there is a salt shaker in front of you. Some of the information you will also remember and be able to recall a second or so later. But you will not consciously perceive all of the information that reaches your brain, and you will remember even less. In other words, we have **bottlenecks in perception and memory**. Only a limited amount of the sensory signals coming in from the eye are fully processed for memory or even for perception. *Attention* refers to the control of which signals are processed the most. That is, if you attend to an object, your brain is trying to ensure that it is fully processed.

For perception, some signals from the senses won't be perceived unless the associated sensory signals are routed to processes in the brain that can only process a few things at a time. Memory also has a limitation, and not just in how much storage is available. There is also a bottleneck for getting things encoded into memory. That is, many things won't be remembered unless you attend to them so that the associated perceptual representations are routed into the memory processes that, sadly, can only process a few things at a time. We don't think the situation in the brain is as simple as a single bottleneck like a CPU. Instead, the different things the brain does are subject to different kinds of bottlenecks.

3.3 Feeling a bottleneck

How many math problems can you do at one time? Although your brain contains more than 80 billion neurons, you can probably only do math problems one at a time. How about other tasks - if I showed you twenty words on a page, how many could you read simultaneously? A large body of psychological research suggests that humans can read only a few words at a time, or possibly only one at a time (White et al., 2018). Yet when light hits the back of our eyes, it is greeted by six million cones arrayed across each retina. Each bit of the image is

Figure 3.2: Sensory information entering the brain eventually hits bottlenecks!

simultaneously processed, as each has its own photoreceptors devoted just to it. Retinal stages, then, process millions of small regions of the retina at the same time.

Explicit thinking like math problems, however, involve a series of steps, and our brains appear to be particularly capacity-limited at doing this.

But given that explicit thought is so limited, seemingly to only one thing at a time, something has to control what it thinks about - attentional selection. The earlier lectures in this unit were about memory, decision making, and problem solving. Attention is what routes information to the memory, decision-making, and problem-solving processes.

That memory encoding and problem solving are so limited is quite intuitive - we've all had the experience to know that it's hard to think about solving more than one problem at a time. A bit more surprising, perhaps, are the limitations on perceptual tasks. But it's important to realise how pervasive bottlenecks are in the mind - they afflict even perception.

Imagine you were to travel back in time several decades ago as a fighter pilot, patrolling the skies of Cuba to deter the expected American invasion back in the 1960s (or the 70s or the 80s, the Americans were always a big threat!). You'd be flying a Soviet-made MIG fighter. When I visited Havana in 2003, I got to see one of these planes up close.

Figure 3.3: A MIG-21R fighter plane

From the top of the stairs they've constructed next to the plane, I was able to take a photo of the inside of the cockpit: 

As you can see, the pilot has a *lot* of dials to monitor. Information overload? Yes, although the large number of dials and switches don't overload the *initial* stages of our brain's visual processing, some stage after that are overloaded.

That later stage, where more extensive processing is done, have a bottleneck. We will describe this in Chapters 7 and 9.

In World War II, dials on planes proliferated and it became obvious that fighter pilots couldn't fly and effectively monitor all the displays at the same time. This inspired psychologists to begin studying capacity limits .

There is a wide array of information a pilot might potentially need to be aware of, more than the pilot could really process simultaneously. Because a pilot's cognitive system can only process a few things at a time, only some signals can be processed at any one time, and this comes at the expense of others.

Many people have the naive view that if our eyes are open, we are aware of everything that hits our retinas. But as will be illustrated in later chapters, this is not the case, due to the bottlenecks on certain aspects of processing in the brain.

Forgetting about the brain for the moment, let's consider your eyes. They are a lot like an old analog camera. They don't know what they are looking at. You need a brain, with millions of neurons working together, to process the eye's signals and identify an object. And our brain isn't big enough to have enough neurons to identify all the objects currently in your field of view. If you had enough neurons to do all that, your head would have to be much bigger, like that of this cartoon man's:

To save head space, and energy, the brain has only a small number of neural circuits capable of doing math problems, of identifying faces, of tracing shapes, of reading, and of doing other tasks. This raises the problem of *attentional selection.*

The problem of attentional selection is that of getting the appropriate signals from the eye to the limited number of neurons available for each task. If you didn't have attention, there would be no control of what gets fully processed.

Regardless of whether you're paying attention or what you're thinking about, the photoreceptors in your retina are transducing the light emanating from me, the ganglion cells are sending spikes representing my image to the lateral geniculate nucleus, and the LGN is passing that information into the cortex. But, unlike these processes, not all visual processing is mandatory- some is optional. For example, your brain probably doesn't start computing the arm trajectory to grab this pen until you willfully think about grabbing it. When you attend to the pen to think about grabbing it, we say that your attention selects the pen for further processing by the motor system.

Put in computer information terms, Anderson et al. (2005) depicted the situation with the following diagram:

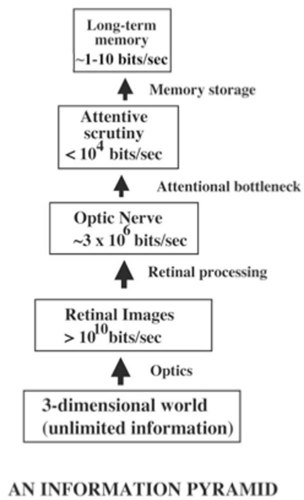


Fig. 1. An information pyramid for the visual system. The estimates of information available in retinal images, encoded in the optic nerves, and passed through the window of attention have been discussed by Van Essen et al. (1991) and Van Essen & Anderson (1995) and are based on information rates of 3 bits/s for each neuron (Eliasmith & Anderson, 2003). Human memory storage rates are based on estimates by Landauer (1986).

Their tentative, rough estimate was that the optic nerve can process 100 times as much information as can get through the attentional bottleneck. In other words, only 1% of visual signals get past the bottleneck.

The same problem occurs for other senses such as hearing. As you've probably experienced at parties, you can't comprehend what everyone around you is saying at the same time. You need to *listen* to one particular conversation.

Figure 3.4: A New Year's Eve party in Barcelona, Spain

A major question of attention research is what kind of processing goes on before you attend to something. Caleb told you about this in PSYC1 when he discussed the debate about the early, late or flexible locus of selection. In this class, we will only discuss this with respect to visual search. For what we will focus on, remind yourself by looking back at the learning outcomes (2) and the chapter titles.

3.4 A bottleneck for object judgments

3.4.1 Simultaneous objects

Duncan (1984) was interested in how many objects we can process at a time.

He asked people to make two judgments about a display that was flashed briefly. His participants could be asked to judge

1. whether a box was small or large
2. whether a gap in the box was on the left or the right side of the box
3. which way a line was tilted
4. whether the line was dotted or dashed

Here are two of his displays:

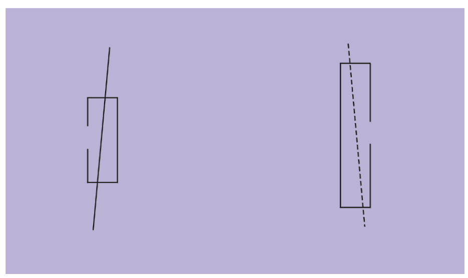


Figure 3.5: On the left, a clockwise-tilted line on top of a box with a gap on the left.

An individual trial in the experiment consisted of a brief presentation of a box with a line passing through it, like so:

This was considered to be two objects, a box (with a gap) and a line. The participant knew they would have to make two of the judgments listed above, such as #1 and #2 or #1 and #3. Those two judgments could be either about one object (the box, or the line) or two objects (the box and the line).

- Judgments about one object: #1 & #2 or #3 & #4
- Judgments about two objects: #1 & #3, #1 & #4, #2 & #3, or #2 & #4

Duncan knew that if he presented the objects for long, everybody would get the judgments right regardless. But he found that if he flashed the objects for

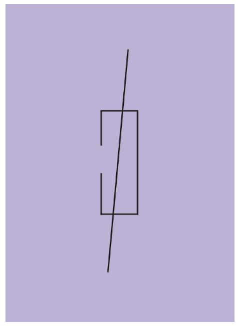


Figure 3.6: On the left, a clockwise-tilted line on top of a box with a gap on the left.

a tenth of a second or less, the task became difficult and participants made frequent errors.

The results of the study were that when the two judgments the participant had to make were about a single object (the box or the line), they made an error about 17% of the time. But when the two judgments were about different objects (one judgment about the box and another judgment about the line), participants made an error about 24% of the time. That's about 50% more errors for two judgments about different objects than two judgments about one object.

This pattern of results suggests there is a *bottleneck* for processing objects. The area the participants had to attend to was approximately the same in the two-object condition and the one-object condition. So the greater difficulty associated with the two-object condition was not down to having to split or spread attention over a greater area. In the conclusion of his paper, Duncan wrote:

> Findings support a view in which parallel, preattentive processes serve to segment the field into separate objects, followed by a process of focal attention that deals with only one object at a time.

In other words, Duncan's proposal was that the early visual system processes the multiple objects of a scene simultaneously, but making certain judgments requires additional processing that proceeds serially.

3.4.2 Successive objects

Subsequent work has yielded evidence that processing of even a lone object is subject to a bottleneck that prevents perception of a second object when a first object is being encoded into memory. The phenomenon that demonstrates it is called the attentional blink. Here is one demonstration and here is another. People do fine identifying a single item from the stream, but not two, if the second one occurs very soon after the first one. In some cases, the performance decrement is quite a bit bigger than that found in Duncan's experiment for two objects.

The relationship of the **memory encoding bottleneck** evident in the attentional blink task to that of the difficulty judging two objects presented simultaneously is still being investigated. We don't really know whether we should think of these as two distinct bottlenecks or whether the same underlying limitation in the brain is causing both.

3.5 Exercises

Answer these questions and relate them to the learning outcomes (2):

- What did the discussion of the CPU of a computer illustrate about the brain?
- What problem does the cockpit of a fighter plane present for a pilot?
- Why do we need attentional selection?
- Write out brief answers to the first two learning outcomes (2).

Chapter 4

Overt and covert attention

Although our brains have a bottleneck problem, our ancestors hundreds of millions of years ago didn't! When our senses first started to evolve in very simple animals, they didn't have the high processing capacity they have today. For instance, the first organism to have vision may have had only one photoreceptor. So, later processing stages were *not* overwhelmed!

However, primitive animals still had one kind of attention. For an object in the environment to be processed, an animal had to point its photoreceptor at it so the light from that object would stimulate its photoreceptor.

When an animal moves its sensory apparatus to better process something, we call that **overt attention**.

4.1 Two reasons for overt attention

Animals like us have millions of photoreceptors, but we still have a limited field of view. Therefore, we move our head and eyes around a lot in order to get objects of interest in our field of view. Again, this is called *overt* attention.

But getting the object in our field of view is not the *only* reason we move our eyes. Just as often, we move our eyes because we need to look directly at something to see it well.

Peripheral vision, away from the center of where you're looking, is low resolution. Many people either don't know this or don't fully appreciate just how bad peripheral vision is. That's because people move their eyes to an object as soon as they're interested in it, so they never notice that their vision of it wasn't good to begin with.

4.2 Peripheral vision

To get some feel for the poor quality of peripheral vision, try reading the text in the red and green squares when you have your eyes on the black dot in the center. Researchers refer to keeping one's eyes fixed on a point *fixating*. So, we call the black dot the *fixation point*.

Peripheral vision has poor spatial resolution, which is why when one is looking directly at the black dot, one can't read the text in the colored boxes

When you fixate on the fixation point, you can't read the words because outside of the centre of vision, the signals passed onto your brain are low resolution, like an image with not enough pixels.

So, **overt attention** is not just about moving your eyes and head so that you can see it, it's also about making sure there's a chance to perceive the object very well, by pointing your eyes directly at it, or at least very near it.

4.3 Covert attention

Have you ever heard the phrase "covert action"? The American CIA (Central Intelligence Agency) describes itself as an agency that "collects and analyzes foreign intelligence and conducts covert action." What they mean is that they conduct activities in secret, such as operations in other countries without other countries knowing what they are doing.

Similarly, **covert attention** refers to attention that can occur without other people knowing what you are attending to. The reason this is possible is because unlike some simpler animals, we can use our minds to choose which part of the sensory world to process more without making any physical movements.

Try staring at the black dot below, and attending to different letters without moving your eyes.

Unlike in the previous display, in this display the letters are probably big enough for you to read even when you are fixating (staring at) the black dot at the center.

You're able to choose to concentrate on an individual letter, like 'H'. Or you can concentrate on two letters, like the 'E' and the 'D'. Try it!

Scientists refer to this as *attending* to those letters. When you do it, your mind processes those letters more than the other letters. This is the phenomenon of *covert* attention.

A previous chapter (3) explained that we have bottlenecks in the brain. It's because of them that we need *covert* attention. When there's a lot of stimulation on our retinas, if there's anything important, we need to attend to them to be sure that those get through the bottleneck into higher-level processes.

4.4 Exercises

Answer these questions and relate them to the learning outcomes (2):

- Explain the difference between covert attention and overt attention.
- What was the first display of this chapter used to illustrate?
- What are two reasons for overt attention?
- What was the second display of this chapter used to illustrate?

Chapter 5

Bottom-up and top-down attention

In the previous chapter we learned that at any one time the sensory signals from only a few objects are being *fully* processed through the bottleneck(s) and thus are likely to enter memory.

Given the existence of the bottlenecks, we really need ways to prioritise what gets selected for high-level processing. How does your brain decide which objects to attentionally select?

Well, where and when attentional selection happens reflects a combination of factors. In the case of this text, you must have decided to read it. That is, your attentional selection of this text occurred because you gave yourself a task of reading it. While your brain is only able to read one, or at most two, words at a time, your eyes and attention hop along to select and fully process the successive words in this line of text.

We call this kind of selection *top-down* attention.

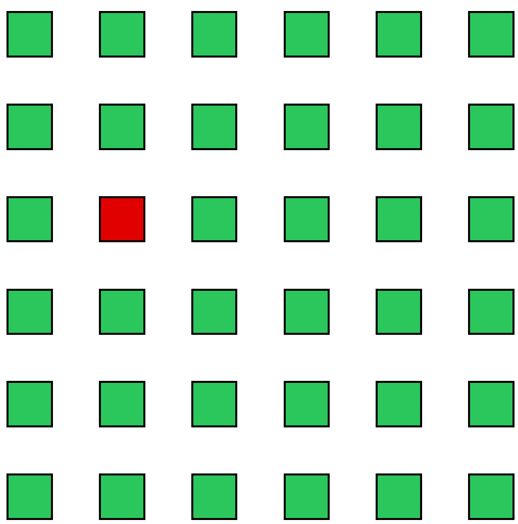
Top-down attention is typically voluntary, and thus guided by your expectations and desires, as represented by this inspector intentionally scrutinising individual bits of a crime scene.

Bottom-up attention is quite different - it's when something in the world grabs your attention. This can sometimes happen even against your will when you are trying to concentrate on something else. The reason for bottom-up attention is a bit like why you have CTRL-C, ESC, or "Force quit" on your computer. After you give your computer a task, sometimes you need to interrupt it. Indeed, every responsive system needs interrupt signals to take them off task when something that might be even more important crops up. That is, no matter how strongly a person is concentrating on a task, there should always be a possibility for unexpected information to trigger attention so that the person remains responsive to unexpected dangers.

If you hear a sudden loud sound, your attention is likely to be taken off, at least momentarily, the task you are performing. This was useful throughout evolutionary history to ensure that our ancestors evaluated sudden movements or sounds that might mark the arrival of another animal such as a predator. Similarly, if someone taps on your shoulder, or another body part, that's pretty likely to get your attention. We have evolved to be quite vigilant regarding possible threats to our body.

The art of concentration, and studying well, is in part knowledge of what distracts one's attention, and placing oneself in situations where your attention won't be distracted.

Unique visual objects in a scene also elicit bottom-up attention. For example, look at the image below - does something in it attract your attention?



If you aren't colorblind, the object with the unique color should have attracted your attention. This is an example of **bottom-up attention**. When they buy a car, some people deliberately pick an unusual color because they know that when they go shopping, if they forget where they parked their car, they will have little trouble finding it. A pink car will "stick out" conspicuously even in a sea of other cars, if those cars have the more typical colors of black, white, grey, and dark colors.

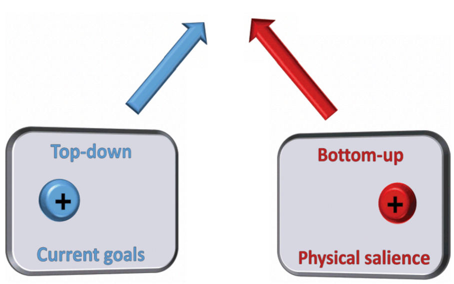
An object with unique motion direction can also summon attention, as you can see here.

However, not all unique objects in a scene will attract attention. In the below image, the animal whose back you see in the foreground is an elk. Can you find the mountain lion that is stalking it?

It's extremely dfficult to find and see for our limited brains. Click here to see the lion circled - you might still need to zoom in to see it! Mountain lions and other animals have evolved to have an appearance, and engage in behaviors, that won't attract the attention of other animals. The next few chapters will be, in part, about what does and doesn't attract attention.

5.1 Bottom-up attention and top-down attention, together forever

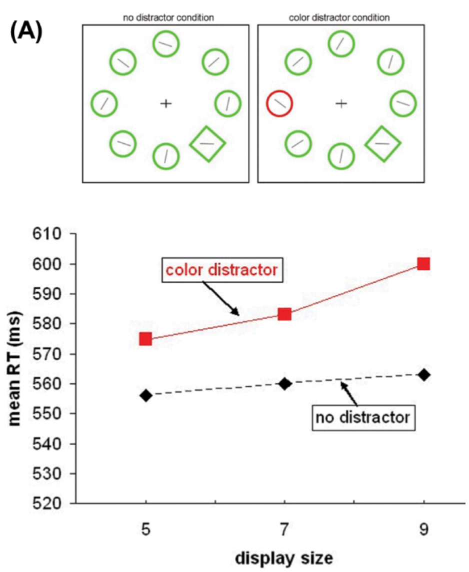
The signals of bottom-up and top-down attention must be somehow combined to determine where your attention ends up going.



As described in the previous section, top-down attention reflects one's current goals and task. Bottom-up attention reflects things in a scene that might grab our attention during almost any task, like a unique color. We don't fully understand how these work together. Sometimes top-down and bottom-up factors compete with each other. This can be seen in the results of an experiment described by Theeuwes (2010).

In the experiment, participants searched for a green diamond presented among a variable number of circles and had to respond to the orientation (horizontal or vertical) of the line segment presented within the diamond shape. So, their task was to find the diamond and pay attention to only it.

In some of the displays, Theeuwes (2010) included a circle that was different in color from all of the rest of the items on the display. The uniqueness of this color tended to attract attention. Because that meant attention was attracted away from the diamond, the results was an elevation in response time for reporting the orientation of the line segment in the diamond.



The graph of the results above shows the average time to indicate what orientation was in the diamond. The horizontal axis shows that the more stimuli that are presented, the longer it takes people to respond. This suggests that the more objects there were, the longer it took to find the diamond. Also notice that the red line is above the black line. This was the most important result - trials with a uniquely-colored distractor slowed response time.

- By how much, approximately, were responses slowed?

This slowing is sometimes called *attentional capture*. Objects or signals that grab, or capture, bottom-up attention are sometimes called *salient distractors* or *exogenous cues*. Due to the changing influences of bottom-up and top-down attention, attention may rapidly shift among different stimuli, depending on a combination of task factors and salience of the items. What we attend to, then, reflects a combination of what is important for our task and extraneous attention-capturing signals.

Some researchers think that top-down and bottom-up attention combine at a “priority map” mediated by a distributed network involving frontal, partial, temporal areas. Top-down signals largely reflect frontal and parietal areas, while bottom-up attention reflects sensory brain areas. These brain areas’ signals feed into the priority map (possibly within the FEF), which ultimately determines selection.

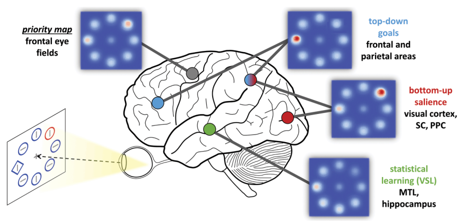


Figure 5.1: A schematic, created by Theeuwes and Failing (2020), indicating brain areas that mediate bottom-up attention, top-down attention, and a priority map.

5.2 Exercises

Answer these questions and relate them to the learning outcomes (2)

- What is top-down attention?
- In the experiment described by Theeuwes (2010), a unique color was used to attract attention. Can you think of something else that might have been used to attract attention?

Chapter 6

Explaining two-picture change blindness

For this unit, the term *change blindness* refers to the failure to notice changes in animations that alternate between two pictures of a scene. Later we will also talk about other situations in which people miss changes, but this chapter focuses on the two-picture alternation animations. In previous years, you probably already saw some of those amazing demonstrations. Here, however, we will learn somewhat different lessons than what you learned before.

First you need to realize that when we view a scene, we typically remember very few details about it. That's true even when we actively try to memorize the contents of the scene. When watching this movie, please scrutinise the scene carefully.

Isn't that amazing? In some ways, we humans are a lot dumber than we think!

6.1 Blindness for gradual changes

Let's start with very gradual changes rather than alternation of two pictures of a scene. Please view this movie. One might expect that when continuously viewing a scene, you would notice any major changes that occur during that time.

6.1.1 The "grand illusion of visual experience"

Most ordinary people are surprised by change blindness. Many researchers were very surprised, too, and some arrived at the conclusion that there is a *grand illusion of visual experience*. This is the claim that while people think that they are simultaneously experiencing the whole visual field, they are wrong about that - it is an illusion. These researchers explain change blindness with the claim that at any one time, you are only experiencing a small portion of the visual field, parts that you are particularly attending to. In other words, these researchers claim that *visual experience* is subject to a strong bottleneck.

However, this conclusion that there is a bottleneck on visual experience may be premature. To understand why, we need to consider in more detail what the failure to notice changes might mean. We need to consider the processing that's needed to detect a change.

6.1.2 What is needed to detect a change?

Let's consider what it takes to detect the change of an object or part of a scene:

1. An internal representation of that object that is different before and after the change.
2. A process that compares what was identified earlier to what is being identified now.
3. A process that calls attention to, or brings into conscious awareness, the instances of change.

Apparently, at least one of the above three processes is lacking. Let's consider #1 first. It is the case that all incoming retinal signals across the scene get processed. Unfortunately, however, if the object is in the periphery, the retinal signals may not be high-resolution enough for the representation to be different before and after the change. This is because vision is low resolution in the periphery (4).

For many real-world scenes, then, #1 above is sufficient to explain why people don't notice changes. However, #1 is not enough to explain all failures to notice changes. Even when researchers create displays in which all the objects are big enough and widely-spaced enough to see in the periphery, still people miss many changes. For example, the changes are large enough in some classic demonstrations like this boat scene.

So, #2 or #3 or both are lacking. This is likely due to a bottleneck. These processes are limited in capacity, so they cannot simultaneously process all objects in the visual scene. And what about the "grand illusion of visual experience?" Well, it seems quite possible that we may have experience of objects without having processes that correspond to #2 and #3. In other words, the conclusion that there is a grand illusion of visual experience may be a hasty one (Noë et al. (2000)). When people are surprised by change blindness, their mistake may be failing to realise that there's various processes required to notice a change, and visual experience may not always involve those.

Only a finite number of neurons can fit in our head, and evolution seems to not have prioritized processing of #2 and #3. The brain has not devoted neurons to constantly comparing what you're seeing now to what you saw half a second ago. Comparing what was present at two different times requires the limited resources of attention to be at that location at the two different times. We don't know why evolution did not prioritize these, but one possibility is that a full comparison process (#2) would require a lot of neurons, and animals like us have been able to get by with other, simpler processes, which we will discuss next.

6.2 Bottom-up attention and flicker/motion detectors

While limited capacity means we can't fully process the whole visual scene simultaneously for changes, brains have evolved some simple tricks that help us catch many changes. One of these is that our brains have flicker or motion detectors that *do* simultaneously process every part of the scene.

At my home, mounted high in the corner of the carport, is an inexpensive motion detector that you can buy at the hardware store. This device is wired so that if it detects motion, the carport light comes on. There is nothing fancy about the processing within it - not much circuitry is required.

When done by neurons, too, crude motion detection doesn't require much work or energy (you can learn more about this in PSYC3013). In the retinotopic maps in visual parts of our brain, each bit of the map has flicker/motion detectors sitting there that fire as soon as something happens in the scene. Sudden disappearance of sudden appearance of an object will drive these flicker/motion detectors to fire.

Firing of those flicker/motion detectors can call attention to a location. This is an instance of bottom-up attention (introduced in the previous chapter). Thus, the brain uses bottom-up attention as a work-around: the flicker or motion ordinarily caused by a change summons attention to a location, and then more limited-capacity processes work out what's changing there.

In other words, we have evolved to process simultaneously across the scene only a few things. Two of these things are flicker and motion. Thus, detecting motion and flicker is NOT capacity-limited. We rely on this to signal the locations where something is happening.

Very gradual changes do not trigger our flicker/motion detectors. But when changes are sudden, this ordinarily stimulates motion or flicker detectors, which in many circumstances calls attention to the associated location.

This is one reason that animals stay very still when they are worried about predators. Thanks to their camouflage, they can be hard to notice when they're not moving, but as soon as they move, they're quite conspicuous (watch this) and predators' attention goes straight to them.

However, what happens if motion or flicker occurs in multiple places? It won't be clear which location attention should go to and our limited-capacity processes may be overwhelmed. O'Regan et al. (1999) have demonstrated this with what they call "splashes" - see Traffic with splashes and Traffic without splashes. The idea is that these movies might resemble the situation if splashes of a puddle hit your windshield while you are driving - the splash would call your attention, preventing your attention from going the location of potentially-important other changes.

Broader background motion can also present a problem - if everything in the scene is moving, then our motion detectors are stimulated everywhere and attention may not go to the location of a change.

Knowing all this is necessary to fully understand why people take a long time to find the change in the classic change blindness animations.

6.3 Two-picture change blindness

Animations like that of the boat scene or this Paris scene sandwich a blank screen in between the two versions of the picture. It's like a blank screen sandwich! (The two pictures of the scene are analogous to the two chocolate biscuits and the ice cream is analogous to the blank screen).

Here is a schematic of the timeline, wherein the arrow represents time.

That blank screen is critical - it creates flicker everywhere in between the two frames. That is, when the picture of the scene is replaced by the blank scene, it creates a flicker signal everywhere, and then more flicker everywhere when the second scene comes on.

Please view this animation of a change blindness video with the blank screen removed. In that animation, the only location that tickles your transient detectors is that of the change. As a result, your attention goes straight to the location of the change.

Without the blank screen, the only location of flicker was the location of the changing object. The flicker called your attention to that location. *With* the blank screen, there's flicker everywhere, so there is no indication of which of the many locations contains the change.

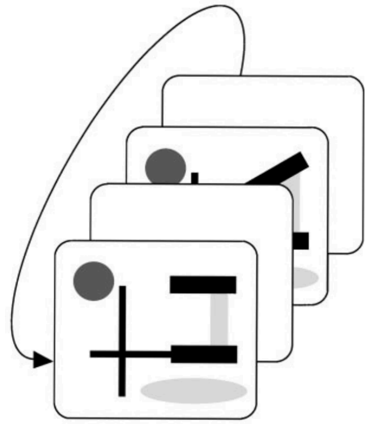


Figure 6.1: Two pictures of a scene are alternated, with a blank screen between them

6.4 Searching without a clue

When the blank screen is in the animation, the flicker/motion detectors provide no clue as to the location of the change. You might think that in this situation, people would search about randomly, or perhaps in a systematic fashion, something like searching from left to right and then top to bottom.

Consider the dinner date change blindness scene.

Researchers have recorded eye movements of people viewing these animations to see where they look - what they fixate on. These eye movements provide a pretty good indication of where people direct their attention. In a previous chapter (4), you learned that most movements of attention are overt - if people are interested in something, they look at it.

The above image shows some data from an eyetracking experiment. The long straight lines represent big jumps of the eyes from one place to another as the participants tried to determine what was changing. As you can see, the eyes dwell mostly on the couple's faces, their hands, and some objects on the table. So, the locations that people looked were not random at all.

You already knew that unique colors and other features are salient and attract attention, but here you can see that other properties of a scene affect attention.

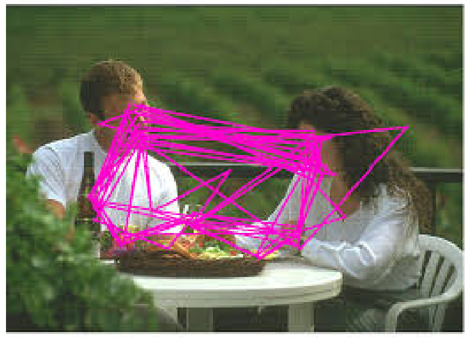


Figure 6.2: Pink lines indicate the trajectory of eye movements made by people searching for a change.

Over two thousand years ago, Aristotle wrote that "Man is by nature a social animal." People are very interested in people, and in working out what they're thinking and feeling. Things like people, bodies, and objects like food and wine are sometimes referred to as a scene's **high-level properties**. The word "high-level" is used in part to indicate that they take more processing to extract, and thus are represented at later ("higher level") stages of the brain compared to, say, color.

To understand the meaning of this scene involves working out the facial expressions of the people and how they are interacting with each other, based on their postures and the objects in front of them. People are so captivated by this that these participants never looked at some of the other objects, like the railing behind the couple (which was what was actually changing in the change blindness animation).

The above image shows only a few participants' data on one particular scene. Does the pattern of fixating the eyes more on face and bodies hold for other stimuli as well? Rigby et al. (2016) investigated this issue. Sixteen participants watched twelve four-second movie clips and twelve still-frame images from several episodes of a TV show (the *Andy Griffith* show) which had a lot of dialogue and characters. No soundtrack was presented.

These data provide further support for the hypothesis that attention is biased towards faces. Rösler et al. (2017) flashed pictures of scenes for just a fifth of a second, so people had time for only one eye movement, and found that people disproportionately looked at parts of the scene with faces or bodies.

These results are used by webpage and advertisement designers who seek to control or guide your attention. You may have noticed that many ads have a picture or animation of a person in them, even when this is completely unnecessary and superfluous to the information provided. Advertisers strive to hack your attentional system to get you to read/watch their ads.

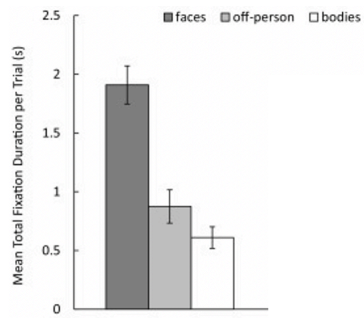


Figure 6.3: Average amount of time spent looking at different parts of the scene.

There is some evidence that the attention of many children with autism spectrum disorder is less biased towards faces than is that of typically-developing children. The study whose results are plotted above also included a group of sixteen adults with autism spectrum disorder, as shown below.

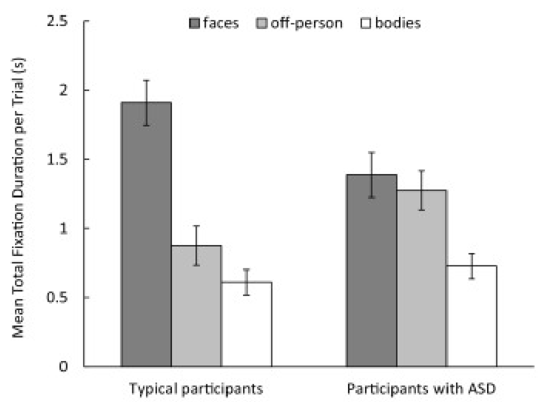


Figure 6.4: Average amount of time spent looking at different parts of the scene, in sixteen adults with (right) and without (left) autism spectrum disorder.

What would you expect, then, for the pattern of performance in change blindness in people with autism spectrum disorder? Kikuchi et al. (2009) conducted a change blindness experiment and varied whether what changed was the head of a person, another object, or a change to the color of the background.

The blank screen sandwich was looped until the participants pressed a key. The participants were then required to report what the change was, by pointing at it or with a verbal description. As one measure of performance, the researchers examined only those trials where the participants correctly detected the change and plotted the response time on those trials. Lower response times suggest that the person detected the change faster.

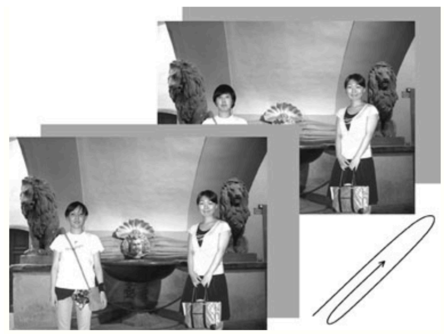


Figure 6.5: A schematic of one of the trials in the experiment. This trial is an example of the head change condition. The head of the person to the left was replaced by another head.

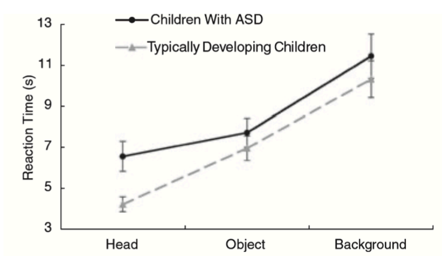


Figure 6.6: Mean correct response time with standard errors. The black line represents children with ASD and the broken gray line typically developing children.

6.5 Exercises

Answer these questions and relate them to the first four, and the last, learning outcome (2):

- Why can classic change blindness animations be described as a "blank screen sandwich"?
- Why are gradual changes hard to detect?
- What effect do mudsplashes and other irrelevant sudden changes in a scene have on our ability to detect important changes? How do they have that effect?

Chapter 7

Three kinds of attentional selection

Because we have limited capacity, we need to attend to things to get them processed. This is assisted by three abilities that we have:

- Location selection
- Feature (e.g., color) selection
- Object selection and tracking

In this chapter we'll talk about location and feature selection - a later chapter will discuss object selection and tracking.

7.1 Location selection

People typically move their eyes to look directly at where they are attending when they are selecting a single location.

But sometimes, people attend to multiple locations simultaneously. Then, their eyes point at one location but they may be attending to other locations simultaneously. In any previous readings or units of study you've taken, location selection may have been the only kind of attention discussed. For example, in standard cuing experiments (Posner et al., 1980), a location is cued, after which participants perform better at processing things in that location, because the participants attentionally selected that location.

The schematic illustrates a shape identification experiment. The arrow cues the participant to attend to the right, which results in them identifying the shape there (the star) more quickly than if the arrow had directed the participant's

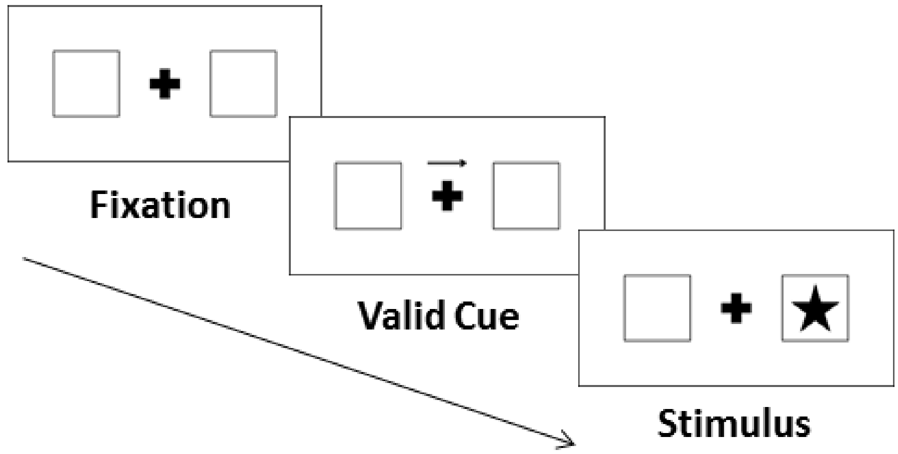


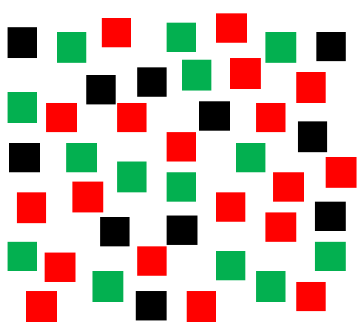
Figure 7.1: A schematic of a location cuing experiment (created by Local870, CC BY 3.0)

attention somewhere else. We'll revisit location cuing in some new contexts in 11.

7.2 Feature selection

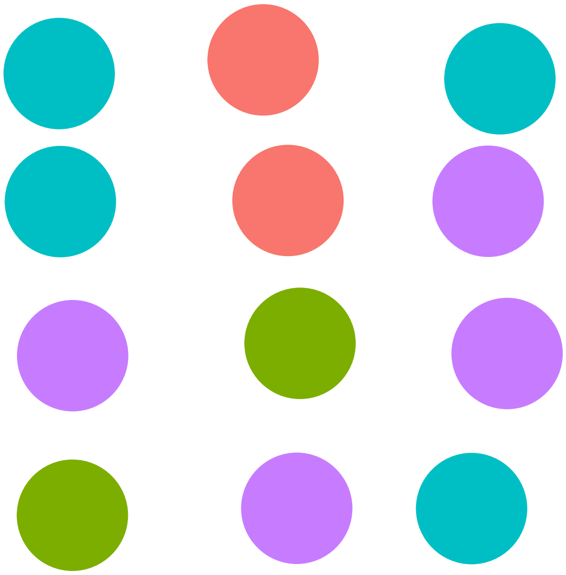
People can select not only an individual location, but also an individual feature value, for certain features such as color.

In the below image, try just attending to the red objects. With some concentration, you may feel you can do it just by thinking of red.



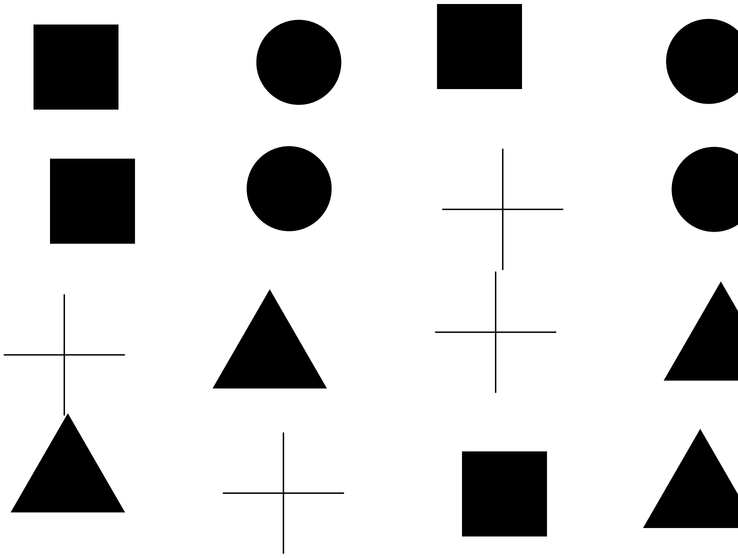
Next, try concentrating on all the black objects. You can use your power of feature selection to do so. Similarly with the green - try it.

For any of these forms of selection to work as effectively as they do, your brain must be processing the color and the shape information simultaneously in parallel across the visual scene. If it didn't, you'd be reduced to moving attention to each location to process its color and shape.



In the below, try concentrating on purple, green, or blue.

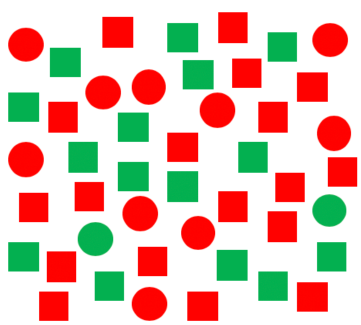
Feature selection also works pretty well for simple shapes. In the below, try



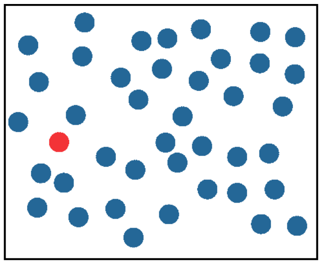
thinking of squares, circles, triangles, or pluses.

7.3 Feature combination selection?

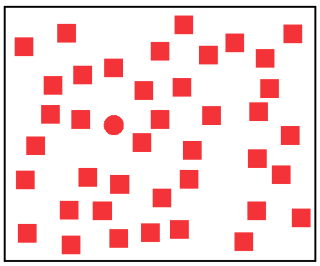
However, selecting a combination of features does not work as well. You can think "green circles", but unless you're so lucky that your attention semi-randomly lands on them right away, it will take longer to find the green circles.



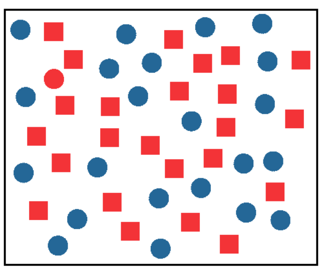
Similarly, in the below image, in *b* it is easy to find the red object, in *d* it is easy to find the circle, but in *f* it takes longer to find the red circle. Selection of a combination of features is not effective.



(b)



(d)



(f)

Feature selection also is effective for direction of motion, such as to select upward or downward motion (Sàenz et al., 2003).

7.4 Recap

Color, shape, and motion direction are processed in parallel across the visual field, in some fashion that allows attention to be guided by them. Researchers don't fully understand yet how this works, but as will be decribed in 9, Anne Treisman proposed that there is a separate map for each feature, which can guide attention.

Whether the mechanism turns out to be a feature map or something else, when we think of a particular color, shape, or motion direction, our brain can enhance activation of the associated neurons rapidly. But we can't do so for a particular *combination* of color and shape. Instead, we end up enhancing the activation of both the objects that are that particular color *and* the objects that are that particular shape; we can't confine our attention to those that have that *combination* of features. For example, if the color is red and the shape is circle, we'll activate all the red objects (even if they're squares) and all the circular objects (even if they're blue), rather than just the red circular object.

7.5 Complex shapes

While shapes that differ dramatically from each other allows parallel selection, like of the circles versus squares above, we do not have the ability to have our visual cortex enhance all instances of a *complex* shape.

In the next search array, your task is to find the shape that does NOT have a squiggly tail.

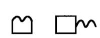
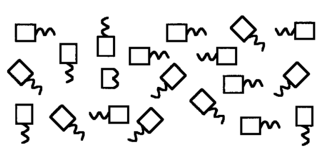


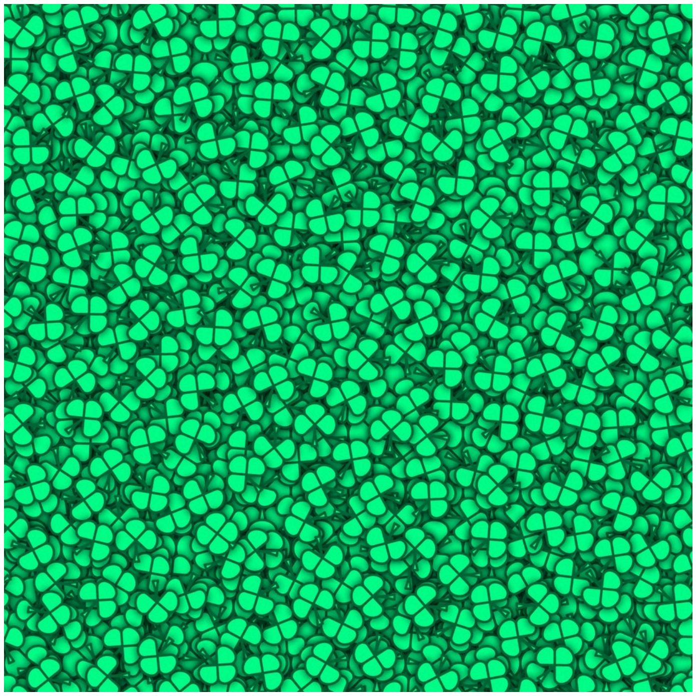
Figure 7.2: The target (left) and a distractor (right)

The search array is below. Find the target - the object without a squiggly tail.



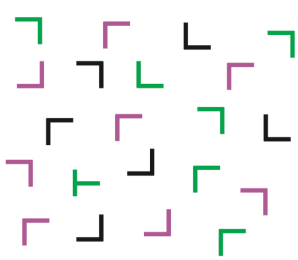
Finding the object without a squiggly tail takes, on average, much longer than in the simple search cases above.

How quickly can you find a lucky four-leaf clover?



Finding a four-leaf clover is hard! One reason is that feature attention is not effective for complex shapes (another reason, in the above display at least, is that the clover is small and crowded, so you have to look at each patch almost directly to perceive its leaves clearly).

Even fairly simple shapes like a 'T' can be beyond the powers of our featural shape selection. I wish we could just think to ourselves, 'T!', and enhance the activation of any T's in the scene, making our attention go straight to their locations, but we can't. So, on average it takes quite a while for attention to go to the location of the T.



In summary, while we have some ability to do featural selection for color, shape, and motion, we can only do so for very simple shapes. The reason for this is not fully understood. Something about how the brain is connected up allows us to think "red" and quickly attend to all the locations in a scene that contain red, but this ability extends to only a few features.

7.6 Exercises

Answer these questions and relate them to the seventh learning outcome (2):

- What does "feature selection" mean?
- What kind of selection is Posner cuing an example of?

Chapter 8

Binding takes time

The previous chapter explained that we can set our attention to process individual locations, or individual features. But we do poorly with combinations of features. The difficulty with encoding combinations of features, but not individual features, can be quite striking in some simple animations, as you will see below.

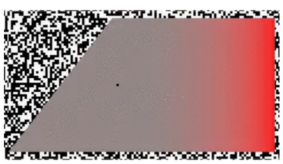


Figure 8.1: Task: judge whether the red color is paired with leftward tilt or rightward title.

If the animation above doesn't work, watch it here.

What's going on in that animation should be easy to perceive - the color is alternating between green and red, and the left side edge is alternating between leftward-tilted and rightward-tilted. You can also easily see how the features are paired; that is, when the left side is leftward-tilted, the right side is red.

If we speed up the alternation of the two frames, however, we can reveal the difficulty our brain has with combining features (Holcombe and Cavanagh, 2001).

If the above movie doesn't work on your web browser, watch the same speeded-up version here. It is still easy to see that the individual features are red/green and left/right, but it's hard to judge which occur at the same time. The brain takes too long to identify combinations of features to know which were presented simultaneously.

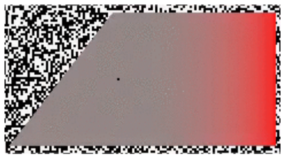


Figure 8.2: Task: judge whether the red color is paired with leftward tilt or rightward title.

The previous chapter explained that we can select objects in the scene by certain individual features such as by location and by color. For example, we can think "red" and our attention will process all the locations containing red more. However, this doesn't work for combinations of features such as red and leftward-tilted. The present chapter showed another side of our trouble with combinations of features - simply identifying pairs of features can be quite time-consuming.

To experience that more, watch the below two-frame animation. This time, the challenge is to combine two different orientations. At the slow rate (top row), you can do it. But in the middle row, you don't have enough time. If the below animation doesn't work in your web browser, try watching it here

Binding can also be time-consuming for color with motion direction, as illustrated by the animation below.

While at the slow rate of the top row, it is easy to judge the pairing of motion direction and white/black color, it is very difficult in the middle row, where the speed is slightly faster.

In each case, the reason for the slow limit on binding, we believe, is that early visual processing stages work very quickly and process the multiple features in parallel. In contrast, determining the pairing of features requires applying limited-capacity resources (there's a bottleneck!). These processes work more slowly. Possibly, attention must select first one location, and then the other location, to bring the features together before the animation changes. I have set the animations' playback speed to a rate fast enough that attentional selection is too slow to get this done (Holcombe, 2009).

In the previous chapter you learned that our brain has a process for attentional selection of an individual feature, but not for combinations of features. In this chapter you learned that combining features can be done, but is time-consuming. In the next chapter, we will put these ideas together to better understand classic findings from visual search experiments that you learned about in first-year psychology.

8.1 Exercises

Answer these questions:

- What do the slow limit on pairing simultaneous features have in common with what was said in the previous chapter about selection?
- How do the demonstrations above relate to Learning Outcome #6 (2)?

Chapter 9

Visual search

From time to time, we all need to find things. Rummaging through our closet for a particular shirt, or wandering about the house trying to find our keys, or for some of us, groping about for our spectacles that we know we put down *somewhere* around here.

Our search performance in some tasks can reveal aspects of the bottlenecks in mental processing. Slow search can suggest a bottleneck is affecting processing. However, sometimes search is slow because the basic sensory signals are not good. For example, when I lose my spectacles, my vision is so poor that I have to bring my face close to each location in the room to check whether my glasses are there. Similarly, wandering about the house looking for one's keys, to evaluate all the rooms of the house, one has to visit each room.

Sometimes, even though something is right in front of our face, the sensory signals aren't good enough for it to be possible to know that the object is there. For example, try searching for the word "wilt" in the below image, which shows the first two pages of Shakespeare's *Romeo and Juliet*.

Did you find "wilt" yet? To find it , your eyes have to move back and forth (it's about 3/4 of the way down the left page). The task is impossible to do without moving your eyes. The main reason for this is that the sensory signals provided by your retinas are only good enough to read small words near the center of your vision. So, you have to move your eyes.

To see this sad fact about your vision more directly, try the following. Stare directly at the black cross and, while keeping your eyes fixed on the cross, try reading any of the words on the bottom of the page. You can't do it. Not because of any bottleneck, or problem with selection, but simply because your photoreceptors are too widely spaced in the periphery. That is, outside of a central region, the spatial resolution of your vision is too low to see many details. The sensory signals from the periphery are coarse.

Enter Sampson and Gregory (with swords and bucklers) of the house

Samp. Gregory, on my word, we'll not carry coals.
Greg. No, for then we should be colliers.
Samp. I mean, an we be in choler, we'll draw.
Greg. Ay, while you live, draw your neck out of collar.
Samp. I strike quickly, being moved.
Greg. But thou art not quickly moved to strike.
Samp. A dog of the house of Montague moves me.
Greg. To move is to stir, and to be valiant is to stand. Therefore, if thou art moved, thou runn'st away.
Samp. A dog of that house shall move me to stand. I will take the wall of any man or maid of Montague's.
Greg. That shows thee a weak slave; for the weakest goes to the wall.
Samp. 'Tis true; and therefore women, being the weaker vessels, are ever thrust to the wall. Therefore I will push Montague's men from the wall and thrust his maids to the wall.
Greg. The quarrel is between our masters and us their men.
Samp. 'Tis all one. I will show myself a tyrant. When I have fought with the men, I will be cruel with the maids- I will cut off their heads.
Greg. The heads of the maids?
Samp. Ay, the heads of the maids, or their maidenheads. Take it in what sense thou wilt.
Greg. They must take it in sense that feel it.
Samp. Me they shall feel while I am able to stand; and 'tis known I am a pretty piece of flesh.
Greg. 'Tis well thou art not fish; if thou hadst, thou hadst been poor-John. Draw thy tool! Here comes two of the house of Montagues.

+

Enter two other Servingmen [Abram and Balthasar].

Samp. My naked weapon is out. Quarrel! I will back thee.
Greg. How? turn thy back and run?
Samp. Fear me not.
Greg. No, marry. I fear thee!
Samp. Let us take the law of our sides; let them begin.
Greg. I will frown as I pass by, and let them take it as they list.
Samp. Nay, as they dare. I will bite my thumb at them; which is disgrace to them, if they bear it.
Abr. Do you bite your thumb at us, sir?
Samp. I do bite my thumb, sir.
Abr. Do you bite your thumb at us, sir?
Samp. [aside to Gregory] Is the law of our side if I say ay?
Greg. [aside to Sampson] No.
Samp. No, sir, I do not bite my thumb at you, sir; but I bite my thumb, sir.
Greg. Do you quarrel, sir?
Abr. Quarrel, sir? No, sir.
Samp. But if you do, sir, am for you. I serve as good a man as you.
Abr. No better.
Samp. Well, sir.

Enter Benvolio.

Greg. [aside to Sampson] Say 'better.' Here comes one of my master's kinsmen.
Samp. Yes, better, sir.
Abr. You lie.
Samp. Draw, if you be men. Gregory, remember thy swashing blow.
They fight.
Ben. Part, fools! [Beats down their swords.] Put up your swords. You know not what you do.

Figure 9.1: The first two pages of Romeo and Juliet, by William Shakespeare.

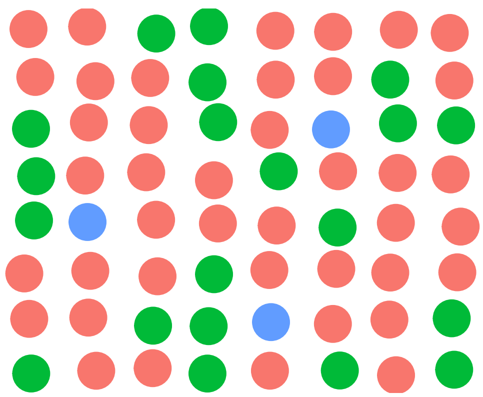
9.1 Information overload

A good way to assess whether there is a bottleneck in a system is to give it more and more things to process and see whether this degrades performance or whether the system can process each just as quickly as when it is given just one. Psychologists did this for visual processing by giving people many stimuli to process, by adding more and more to a display. In doing this, however, they had to be careful to make sure that the brain had a chance, by making sure that a person could see each individual stimulus even when it wasn't in the center of their vision (unlike in the Romeo & Juliet demonstration above). If the person couldn't even see the stimuli, then of course the brain wouldn't process it well even if it didn't have a bottleneck.

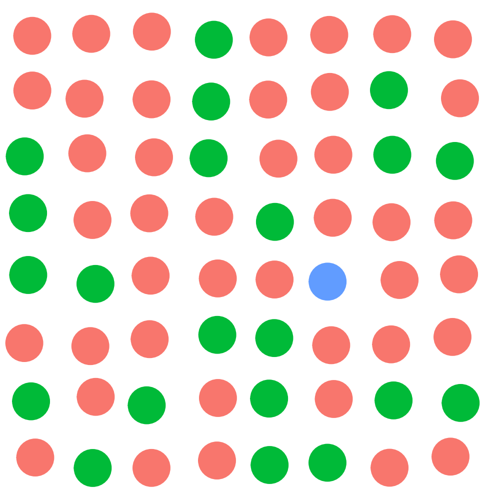
One of the tasks psychologists have used for this is called "visual search". In a visual search experiment, people are shown a display with a particular number of stimuli and asked to find a target. This is discussed this in first-year psychology. The next section is, in part, a review of that.

9.2 Parallel search

In a previous chapter (7), you learned that you can select stimuli by their location, or by an individual feature such as a particular color. That is, if you just think about a particular location, or a particular color, your attention tends to go the appropriate place(s). For example, if you think "blue" while looking at the display below, your attention will go to the blue dots quite quickly. Stare at the center of their display and concentrate on selecting the blue dots.



To capture this effect in experiments, researchers typically present just a single object with the target characteristic. The task is to press a key if the target is present, and in half of trials the researchers present the display with no target.



The data indicate that people can find a blue object quickly no matter how many other objects there are in the display. This is called "parallel search" because the evaluation of objects occurs simultaneously across the entire scene. In other words, the processing happens before the bottlenecks in the brain.

The associated pattern of experiment results is demonstrated by the graph below.

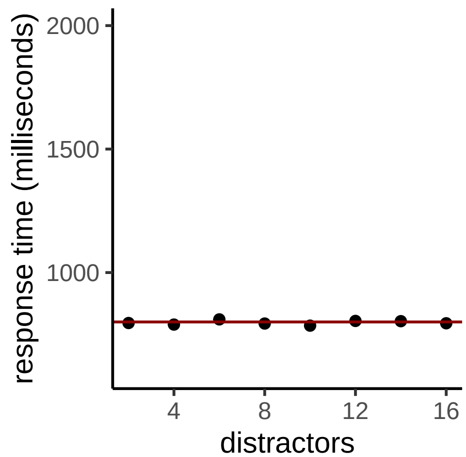


Figure 9.2: Time to find a lone blue circle among red and green circles

The diagram below provides a basic idea of the processing stages involved. First, the first stages of your visual brain determine the color of each object in the display, processing them all simultaneously. Then, by simply thinking about red, the red neurons' activity is enhanced and your focused attention will end up going to the location of any red objects present.

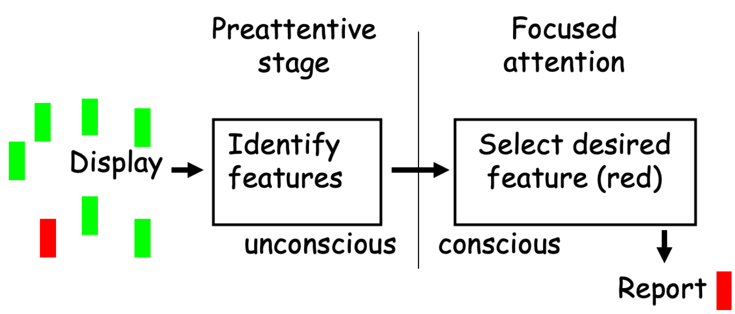


Figure 9.3: A schematic of the processing stages involved in searching for red.

9.2.1 Bottom-up attention also plays a role

A complication in visual search experiments for a lone target (like a lone blue circle in a sea of red and green) is that it is the only item in that display which has that feature. As described in 5, bottom-up attention will drive attention toward the location of the uniquely-colored object.

For a search like for the blue target above, then, both feature selection (thinking "blue") and bottom-up attention combine to make the search particularly easy.

9.3 Processing one thing at a time

But parallel search doesn't happen in most cases for *combinations* of individual features. Instead, there is a bottleneck. To put that in context, I will remind you of aspects of parallel versus serial processing.

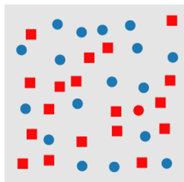
Imagine you were in an art installation where the artist had hung many speakers from the ceiling, and each speaker played a different person's voice, each telling a different story Well this is precisely the situation I was in one day when I visited a museum in Havana, Cuba. What I heard sounded like an incoherent jumble. I didn't hear any of the actual stories being told by the voices until I moved my ear up against an individual speaker. In other words, I could only process a single auditory stimulus at a time, and to do so, I had to select it.

A forest of speakers is not a situation you are likely to encounter. It does illustrate, however, one possibility for sensory processing - for certain things, you may be unable to process multiple signals at once. In that case, you need to select one stimulus to concentrate on it.

Fortunately, our visual brain can process certain aspects of the visual scene in parallel. But for combinations of features, you are in much the same boat as I was in that day in Havana, having to select individual locations to evaluate an aspect of what is present - specifically, the combination of features there.

9.4 Combinations of features - serial search

In the below display, your task is to search for the red circle, which is a combination of features - red and circular.



As a result, instead of being able to rely on parallel processing to rapidly tell you where the target is, you have to bring limited attentional resources to bear. Those resources can only process a few objects at a time (they impose a bottleneck). So, the more distractors there are, the longer it takes (on average) to find the target. The below plot shows the average response time as a function of number of distractors.

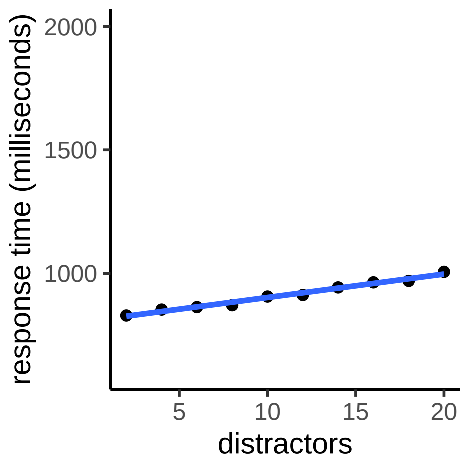


Figure 9.4: Searching for a lone red circle among blue circles and red squares. The average time it takes for a participant to find the target increases steadily with number of distractors.

Now view the below image, which is made up of two displays, the left half and the right half.

On the left half of the display, it is very easy to find the red vertical item, simply by using feature selection for red.

But in the right half of the display searching for the red vertical item is much more difficult. This is because the target differs from the distractors not by a single feature, but rather by a combination of features. This is called **conjunction search**. That is, conjunction search is search for a target among other objects that have the constituent features of the target, but in different combinations. Here, the target is the only red vertical object, but some other objects are red (but with a different orientation) and some others are vertical (but with a different color).

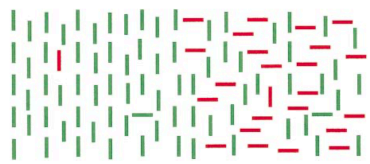


Figure 9.5: Two search displays. Left half: feature search. Right half: conjunction search

Assessing combinations of features requires a limited-capacity process. Therefore, attentional selection must rove about the display until the target is found. This was first suggested by Anne Treisman and called "Feature Integration Theory". Treisman specifically proposed that attention must individualy select each object, to one-by-one to evaluate what combination of features it has. That's quite the bottleneck!

Conjunction search search thus yields a very similar result for the pattern of response times as did the previous search.

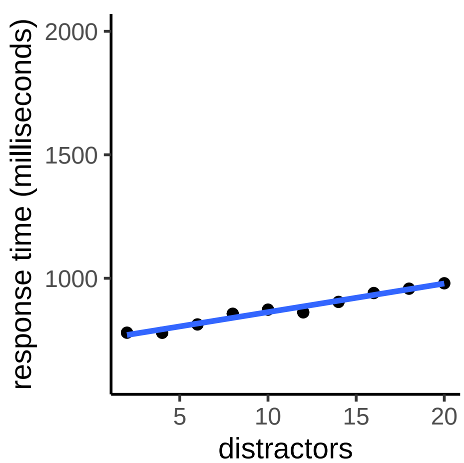


Figure 9.6: Searching for a vertical red line among vertical green lines and horizontal red lines.

The linear slope indicates that each additional distractor imposes a cost. This is expected if one can only evaluate one, or a few, items at a time. The longer it takes to evaluate each item for whether it is the target, the steeper the slope will be. For the graph above, the slope of the line is such that it rises by 100 milliseconds (one-tenth of a second) for every additional ten distractors.

Serial search is the theory that to complete a particular search task, some process in the visual system has to evaluate the stimuli one-by-one, or maybe two-by-two or three-by-three; the main point is that it is capacity-limited and thus can't process all the items at once. Because conjunction search typically has a linear positive slope, many researchers have concluded that such searches occur serially.

Here is a basic idea of the brain stages thought to be involved in such a serial search.

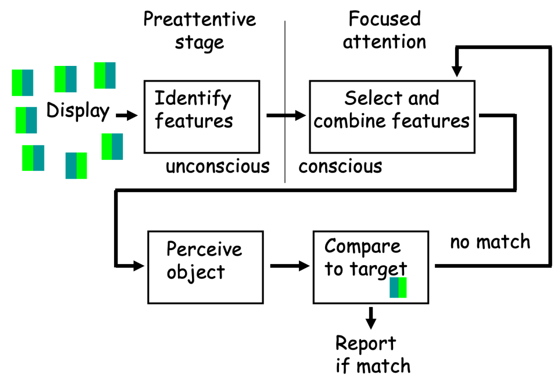


Figure 9.7: The figure schematizes for a target with dark green on left, light green on right. Combining the two features is required, which happens just one object at a time.

In visual search, if one starts at a random place in the scene and then evaluates each item in a random sequence, on average one will only have to visit half of the items before one lands on the target. So if response time is 100 milliseconds slower when there are 10 more distractors, on average you only had to evaluate half of those distractors in that time, so the search rate is 100 milliseconds / 5 distractors = 20 milliseconds per distractor. So if search was happening one-by-one, people were searching the scene at a rate of about 20 milliseconds per item.

Twenty milliseconds per item is pretty fast! After all, that'd be fifty items per second. But researchers are not sure whether or not the serial search happened one-by-one. Instead, people might be able to evaluate, say, three items simultaneously. If so, then it'd be evaluating each group of three every sixty milliseconds (about 16 groups of three per second).

Next is another search that does not occur in parallel. Look for the red vertical line.

In the above display, vertical items and red items are interspersed throughout, making individual feature selection completely useless - you really have to evaluate each location for what combination of features is present. Therefore, the more items are in the display, the longer it takes to find the target.

The slope of this graph is even steeper than that of the previous two searches.

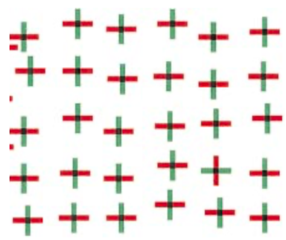


Figure 9.8: A conjunction search for red and vertical.

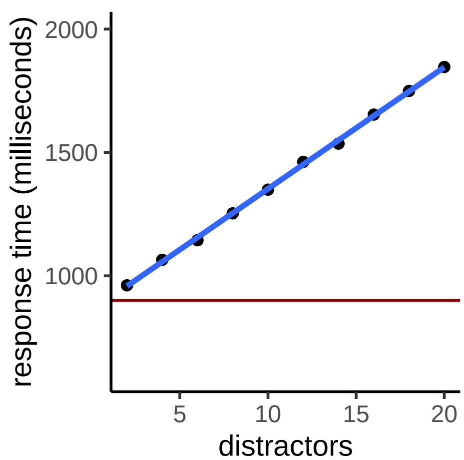


Figure 9.9: Searching for a vertical red line in the crosses display.

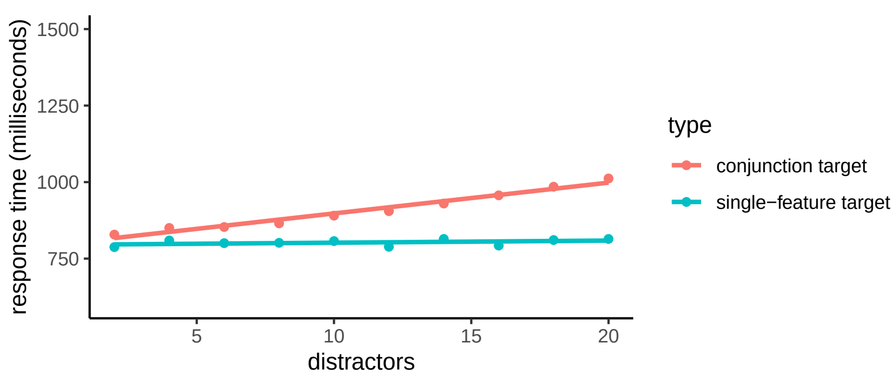
Here, each additional ten distractors increases response time by a full half second (500 ms). That is, the search rate is 50 ms per stimulus, or twenty per second.

To put this time in perspective, we can compare the rate at which this feature-combining bottleneck can process things to the rate at which the CPU of a modern computer can process things. Recall that in 3, we explained that the CPU of a computer is a bottleneck. The CPU can only process a little bit of information at a time.

Engineers haven't found many ways around this problem of the CPU being a bottleneck. However, technological progress has meant that the CPU can make up for its limited capacity with raw speed. The iPhone 11, for example, has a CPU that can perform 2660000000 operations per second. That's a lot faster than the twenty per second of the above human visual search and helps explain why, despite having a bottleneck like we do, computers can do some things much faster than we can. If you're interested to learn more about human-computer comparisons, you might consider taking PSYC3014 (Behavioural and Cognitive Neuroscience) next year.

Going back to humans, another case where features need to be combined is when searching for a colored letter. Click on this link and you will be asked to search for an upright orange T with inverted orange Ts and blue Ts as distractors.

9.5 Serial search versus parallel search



Watch an 11 min video about the 'human visual search engine' (with an accompanying transcript) starring Jeremy Wolfe that explains more.

Some points to take away from his video:

- Searching for your black cat in a white carpet, it's easy alone on the white carpet, hard among many other cats (if they have some black).
- Anne Treisman in the 80s suggested there are two kinds of searches, serial ones and parallel ones.

- Ts among L's don't jump out; search is not parallel. But most people can do 20 to 30 per second because they have so much practice reading.
- For visual scenes, there's two factors reducing the information you can process in parallel.

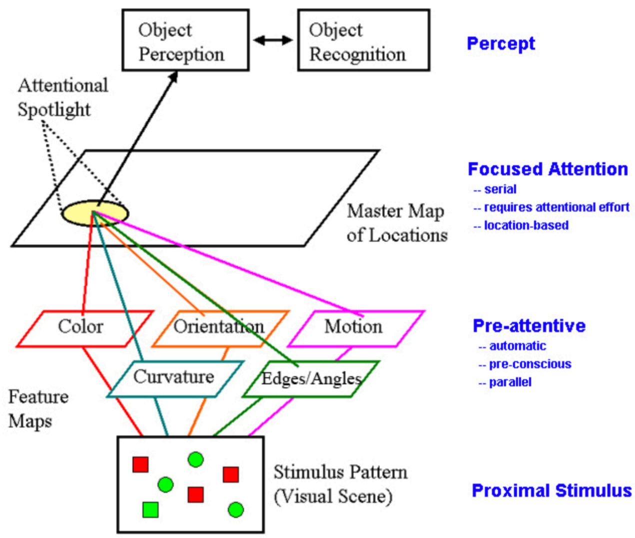
1. The poor resolution of the visual periphery.
2. The lower processing capacity of higher levels.

Figure 9.10: Anne Treisman

Now we can connect back to the early versus late selection question that was discussed in PSYC1. **Early selection** was the idea that there is an early bottleneck - that sensory information is not processed much before the bottleneck. If selection were very early, that would mean identifying features occurs after the bottleneck. Thus the brain identifies only a few features at a time. It's called "early selection" because selecting something for further processing means doing it early in the system.

Late selection was the idea that the brain is able to process sensory information from across the visual field in parallel. The bottleneck does not occur until much later. It's called late selection because selecting something for further processing means doing it after a lot of information has already been extracted.

These visual search results suggest that some basic features get processed in parallel, but these features are not integrated into complex shapes or objects. To do that, selection is required. Anne Treisman's theory of the processing architecture is schematized here

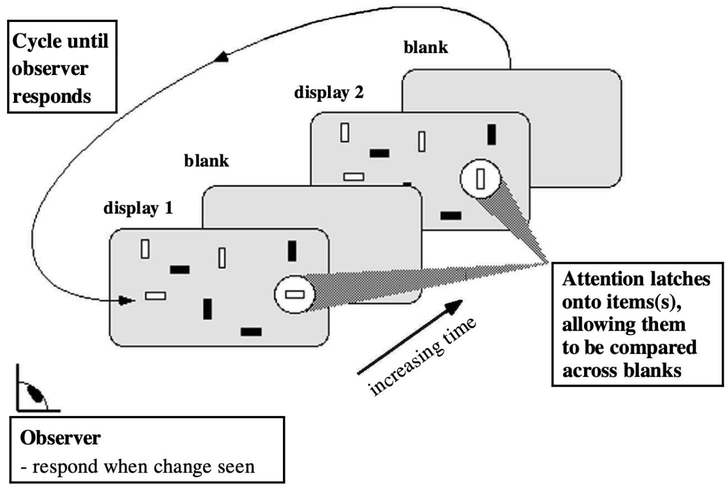


So, selection is after feature processing. This makes it late, relative to feature processing. But selection is before complex shape and object processing. So it is early, relative to complex shape and object processing.

9.6 Visual search and blank-screen sandwiches

Recall the blank-screen sandwich change detection animations of Chapter 6. In a typical blank-screen sandwich experiment, people are timed for how long they take to find the change happening in a photo of a scene. To better assess what is happening with attention, one can use a carefully crafted visual search display instead of a natural scene.

Rensink (2000) developed this technique.

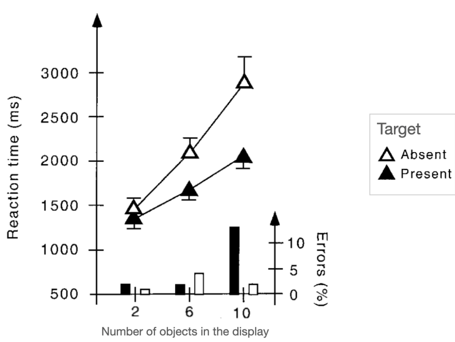


As schematized above, the participant was shown blank screen sandwiches with one object changing, and how long it took them to indicate the location of the changing object was recorded. The displays were shown for 800 ms and the blank screen was shown for 120 ms.

Rensink's hypothesis was that evaluating whether a change is present can only be done for one or a few items at a time, by attentionally selecting that location.

- What do you predict should be the effect on number of items in the display on response time?

Here are the results:



On some trials, the target was absent (the unfilled triangles), and participants likely did not respond until they had evaluated every object in the display so they could be sure nothing was changing.

More important is the results for the trials when the target was present. The filled triangles show a steep increase in search time with number of objects in the display.

- Was this predicted by the hypothesis?

9.7 Exercises

- Why do people need to move their eyes for many searches?
- What factors can make visual search slow?
- Describe how the kinds of selection connect to visual search performance for different types of display - learning outcome #5 (2).
- How does the finding for visual search performance for feature conjunctions relate to the rate limit found for pairing simultaneous features in the previous chapter?

Chapter 10

Real-world search

You may remember that it was a technological problem that helped get psychologists going in studying capacity limits - the problem of fighter pilots having to deal with flying and monitoring things in the cockpit at the same time.

More recently, in the aftermath of the September 11 attacks on the U.S., a different problem with planes became prominent, causing governments to want to closely inspect passengers' luggage.

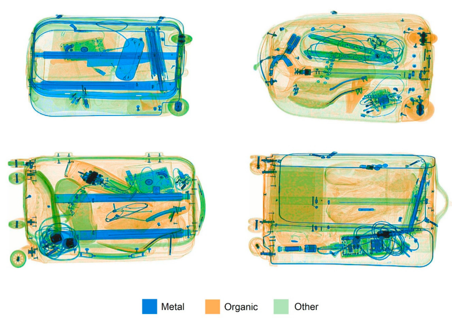
Figure 10.1: A plane crashing into a building in New York

In recent years, tens of thousands of workers around the world have spent hours every day doing one particular kind of visual search - baggage searches! These are the people that sit at airport scanners and search through scanner images for prohibited items.

This kind of search involves additional complications beyond those described in the previous chapter (9). In the searches of the previous chapter, participants only had to worry about finding one kind of target, for example a red vertical line. But airport workers have to keep in mind many targets - scisssors, knives, and lighters as well as guns.

Figure 10.2: Some items prohibited on flights in some parts of the world

Below you can see some examples of the baggage scan images they have to search.



Can you find the prohibited items in the bags above?

Which type of response time pattern do you think this kind of search would result in - parallel search or serial search? See the previous chapter 9 for a reminder of what this means.

10.1 Individual differences

Some questions for managers of airport security are:

1. Do people improve much with training at detecting threats?
2. How much do people improve?
3. With enough training, can anyone be made into an expert and have a similar level of performance?

Ericson et al. (2017) investigated this by making a game out of airport scanner inspections:

When Ericson et al. (2017) plotted the data from users of the game, they discovered something important.

Figure 10.3: A screenshot from the game Airport Scanner.

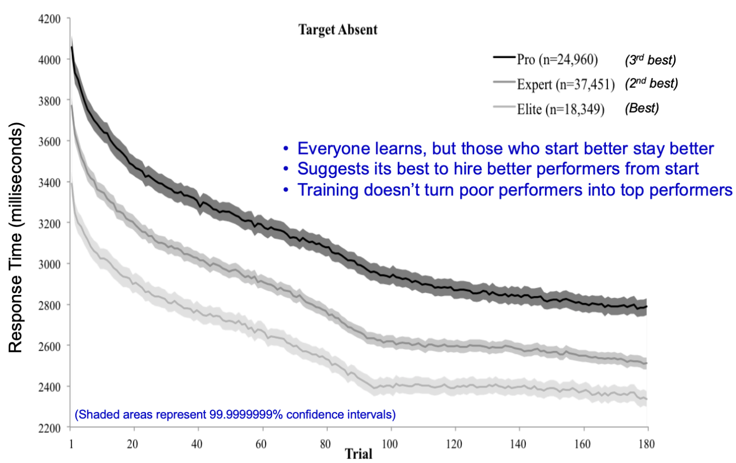


Figure 10.4: Response time when searching for a prohibited item versus trial number, for participants with 3 different levels of performance.

The horizontal axis is trial number, where "trial" means an individual try. The three lines are three different groups of people, those who start out (at trial 1) with relatively slow performance (top line), those who start out a bit faster (middle line) and those who start very fast (bottom line). Each group of players gets better from left to right - the more trials they participate in, the faster their response time.

It's good that everyone learns with practice. What's unfortunate, however, is that the curves do not converge - the slow people get much faster but the fast people get even faster, so the people who start out fast maintain their lead. In other words, the individual differences are stable.

From these results, the airport security managers might have some tentative answers to their three questions. What do you suggest for answers to the three questions?

10.2 Running, cycling, and driving are like continuous change blindness experiments

Cycling (private video shown during lecture)

Chapter 11

Attentional cuing

An attentional cue is something that directs one's attention to something.

Posner et al. (1980) showed that when a location is cued, participants perform better at processing things in that location. This is because the participants attentionally selected that location.

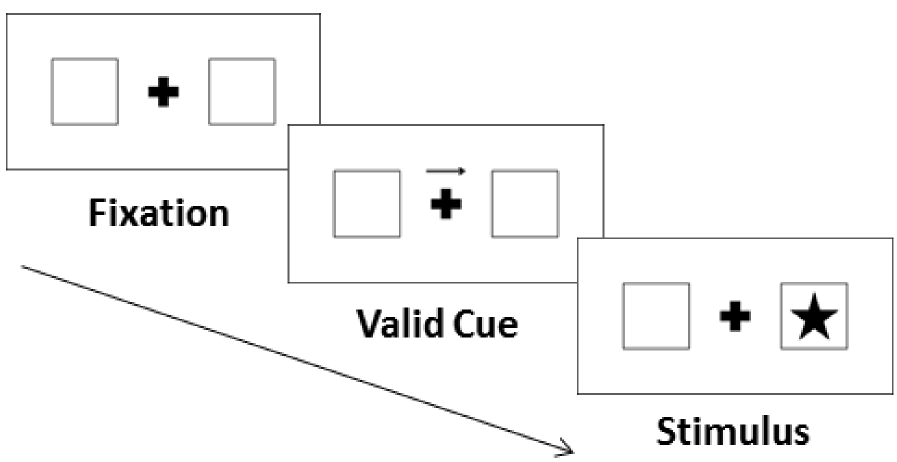


Figure 11.1: A schematic of a location cuing experiment (created by Local870, CC BY 3.0)

Some cuing is almost purely bottom-up, while other cases of cuing are a combination of bottom-up and top-down. Let's explore some interesting cases of cuing.

Because babies don't know anything, they can't go around directing their attention the location of food and other important things based on their knowledge of where they might be, or even what it looks like! They have to rely on bottom-up attention, which is often used by caregivers to direct the baby's attention. At two months old for example, many babies will turn their attention to a sound.

11.1 Side-eye or gaze cuing experiment: babies

The eyes play an important role in human communication; they are very expressive.

Most humans start interpreting others' eyes very early. Babies will not only turn their attention in response to the bottom-up cue of a sound, but also they learn to use the cue of others' eyes.

Babies' attention is cued by the direction of adults' gaze. This was documented scientifically by Hood et al. (1998), who tested 16 infants between 10 and 28 weeks old.

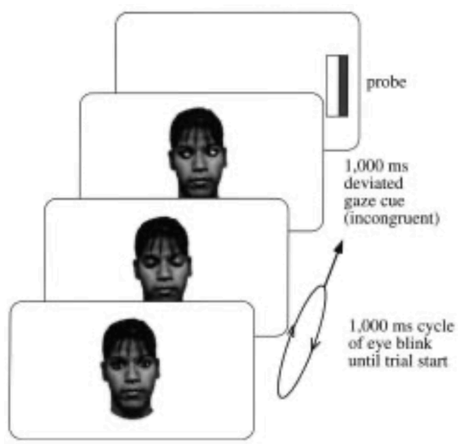


Figure 11.2: A schematic of the task used by Hood et al. (1998)

It might be a bit hard to see in the above image, but on the last screen before the probe, the eyes of the face are looking to the left. In half of trials, they looked to the left, and in half of trials to the right. After that face disappeared, a rectangle suddenly appeared on one side. The sudden appearance of the rectangle tends to attract the infants' attention.

The question was, would the frequency with which the rectangle attracted the infants' attention be affected by the direction the eyes before were looking in?

There were two kinds of trials:

- Congruent trials: the rectangle was on the side the eyes pointed at.
- Incongruent trials: the rectangle appeared opposite the side that the eyes were pointing at.

Usually, when the rectangle appeared, the babies looked at it (bottom-up attention). When the baby *didn't* look at the rectangle, the researchers scored that as an "error". The babies made *fewer* errors in the congruent condition than in the incongruent condition.

Table 1. *Mean percentage of orienting away from the probe (errors)*

	Cue validity	
Condition	Congruent	Incongruent
Face off (Experiment 1)	12.8	21.4

Figure 11.3: Results of the Hood et al. (1998) study

From this result, the researchers inferred that the face's gaze direction often directed the babies' attention.

Eye gaze is classifed as a top-down attentional cue. The bottom-up attentional cues are all things that cause attention to shift to the cue's location. Eye gaze directs your attention *somewhere else*. This requires interpretation likely involving limited-capacity (bottlenecked) processing at higher levels of the brain. After the cue is interpreted, top-down attention works to direct attention in the direction of the eyes.

11.2 Side-eye or gaze cuing experiment: adults

For adults too, following the gaze of others can be important. In evolutionary history, it made sense to look in the direction other people and animals were looking, because often what they were looking at was important, such as a possible threat, or a food source.

Friesen and Kingstone (1998) conducted an experiment much like that done on the babies. Here are the stimuli they used:

The experiment showed the participants the stimuli at a number of different intervals between cue (cartoon face) and the presentation of the letter. That interval from onset of the cue to onset of the letter is called the "stimulus onset asynchrony". The task of the participants was to hit a key to indicate whether the letter presented was an F or a T. The variable of interest to the researchers was how long it took the participant to hit the key. The idea is that if the cue directs the participants' attention toward the letter, the participant will respond more quickly. But the participants were informed that the direction in which the eyes looked was not predictive of the location or identity of the target letter or of when it would appear. Thus, the participants didn't have any reason to voluntarily move their attention in the direction of the eyes.

The plot below shows the average response time for trials where the participant typed the correct key.

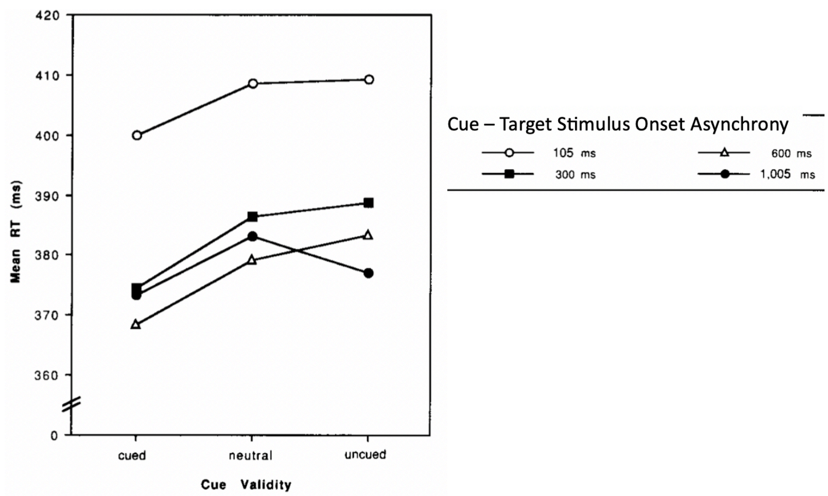


Figure 11.4: Response time against cue type, for four different stimulus onset asynchronies

- How was the response time different in the cued conditions compared to the uncued conditions?
- What condition yielded the slowest responses?
- Why did that condition yield the slowest responses?

These eye cuing effects may be even stronger with a person in the real world. The magician and pickpocketing artist Apollo Robbins explains how he uses eye cuing in this video (I've set it to begin playing at the relevant part).

11.3 Technology and media

In the commercial world, there are many companies and media outlets trying to get your attention. They use a lot of what we know about visual attention and cuing to push your attention to where they want you to attend. To draw attention to a company's logo, for example, they'll position a face near it, looking in the direction of the logo they want you to read. You may have noticed this on websites as well as on TV commercials.

Chapter 12

The role of memory and expectation

12.1 The contents of memory can guide attention

12.1.1 Working memory

In his lectures for this unit, Caleb told you a bit about working memory. As far back as 1893, it was suggested that "Impressions which repeat or resemble ideas already present in consciousness are especially liable to attract the attention" (Külpe, 1893), which was Külpe's way of saying that having something in one's working memory can guide attention to associated sensory stimuli.

Huang and Pashler (2007) did an experiment that provided quantitative evidence for Külpe's claim. Participants had to memorize a shape (the "prime shape"), knowing they'd be shown a shape at the end of the trial and asked whether it was the same one. They also knew that after the prime shape was presented, they'd be presented with three digits that they should memorize because they'd be asked to report them at the end of the trial.

Criticallly, the digits to be memorized were presented on a background of three shapes, and one of those three shapes had been presented previously as a prime shape to be memorized. On 10% of trials, that shape was the shape presented at the beginning of that very trial. The hypothesis was that if having a shape in mind guided attention, then people would do better at reporting a digit that was presented on that trial's prime shape than at reporting other digits.

The results show that if the shape that a digit was presented in was the prime shape for that trial, people did well at reporting that particular digit. But

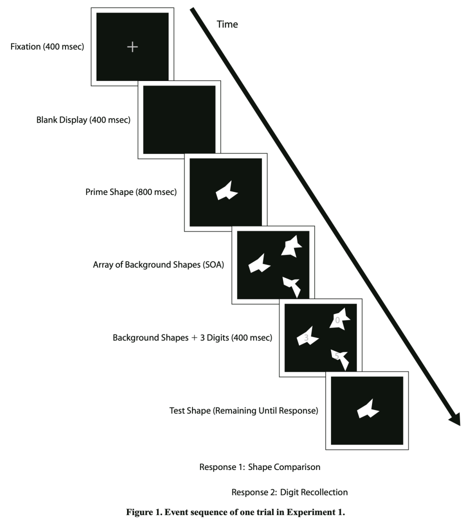


Figure 12.1: Schematic of the trial sequence in the Huang and Pashler (2007) study.

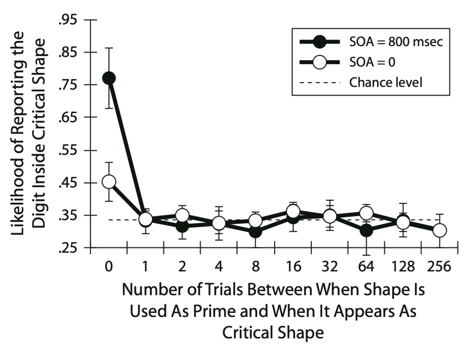


Figure 12.2: The result for '0' represents performance for the digit presented on the prime shape for that trial. It is much higher than "chance level" - the rate of performance if participants merely guessed a digit (the dotted line).

if the shape had instead been presented on a previous trial, performance was indistinguishable from the level one would find if the participant simply guessed the digit. This pattern of results shows that holding a shape in mind (the prime shape) caused participants to attend more to that shape (and the digit presented on it) subsequently.

- When there was no time from background shape onset to digit onset (0 SOA), performance was not nearly as good as when the background shapes were presented for 800 milliseconds before the digits came on. Why might that be?

12.1.2 Long-term memory

That the contents of long term meory guide attention is suggested by the Baader-Meinhof phenomenon, which is also known as the "frequency illusion."

Read this article about it.

- Why is it called the Baader-Meinhof phenomenon?

Zwicky theorized that an important cause of the phenomenon is an aspect of attention. As Huang and Pashler (2007) showed, once you have something in mind, you pay more attention to it when you encounter it. The Baader-Meinhof effect suggests that this phenomenon extends to things you recently learned about. Once you know about something, such as what the Baader-Meinhof gang was, it's more likely that you'll pay attention to and thus remember any individual encounter with it.

12.2 Apollo Robbins, misdirection, and pickpocketing

Magicians use their knowledge of how attention works to prevent us from noticing things that are right in front of us, in the real world, not just on a computer screen in a laboratory experiment video.

Watch Apollo Robbins talk about misdirection.

- What does he say about attention?

Now, watch again the sleight of hand trick he does in that video (I've set the video to begin playing at the critical part). How did the pen end up behind his ear?

Sometimes the movements that disappear objects in a magic trick are highly visible, but because our eyes and/or attention were in the wrong place, we don't notice them.

Watch this video of Apollo performing sleight of hand and pickpocketing

- How does Apollo manage to repeatedly place the coin on the woman's shoulder without her attending to her shoulder?

If the women could replay all the arm and hand movements of Apollo in her head, she could perhaps work out how the coin got on her shoulder. Unfortunately for them, they don't remember them all, because of a combination of the attentional bottleneck and the memory bottlenecks, not many of the movements got into memory. And when the movements were actually happening, Apollo subtly led the women to believe that the critical ones were unrelated to the coin.

- What role does eye-gaze cuing play in Apollo's video?

Optional reading: A NYT article about magic.

12.3 Four factors for managing attention

Apollo uses the following factors to manage the attention of people.

- Cuing
  - By directing attention away from where the coin actually was, Apollo prevented the women from focusing attention on the critical movements.
- Expectation
  - Apollo says things and makes gestures to cause viewers to expect something critical to occur in one location, so all eyes are on that location. This mostly happens via the viewers' top-down attention. Meanwhile, the critical action is happening somewhere else.
- Engagement in a task
  - When one is fully engaged in a task, bottom-up cues can be less effective in attracting attention. When one is not engaged in a task, one's attention can be widely distributed, ready to go to a location indicated by a cue. Magicians like Apollo sometimes ask people to concentrate on something (while the critical event for the trick is happening somewhere else). Also, the magicians may keep telling the participants things to ensure their mind doesn't wander in an unwanted direction.

- Encouragement to adopt a certain interpretation of an action
  - For example, Apollo getting the participants to get used to and expect him to touch their shoulder, until they interpret it as just a touch rather than that he might be putting something on their shoulder.

The last factor listed above is more of a high-level cognition phenomenon than the basic attentional phenomena that are the focus of this class. Specifically, the interpretation of the actions is being shaped by an implicit cognitive schema or narrative. The sections above on how the contents of memory influence attention are a precursor to this. Narratives and stories shape our interpretation of many things in the world, from peoples' body movements to news stories.

12.4 Inattentional blindness

In a previous unit of study, you may have already seen the inattentional blindness movies in the next section, but now you are equipped with some new ideas to apply to them. Keep the four factors that manage attention in mind.

Inattentional blindness refers to people not noticing something that's highly conspicuous if you know it might be coming, but is not conspicuous if you don't. This is different from the change blindness videos and the pickpocketing demonstration. In the change blindness videos, people knew there was a large change happening somewhere in the scene, but it still wasn't conspicuous. In Apollo Robbins' pickpocketing demonstration, people knew they were about to be pickpocketed .

People notice even fewer things if they aren't expecting them, or if they're given an unrelated task.

Watch this video for a demonstration and explanation (you can ignore the stuff about "smooth pursuit" at the end).

Take-aways:

- When you're focused on the task of evaluating the kick, you don't focus your attention on other things.
- The constant motion of the cheerleaders helped prevent the changes from grabbing your attention.
- Because noticing changes is capacity-limited, not having one's attention on the change meant it wasn't encoded or noticed.

The most classic demonstration of inattentional blindness involves the video created for an experiment by Simons and Chabris (1999). When watching the video, your task is to count the number of times the people in white shirts pass the basketball. Watch the video here.

Many of you may have seen that video before, which can create some expectation of what would happen and helped you notice what many people miss the first time they watch the video.

Figure 12.3: A shot from the movie used by Simons and Chabris (1999)

The associated experiment was done in 1999. The participants' task was to count the number of passes of either the white-shirted or black-shirted team. Doing the basketball task made their minds less likely to notice anything not basketball-related. It was as if they were "blind" to the gorilla in the scene - hence the name of the phenomenon, "inattentional blindness". The participants weren't expecting to be asked about anything besides their count of basketball passes, so all their attention was oriented towards the basketball task.

They were much more likely to notice the gorilla if they were counting the passes of the black team than if they were the white team. This is an example of the effect of feature selection. In one version of the experiment, only 42% noticed the gorilla if they were counting the passes of the white team, while 83% did if they were counting the passes of the black team!

12.4.1 The failure of bottom-up attention in inattentional blindness

These findings of inattentional blindness do not mean that people, when focused on a task or not expecting something, will never notice other things. Bottom-up attentional cues such as sudden flashes or sounds will still usually be effective in grabbing someone's attention. In most of the inattentional blindness videos and experiments, the displays are arranged so that the cues are not very effective. For example, in the gorilla video, there was lots of continual motion throughout the scene. If the gorilla had been the only thing moving in the scene, everyone would have noticed it, thanks to bottom-up attention.

12.4.2 The door swap

Perhaps the funniest inattentional blindness demonstration is the door swap. Watch this version created by Derren Brown.

In the Derren Brown video, it seems clear that some people definitely didn't notice the swap, and it's pretty amazing if even only a minority of people didn't notice the swap. To make the video even more amazing, the TV program producers may have edited the video to omit some cases of people who did notice and say something when the person changed. Because of such hijinks, you should always be suspicious of Youtube videos!

From the Simons and Levin (1998) article reporting the original scientific study, we knew that in the main study, seven of fifteen participants reported noticing that the person changed. That's not as incredible as the impression one gets from the Derren Brown video, but still it seems amazing that eight of fifteen people did *not* notice the person changed. The experimenters were worried that perhaps those people were too shocked, embarassed, or weirded out by the swap to say anything. For that reason, the experimenters took the people aside afterward to ask them whether they noticed anything strange. You can watch excerpts of the study here.

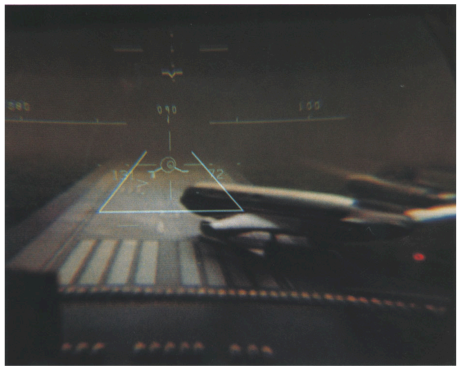
- Go back to the chapter on explaining two-pic change blindness (6). Which necessary processes for change detection do you think failed to occur?

- What factors do you think contributed to the participants not noticing the surprising event in the door study?

- What might be different between the people that did notice and the people who didn't notice?

12.4.3 In the air

Inattentional blindness also occurs in the real world. Landing a plane is a demanding task that involves monitoring several instruments. Haines (1991) found that when landing in foggy conditions in a flight stimulator, several pilots did not notice a large airplane obstructing their path on the runway, one that was clearly visible as they approached:



The pilots' strong engagement in the task of attending to their instruments probably contributed to some of them not noticing the plane.

Inattentional *deafness* is also a thing. People sometimes do not notice important sounds when they are not expecting them or they are highly engaged in a task. Read this article on pilots not hearing alarms.

A major factor contributing to cases of inattentional blindness and deafness has been the engagement of people in a task; the third factor mentioned in 12.3, which you should now review.

12.4.4 Exercises

- Why do bottom-up attentional processes fail to reliably draw attention to the gorilla in the inattentional blindness video?

Chapter 13

Divided attention

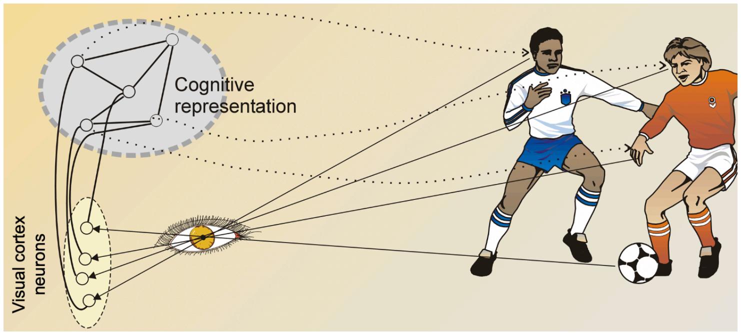
You already know about location selection. You can select an individual location with attention. You can also select multiple locations, and it seems you can select them simultaneously rather than your attention having to switch back and forth between them. However, as we'll see in the next section, we can only select a very limited number of stimuli at a time, which reflects that we have bottlenecks.

Given how narrow the bottlenecks in our brains are, a critical topic of research has been exactly what kind of processing goes on before signals reach the bottleneck. The study of how we pay attention to moving objects provides some important information about this.

You'll recall from Duncan's experiments (3) that higher-level processing is so limited that even identifying a simple feature of two objects yields much worse performance than identifying one. However, if all a person needs to do is keep your attention on multiple locations, without trying to identify things there simultaneously, a person can do pretty well with more than two objects.

13.1 Tracking moving objects

In the real world, objects of interest are often moving,or your eyes are moving, or both. When playing football, for example, while every point in the scene is processed by the first layers of neurons in cortex, certain bits are of special interest to continuously be aware of, like the location of the ball and of a defender. Thus, it pays to keep attention on those bits to ensure they are fully processed.



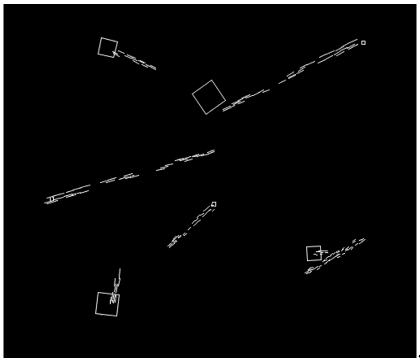
If all we had was feature selection and location selection, then if we were interested in monitoring a particular object, like the football, or a child playing in the ocean at a crowded beach, every time the object moved, we would have to find it all over again to attend to it. Finding it would likely require a visual search, which as you know from (9) can be time-consuming. Even if it only takes a third of a second, during that time a penalty kick traveling at 30 metres per second will have travelled a full ten metres.

Fortunately, attentional selection is able to follow a moving object (this was mentioned very briefly in 7). This is usually studied with what's known as the "multiple object tracking" procedure. In multiple object tracking, first a bunch of identical objects appear on the screen. Then, some of them are cued by briefly appearing in a different color or by flashing. The participant's task is to keep track of the cued objects as they move around. Try it here.

The task feels fairly natural. Indeed, it seems that once one selects an object with attention, if it starts moving, your attentional focus will tend to move along with it. It may be that what we have previously been referring to as location selection may best be thought of as **object selection**, because if an object in a selected location starts moving, it's hard to have your attention **not** follow along and stick to the location instead. Magicians like Apollo Robbins exploit this, moving their hand or another objects of interest with smooth gestures, knowing that if they do that, the viewer's attention is likely to come along.

The idea that attention selects objects has important implications. For attention to select objects suggests that the visual system must have processed the image preattentively into objects, otherwise attention may not have been able to select them. This is similar to the logic that Treisman used when arguing that processing of individual features occurred prior to the action of attention (*preattentively*) - otherwise, attention could not go straight to a stimulus with a particular color.

But what does the preattentive visual system consider an object, that attention can then select? This was investigated by VanMarle and Scholl (2003). In what they called a "substance" condition, objects sort of *poured* from one location to another.

.

People performed much more poorly when trying to track the pouring substances than they did when tracking the normal objects. Performance was similarly poor in a related 'slinky' condition. But the problem was not simply due to the nonrigidity or lack of cohesiveness of the objects, as shown by the fine performance in two control conditions.

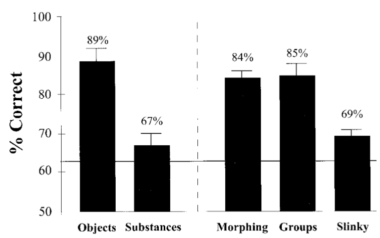


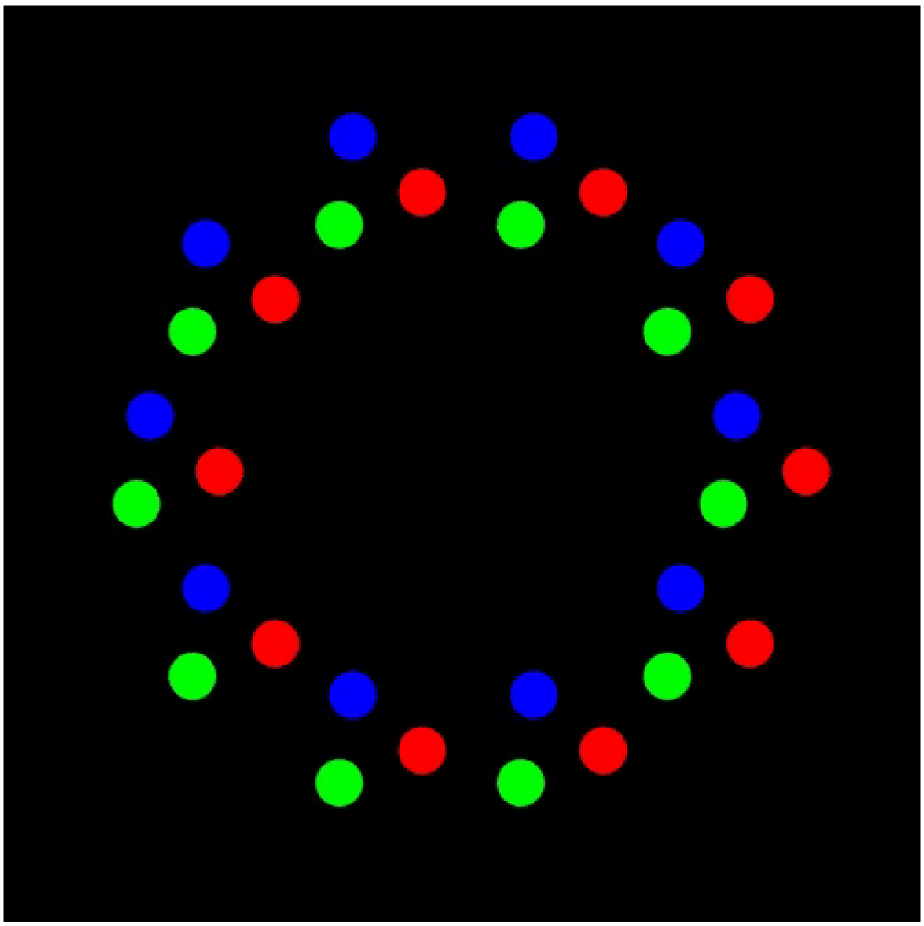
Figure 13.1: Mean tracking accuracy and standard error for each condition. The horizontal line running through the graph represents the performance that would be obtained if one tracked only one object and guessed on the others — 62.5% when tracking four items in a display of eight.

Experiments like this seem to be revealing how our visual system carves up the world into objects prior to the action of our attention. It also groups moving things based on how they move in relation to one another or are connected to each other. The following animation was taken from here.

You might be able to sometimes see an individual triplet of red, green, and blue as rotating around each other, and at other times see each color as being part of a global circle.

13.2 A temporal limit on tracking

The processes that allow attention to keep up with moving objects are still mysterious. But we do know something about the limits on those processes,



as illustrated here. These limits are much slower than those of basic motion perception (Verstraten et al., 2000; Holcombe and Chen, 2012).

13.3 Test yourself

You can test your multiple object tracking more quantitatively at the following site:

https://www.testmybrain.org/classroom/

Username: education

For the password, see Alex's Canvas module

The test first tests your forward digit span, then your backwards digit span, tests that you may have heard about in memory and intelligence lectures, and then measures speed thresholds for tracking.

Is there any relation between tracking ability and intelligence? We don't really know. But in the Testmybrain sample of thousands of people, those who did their degree in a STEM field (science, technology, engineering, or medicine) did better than those who did their degree in another field, such as arts or law. If you took a random STEM person and a random non-STEM person, the STEM person would have a 55% chance of having gotten a higher score on the MOT test. In other words, there are plenty of non-STEM people who have higher tracking performance than the average STEM person, but on average, STEM people do slightly better.

13.4 Exercises

- What does the fact that attention can select objects have to do with magic tricks?
- What have you learned about what processing occurs prior to a bottleneck?

Chapter 14

Bottleneck - memory (INCOMPLETE)

How long does it take to form a memory?

14.1 The attentional blink

ADD STIMULUS SEQUENCE PICTURE, which depends on which task I use, or use the Goodbourn Holcombe Keynotes for slides.

The attentional blink. As soon as you choose one item for memory encoding, people fail to do so for the second. In fact, they don't even seem to perceive the second item!

14.1.0.0.1 Recap Now, reflect back on the door video. What role might memory have played?

Canvas QUESTIONS: what would people have to remember in the door video to notice the change?

14.1.0.1 VIDEO: What can we expect for reading?

- If someone hands you a page of text to read, do you just stare at the middle of the page and process all the words at once? No, and there's two reasons for that
- refer back to the visual search stuff
- Reading seems to be fast but serial

Make some discussion questions for this video.

If we then add a bunch of O's everywhere, you can still detect it in parallel.

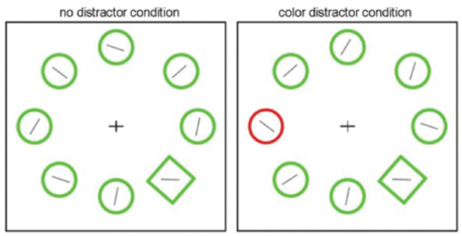
Chapter 15

Distraction

The "Three Wise Monkeys" is a 17th-century carving in the Tōshō-gū shrine in Nikkō, Japan and represents a Japanese Buddhist legend. Occasionally, all of us feel a bit like the monkey on the left, trying to ignore the sounds around us so that we can pay full attention to something we are looking at. Other times, we may want to concentrate on something we are listening to, in which case shutting our eyes like the monkey on the right might be a good idea.

It would be nice if we could use our brains to completely shut down hearing or vision temporarily and thereby avoid distractions. As we discussed in Chapter X, however, as creatures that evolved in a dangerous world, our brains are specifically evolved to *not* do that. Rather, our brains are constantly monitoring incoming signals from all our senses for things that might be a danger, or be rewarding. Sleep is a little different, but even during sleep a loud sound, a bright light, or a strong touch will pretty reliably wake people up. Babies sleep much more soundly and are harder to wake up, in part because they evolved to rely on adults for their defense and for their feeding, so they can afford to concentrate on brain development rather than monitoring the environment.

The top-down attention chapter CROSS-REFERENCE introduced the idea of *attentional capture*. The experiment discussed used these sorts of displays:



The task was to find the diamond and determine the orientation of the line within it, and people took longer to do that in the presence of an odd-colored distractor like the red circle at right.

In today's world, imminent physical threats are not as much of a concern. Neither is the sudden appearance of a food resource that we might never see again if we don't chase after it.

realistically, biggest obstacle to our studying

earplugs, hat

Noise-cancelling headphones Temptation - motivation. That's not really a topic for us. However, just like our brain evolved to respond to sensory information that may be important, our brain also evolved to not get too stuck on thinking about a single thing. This is because through most of our evolutionary history, it paid to switch tasks pretty frequently between looking for food, looking for other resources, attending to children, etc.

Figure 15.1: Horse. Credit: Alex Proimos from Sydney, Australia, CC BY 2.0, via Wikimedia Commons

A figure in Chapter blah illustrated that feature singletons can draw your attention. Does that happen even if you're concentrating on another task? NEED DEMO

- I can't work when the TV is in the background, even on silent
- auditory changing-state

15.1 Consequences for memory

VISION Morgan et al. (2020) - Dynamic visual noise eliminates the standard benefit of concrete over abstract words (but only in delayed free recall and delayed recognition tasks, Parker and Dagnall, 2009; see also Chubala, et al., 2018).

even quite modest changes in the visual texture of the scene – without any sort of accompanying threat or startle – have effects on a range of memory tasks (see Chubala et al. 2018, for an overview). By way of distraction each pixel on a display screen is randomly set either to black or white and every second a small random number of them changes state, a manipu- lation known as dynamic visual noise (see Quinn and McConnell 1996) while a short-term memory task is undertaken auditorily. Although task and distractor are in different modalities, mere exposure to the dynamic visual noise produces a reduction in memory, suggesting that some automatic processing of the visual display occurs, and that the result enters the cognitive system and proves disruptive. However, the results seem to vary across task type and task stage. Not all memory tasks are equally susceptible to dynamic visual noise. An early study (Quinn and McConnell 1996) found that dynamic visual noise produced no effect on rote memory, only when the words involved a visual imaging strategy for their retrieval. This result has been replicated a number of times (e.g. Andrade et al. 2002; Chubala et al. 2018; McConnell and Quinn 2000; Quinn and McConnell 1999). An analogous finding is that paired associate memory was vulnerable to dynamic visual noise, but serial recall was not (Ueno and Saito 2013). By contrast if irrelevant speech is presented while a visual memory task is undertaken the effects are strongest for serial recall, generally speaking serial pro- cessing tends to be most sensitive, with tasks that do not involve serial order showing mark- edly less sensitivity (e.g. Beaman and Jones 1997; Macken and Jones 2003; see also below).

AUDITION the more changes to the speech, also the greater the disruption (see for example, Beaman & Jones, 1997 and below)

15.2 Exercises

- What does the fact that attention can select objects have to do with magic tricks?
- What have you learned about what processing occurs prior to a bottleneck?

Chapter 16

Tutorial material

16.1 Visual search

Here's data from 1,725 PSYC2016 students for the visual search experiment you did.

Why is the slope of the green 'N' target data steep and the other graphs shallow?

16.2 Traffic lights

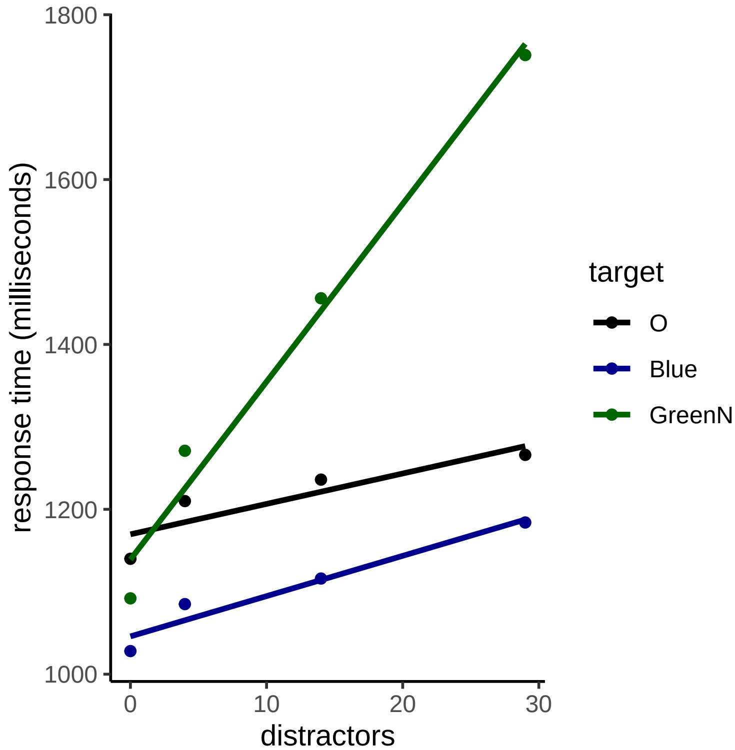


Figure 16.1: Results for a blue target (among differently-colored distractors), a green 'N' (among Ns of other colors and other letters that are sometimes green), and an 'O' (among other letters).

Figure 16.2: A Sydney suburban intersection. How many traffic lights do you see?

Chapter 17

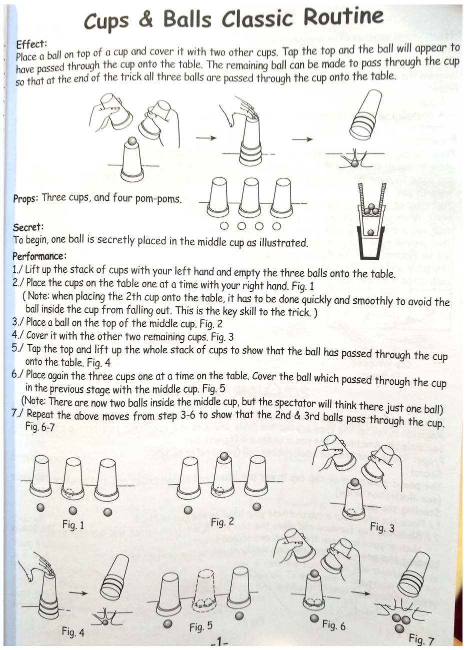
The cups and balls trick (INCOMPLETE)

VIDEO:

Figure 17.1: Teller doing some sleight of hand

(Youtube version)

- The role of assumptions



Cups & Balls Classic Routine

Effect:
Place a ball on top of a cup and cover it with two other cups. Tap the top and the ball will appear to have passed through the cup onto the table. The remaining ball can be made to pass through the cup so that at the end of the trick all three balls are passed through the cup onto the table.

Props: Three cups, and four pom-poms.

Secret:
To begin, one ball is secretly placed in the middle cup as illustrated.

Performance:
1./ Lift up the stack of cups with your left hand and empty the three balls onto the table.
2./ Place the cups on the table one at a time with your right hand. Fig. 1
(Note: when placing the 2th cup onto the table, it has to be done quickly and smoothly to avoid the ball inside the cup from falling out. This is the key skill to the trick.)
3./ Place a ball on the top of the middle cup. Fig. 2
4./ Cover it with the other two remaining cups. Fig. 3
5./ Tap the top and lift up the whole stack of cups to show that the ball has passed through the cup onto the table. Fig. 4
6./ Place again the three cups one at a time on the table. Cover the ball which passed through the cup in the previous stage with the middle cup. Fig. 5
(Note: There are now two balls inside the middle cup, but the spectator will think there just one ball)
7./ Repeat the above moves from step 3-6 to show that the 2nd & 3rd balls pass through the cup. Fig. 6-7

Fig. 1 Fig. 2 Fig. 3

Fig. 4 Fig. 5 Fig. 6 Fig. 7

-1-

17.1 Questions

When people view the trick for the first time, why does the appearance of the balls under the cups seem like magic?

Chapter 18

Dual task interference (INCOMPLETE)

and driving

Video of people driving real cars while taking Strayer's test and trying to operate voice controls

What role did the presence of a second task play in the gorilla video discussed at 12.4?

Canvas QUESTIONS: what task were people doing in the gorilla video? Would they haven noticed the gorilla without that task?

Simons and Chabris (1999)

18.0.0.1 VIDEO: Dual-task interference at fraction of a second scale

- PRP

REQUIRED READING

Hands-free cellphones not a solution to distracted driving

"The 27 seconds means a driver traveling 25 mph would cover the length of three football fields before regaining full attention." https://unews.utah.edu/up-to-27-seconds-of-inattention-after-talking-to-your-car-or-smart-phone/

Bibliography

Anderson, C. H., Van Essen, D. C., and Olshausen, B. A. (2005). Directed visual attention and the dynamic control of information flow. Elsevier, San Diego, California.

Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113(4):501–517.

Ericson, J. M., Kravitz, D. J., and Mitroff, S. R. (2017). Visual Search: You Are Who You Are (+ A Learning Curve). *Perception*, 46(12):1434–1441.

Friesen, C. K. and Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5(3):490–495.

Haines, R. F. (1991). A Breakdown in Simultaneous Information Processing. In Obrecht, G. and Stark, L. W., editors, *Presbyopia Research: From Molecular Biology to Visual Adaptation*, Perspectives in Vision Research, pages 171–175. Springer US, Boston, MA.

Holcombe, A. O. (2009). Seeing slow and seeing fast: two limits on perception. *Trends in cognitive sciences*, 13(5):216–21.

Holcombe, A. O. and Cavanagh, P. (2001). Early binding of feature pairs for visual perception. 4(2):127–128.

Holcombe, A. O. and Chen, W.-Y. (2012). Exhausting attentional tracking resources with a single fast-moving object. *Cognition*, 123(2).

Hood, B. M., Willen, J. D., and Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, 9(2):131–134. Publisher: SAGE Publications Sage CA: Los Angeles, CA.

Huang, L. and Pashler, H. (2007). Working memory and the guidance of visual attention: Consonance-driven orienting. *Psychonomic Bulletin & Review*, 14(1):148–153.

Kikuchi, Y., Senju, A., Tojo, Y., Osanai, H., and Hasegawa, T. (2009). Faces do not capture special attention in children with autism spectrum disorder: A change blindness study. *Child Development*, 80(5):1421–1433. Publisher: Wiley Online Library.

Külpe, O. (1893). *Outlines of psychology.* Sonnenschein, London, 3rd edition. E. B. Titchener, Trans.

Noë, A., Pessoa, L., and Thompson, E. (2000). Beyond the grand illusion: What change blindness really teaches us about vision. *Visual cognition*, 7(1-3):93–106. Publisher: Taylor & Francis.

O'Regan, J. K., Rensink, R. A., and Clark, J. J. (1999). Change-blindness as a result of 'mudsplashes'. *Nature*, 398(6722):34–34.

Posner, M. I., Snyder, C. R. R., and Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109(2):160–174.

Rensink, R. (2000). Visual Search for Change: A Probe into the Nature of Attentional Processing. *Visual Cognition*, 7(1):345–376.

Rigby, S. N., Stoesz, B. M., and Jakobson, L. S. (2016). Gaze patterns during scene processing in typical adults and adults with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 25:24–36.

Rösler, L., End, A., and Gamer, M. (2017). Orienting towards social features in naturalistic scenes is reflexive. *PLOS ONE*, 12(7):e0182037. Publisher: Public Library of Science.

Simons, D. J. and Chabris, C. F. (1999). Gorillas in our midst: sustained inattentional blindness for dynamic events. *Perception*, 28(9):1059–1074.

Simons, D. J. and Levin, D. T. (1998). Failure to detect changes to people during a real-world interaction. *Psychonomic Bulletin & Review*, 5(4):644–649. Publisher: Springer.

Sàenz, M., Buraĉas, G. T., and Boynton, G. M. (2003). Global feature-based attention for motion and color. *Vision research*, 43(6):629–37.

Theeuwes, J. (2010). Top–down and bottom–up control of visual selection. *Acta Psychologica*, 135(2):77–99.

Theeuwes, J. and Failing, M. (2020). Attentional Selection: Top-Down, Bottom-Up and History-Based Biases. ISBN: 9781108891288 9781108813068 Library Catalog: www.cambridge.org Publisher: Cambridge University Press.

VanMarle, K. and Scholl, B. J. (2003). Attentive tracking of objects versus substances. *Psychological Science*, 14(5):498–504.

Verstraten, F. A. J., Cavanagh, P., and Labianca, A. (2000). Limits of attentive tracking reveal temporal properties of attention. *Vision Research*, 40(26):3651–3664.

White, A. L., Palmer, J., and Boynton, G. M. (2018). Evidence of serial processing in visual word recognition. *Psychological science*, 29(7):1062–1071.