From Attention to Distraction

Alex O. Holcombe

Updated on Fri 13 Oct 09:00 AEDT 2023

Contents

List of Figures

Preface

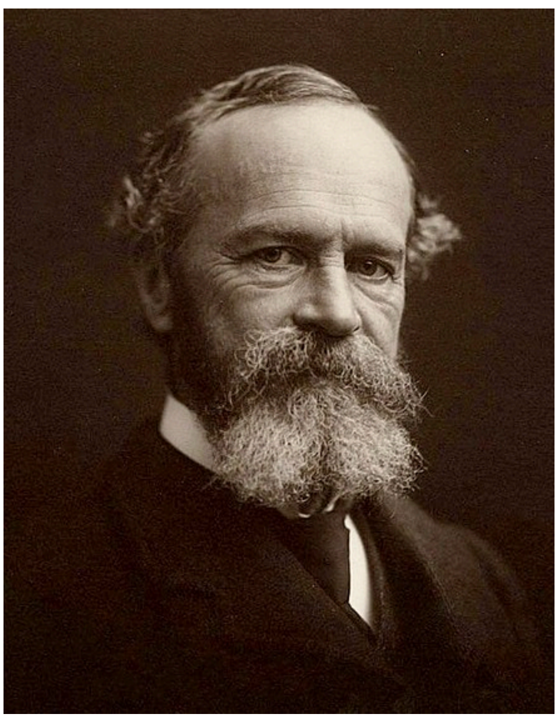
This is a mini-textbook on attention. It was written for PSYC2016 at the University of Sydney, but is likely to be suitable for everyone. Please let me know about any typos or unclear bits!

You can contact me (he/him) via email (alex.holcombe@sydney.edu.au), Mastodon, or twitter.

You can read this mini-text here on the web, download a PDF file, or an e-book version, for your Kindle or other e-book reader. However, the web version is the one you should rely on - some movies and possibly some images may be missing from the PDF and e-book files.

Chapter 1

Background



> 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

That was William James, one of the founders of modern psychology, writing in 1890, which was the very beginning of the scientific study of psychology. In the quote, he appeals to common sense. Common sense was relevant because some things about attention have been understood for a very long time, even if not studied scientifically.

The Cardsharps was painted by Caravaggio more than four hundred years ago. It depicts a young man getting duped in a card game. The cardsharp in front is assisted by an older man who looks over the young softie's shoulder. The old cardsharp signals something about the cards to the younger cardsharp, who is reaching behind his back to pull a new card from his breeches.

The painting tells a compelling story. Clearly people of that time knew how to manipulate the attention of others so that they would notice some things and not others. And I don't mean just the cardsharps; in order to create paintings that would direct the viewer's attention in the way he intended, painters also must have understood some things about attention.

Much of the knowledge of attention possessed by painters, as well as by con men like those depicted in the painting, was likely implicit rather than as spelt out as we expect for a scientific theory. It was only in the twentieth century that scientists began doing extensive experimentation on attention, and even then attention was studied very little relative to memory, perception, or cognition.

Around the 1970s, that began to change, but there is still much for science to learn about attention. A good starting point continues to be people's everyday understanding of attention - what is sometimes called *folk psychology*. Some aspects of the folk psychological understanding of attention are so old that they have become embedded in our language. Here are some examples:

- Please “pay attention!”
  - Implies that one has to choose something (often the teacher) to fully process its signals.
- “Sorry, I wasn’t listening”
  - In English, there’s a difference between hearing something and listening to it. Not listening means although it may have been processed by our ears, we either didn’t experience it or didn’t retain it.
- “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 do you think people might mean when they say that rather than choosing to instead say “I didn’t see that.”?

Failures to notice things explain a substantial proportion of accidents such as car crashes. But **why** do we have to pay attention to comprehend or retain some information? That is the subject of two coming chapters (3, 4).

If you took PSYC1 at the University of Sydney, you already heard Caleb Owens’ “Cognitive Processes 2” lectures on visual attention, which were related to what you will learn here. Caleb’s lectures 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.

This year has its own list of learning outcomes (2).

We will review material from Caleb's lecture, bring new ideas to bear on it, and add more material. In the parts that review what Caleb's points, we 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

After reading this mini-text and participating in the lectures, 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
11. Explain why learning is important for determining what attracts our attention

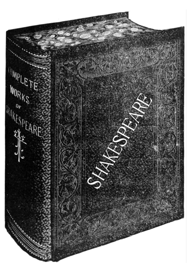
These are high-level outcomes. You'll see more detailed points in each chapter, often with questions at the end to check your understanding of individual points.

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, one form of which is called RAM. For example, the laptop that I am writing this on has 16 GB of RAM. That is enough to store a lot of information- 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.

Imagine you had the complete works of William Shakespeare in your computer's RAM, 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. To do this task, we'd 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, as it does letters one-by-one.

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

3.2 Brains

The architecture of brains has some similarities with that of computers. In particular, brains have bottlenecks.

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

The associated processing results in perception. 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 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.

Figure 3.1: Sensory information entering the brain eventually hits bottlenecks.

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 3.2: A MIG-21R fighter

By climbing the stairs 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 stages 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.

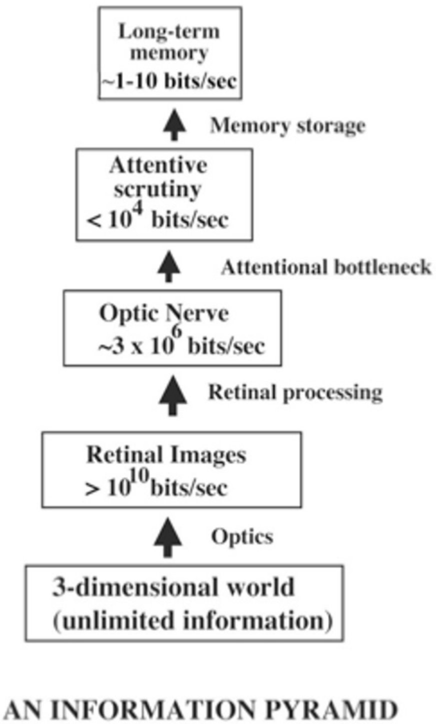
Forgetting about the brain for the moment, let's consider your eyes. They are a lot like an old analog camera. They don't know what they are looking at. You need a brain, with millions of neurons working together, to process the eye's signals and identify an object. And our brain isn't big enough to 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:



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 3.3: **A New Year's Eve party that I attended 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; remind yourself of what this class focuses on 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:

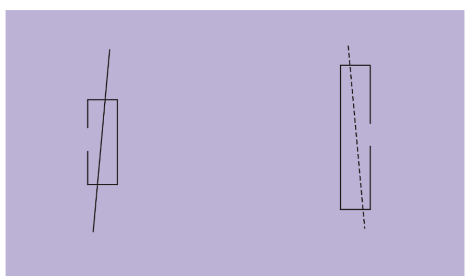
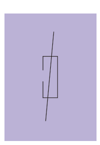


Figure 3.4: Two displays. On the left, a **clockwise-tilted** line on top of a box with a left gap.

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



This was described as 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 a reasonable amount of time, everybody would get the judgments right regardless. But he found that if he flashed the objects for a tenth of a second or less, the task became difficult and participants made frequent errors.

The rate of errors participants made for the different tasks was illuminating. 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 judgments about different objects than for judgments about one object.

These results suggest that 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.

3.5 Reconciling a bottleneck with our visual experience

While it's difficult to know whether processing is so capacity-limited that it is truly one-by-one, subsequent results have supported Duncan's main idea that many judgments involve processing that is far from unlimited-capacity and parallel. And as alluded to above (3.3), Duncan's idea comports with the feeling that it's hard to think about more than one thing at a time.

For visual experience, however, in some ways it *doesn't* feel like we can only process one thing at a time. As I write these words, for example, I seem to be experiencing the whole visual field simultaneously. It doesn't feel like I can only experience one object at a time. Yet when we carefully test people on what they can report about multiple objects, if we don't give them time to move their attention around, the results suggest much processing is limited to only a few objects.

The discrepancy between what people think they experience and what information they can report is sometimes referred to as the puzzle of phenomenal experience. The simultaneous experience of the entire visual field is labelled "phenomenal consciousness" and the more limited ability to report information about the contents of visual experience has been labelled "access consciousness" (Block, 2011).

Researchers can't yet fully explain the difference between phenomenal consciousness and access consciousness. However, they have discovered various visual processes that are *not* very capacity-limited, which can help account for some aspects of phenomenal consciousness. In the following chapters, we will go through some of these processes.

3.6 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 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?

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.

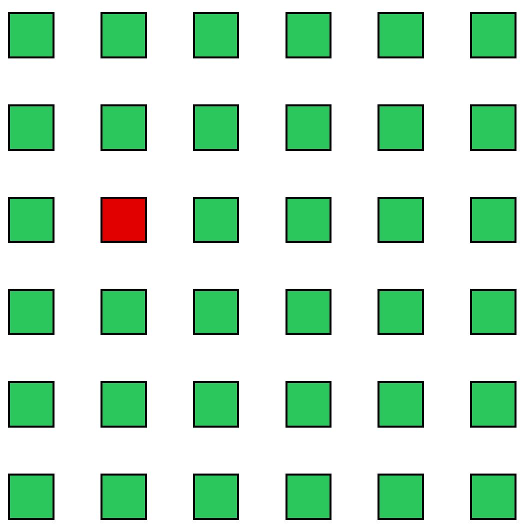
Top-down attention is typically voluntary, and thus guided by your expectations and desires, as represented by this inspector intentionally scrutinising individual bits of a crime scene.

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**.

The neural processing that underlies bottom-up attention is sometimes called **salience** computation. Color-tuned neurons respond to the color features, with different neurons responsible for each location of the visual field. Neighboring colors are then compared to each other, with the aim of strengthening the neural response to colors that are different than most of their neighbors. This is achieved partly through neural inhibition, where neurons that respond to the same feature inhibit each other's activity. In the example above, then, green-responding neurons inhibit green-responding neurons nearby. Because the red patch has no red neighbors, nothing inhibits it, so the corresponding neural activity is higher for it than to the green elements.

An object with unique motion direction can also summon attention, as you can see here, which probably works in the same way.

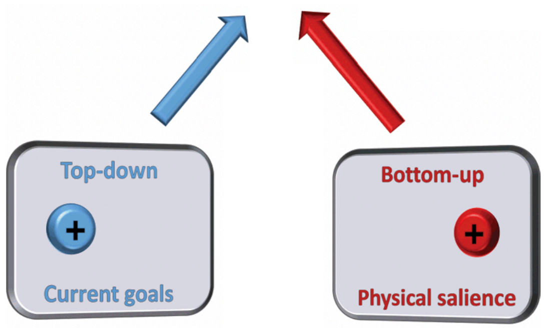
When choosing a car to buy, some people deliberately pick an unusual color because they know that when they go shopping, if they forget where they parked their car, they will have little trouble finding it. A pink car is salient or "sticks out" conspicuously even in a sea of other cars, if those cars have the more typical paint jobs of black, white, grey, and dark colors.

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).

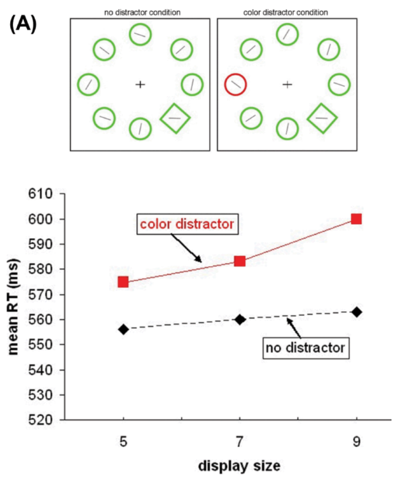
In the experiment, participants searched for a green diamond presented among a variable number of circles and had to respond to the orientation (horizontal or vertical) of the line segment presented within the diamond shape. So, their task was to find the diamond and pay attention to only it.

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 presented, the longer it took 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.

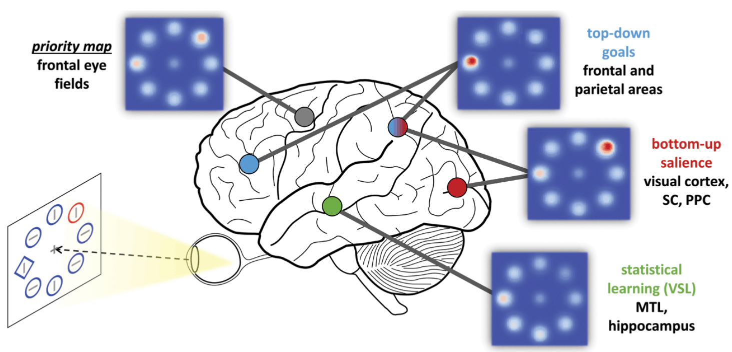
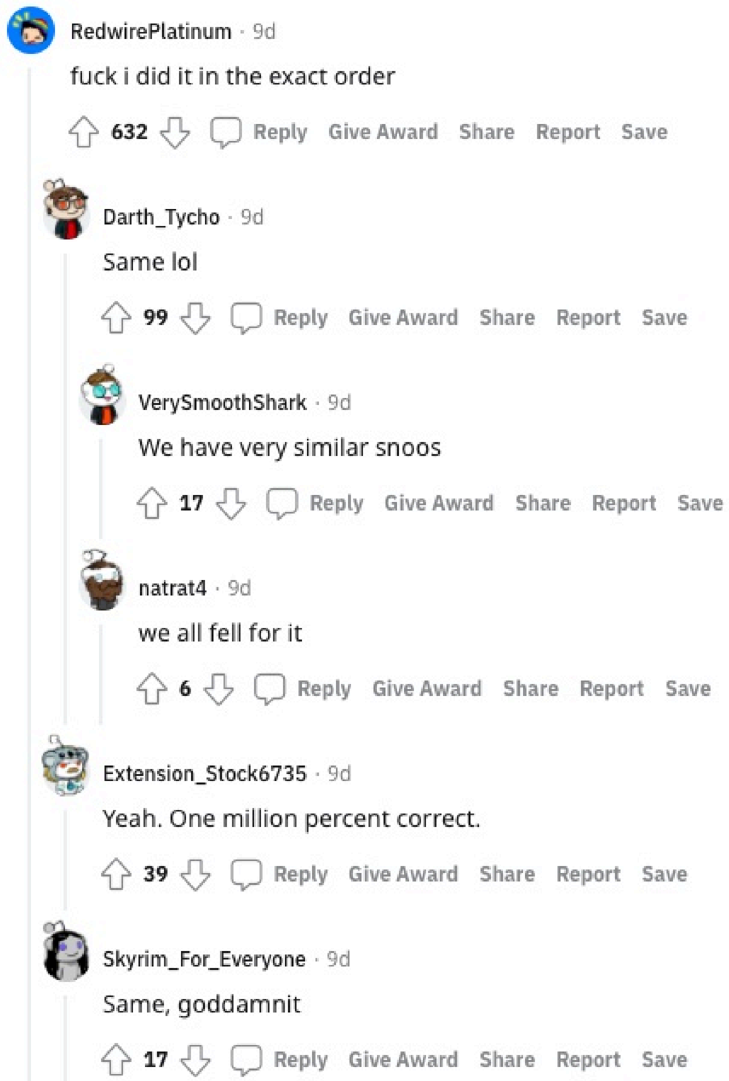


Figure 5.1: A schematic, created by Theeuwes and Failing (2020), indicating brain areas that mediate bottom-up attention, top-down attention, and a priority map.

5.2 Bottom-up and top-down attention when reading

Web designers, game designers, and graphic designers all need to have some understanding of what in a display attracts attention. One of the principles that they learn is that unique colors attract attention. But there's more to it than that.

The above image illustrates some of the factors that affect the attention of people when they look at text. The content of the image claims to successfully predict the order in which you will read the different printed phrases. I don't know that there's been any scientific test of this, but the comments on Reddit provide some anecdata suggesting that it did work for many people:



How was the designer of the image able to fairly accurately predict reading order? Some of that comes from the principles of bottom-up attention. The "First, you read this" text is an odd color, which you know will attract attention. Two other factors that help are that it has a large font and it occupies a central position. For various sorts of objects, all other things being equal, people will tend to look at the center image. This may be particularly true of designed images with a frame like this one, because we come to learn that designers often position the most important information in the center of a frame.

"Then you will read this" is actually slightly larger than the "First, you read this" text, but because it is at the bottom of the image and it is not in an odd color, the designer could safely bet that most people wouldn't look at it first. Why is it (often) the next text looked at? Its size certainly helps, but another factor is that it is below the text that most people read first. Once you put a person into reading mode, they tend to shift their attention according to reading order: in English, from left to right and from top to bottom.

You'd have to expect that when people open the pages of a book, or land on a webpage with a lot of paragraphs, laid out like prose. To make sense of a paragraph, we of course will read the words from left to right and the lines from top to bottom. I was interested in whether participants would do this even when they were looking at just two individual letters, rather than a bunch of text.

In experiments a few years ago, we investigated this by testing several hundred students who were taking PSYC1002 at the University of Sydney :)
We repeatedly flashed two widely-spaced letters at the students and asked them to identify both of them. If the students attended to one location before another, we reasoned, their performance would be higher on the letter that occupied the position attended to first.

One configuration we tested was much like the below, with one letter placed

g

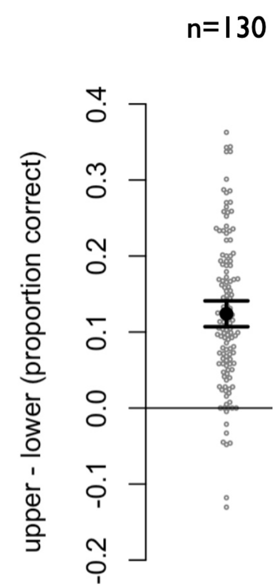
h

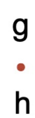
above another, like this:

People are extremely good at reading letters, so we knew that if the students were to ever report a letter incorrectly, we would have to flash the two letters extremely briefly. But even when the letters were flashed for only about 2 hundredths of a second, we found that students still performed above 90% correct! To reduce performance enough, down to around 75% correct, we had to present the two letters very briefly, with low luminance contrast, and immediately afterward flash a bright "mask" that helps limit the amount of time one's brain has to process the letters. In slow motion, then, a trial looked like this movie.

Prior to the presentation of the two letters, the students were looking directly at the center of the screen. The presentation was so brief that they did not have time to move their attention to either location, so any difference in performance for the upper versus lower location would be a result of covert attention (assuming that both parts of the visual field are equally good, which had been confirmed).

To examine whether there was any difference between the top and bottom positions, we subtracted each student's performance on the bottom letter from that of the upper letter. Each dot in the below plot represents the difference score for one of the one hundred and thirty first-year students.

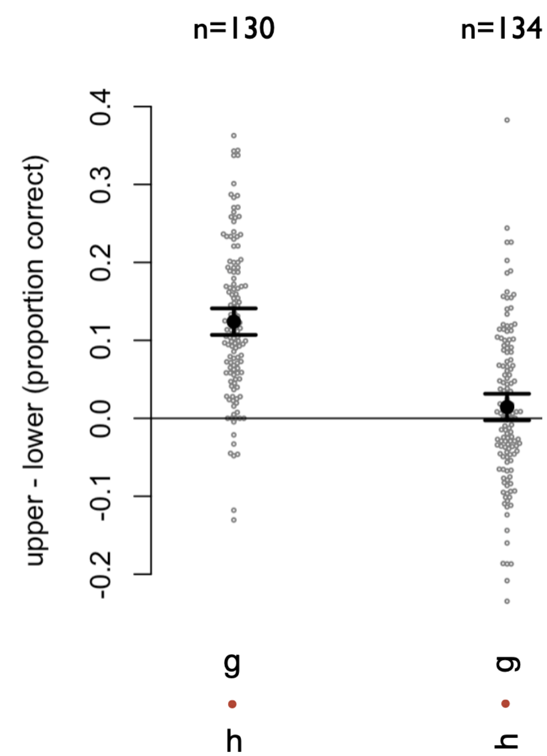




The average difference score of the students is 0.13, meaning that they were thirteen percentage points more accurate at reporting the top letter compared to the bottom letter. Notice that this upper-letter bias was not true of everyone - eight students actually identified the bottom letter correctly more often than the top letter. It appears, however, that the overwhelming majority of students did have an upper bias with this display.

A second condition of our experiment found evidence that this upper bias reflects an automatic appraisal by the mind of what the correct order should be for reading text. In this condition of the experiment, a different group of the students had to identify two letters that were rotated anti-clockwise by 90 degrees, so that they faced upward:

The question was whether this would affect the upper bias documented in the basic condition. With the two letters facing upward, what you might call the implicit reading order should favor the bottom letter. That is, if you were to rotate the page of a book in this way, you would then start from the bottom and read toward the top.



The above plot shows the results for the two conditions side-by-side. You can see that the average top bias found previously is much smaller in the facing-up condition. Indeed, plenty of students now showed a bottom bias rather than a top bias.

Remember, these letters were presented so briefly that there was no time for anyone to move their eyes. It was so brief, in fact, that the students are very unlikely to have even been able to move their attention while the letters were being presented. But the students seem to have allocated more of their attention to one part of the display to another, which led them to perform better for the letter in that location.

This finding is an example of how biases in our attention that we often aren't even aware of contribute to what we notice and what we can report. This particular bias is not a bottom-up effect like an odd feature such as color or size. Instead, this attentional bias is one that we *learn* from our experience reading English text. In other experiments, we found that English/Arabic bilinguals had different biases when asked to report two briefly-presented English letters than when they were asked to report two briefly-presented Arabic letters (Ransley et al., 2018). As learned biases, these phenomena are more like top-down attention than bottom-up.

5.3 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?
- What are interrupt signals?

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.

Wasn't that amazing? In some ways, we humans are a lot dumber than we think! People usually don't notice any of the several changes made to the scene. We'll circle back to this Whodunnit movie, but first let's talk about a case that may be even more striking.

6.1 Blindness for gradual changes

In the Whodunnit movie, the changes occurred off-screen, when the camera was focused tightly on the detective on the left. One might expect that if the changes happened right in front of your eyes, you would notice them.

Amazingly, we fail to notice changes right in front of us, too, if they happen very gradually, as illustrated in this movie.

6.1.1 The "grand illusion of visual experience"

Most people are surprised by the blindness for changes in the gradual-change movie. Many researchers were very surprised, too, and some concluded 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 might be involved in detecting the change of an object or part of a scene:

1. Internal representations of the object before and after the change.
2. A process that compares compares the before representation to the after.
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. The internal representation of the object is too low in fidelity. However, this 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. This device is wired such that if it detects motion, the carport light comes on. There is nothing fancy about the processing within it - not much circuitry is required for it to work.

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

Firing of those flicker/motion detectors can call attention to a location. This is an instance of bottom-up attention (5). 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 summary, 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 will stimulate our motion or flicker detectors, which in many circumstances will call attention to the associated location.

These facts about the brains of humans and other animals are one reason that animals stay very still when they are worried about predators. Thanks to their camouflage, many animals 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 of the associated locations attention should go to. O'Regan et al. (1999) have demonstrated how this can enable changing blindness with a display feature that they called "mud 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 trigger your motion detectors and thereby 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.

Now you have all the knowledge you need to understand the full explanation, in the next section, of why people take a long time to find the change in the classic change blindness animations.

6.3 Two-picture change blindness

Many of you have seen animations like that of the boat scene or this Paris scene, which 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 chocolate biscuits and the ice cream is analogous to the blank screen… yum.

Here is a schematic of the timeline of a blank screen sandwich.

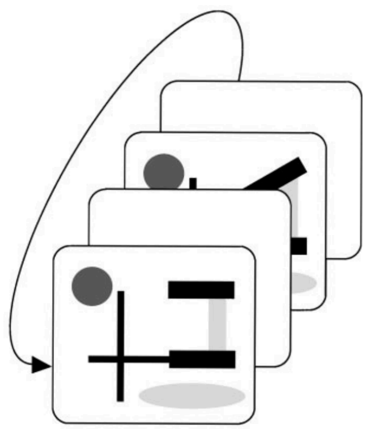


Figure 6.1: **Two pictures of a scene are alternated, with a blank screen between them.**

The 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 flicker everywhere again when the second scene comes on.

When the blank screen is removed from a blank screen sandwich, the scene change is conspicuous; this animation is an example. In it, 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.

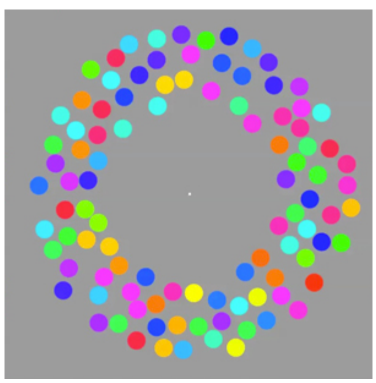
Without the blank screen, the only location of flicker was the location of the changing object. The flicker called your attention to that location. *With* the blank screen, there's flicker everywhere, so there is no indication of which of the many locations contains the change.

Then how do people ever find the change? We will return to this topic at the end of Chapter 9.

6.4 When change is everywhere but you still can't see it

A blank-screen sandwich isn't the only way to prevent motion or flicker signals from signalling to the brain that an object has changed.

Adding motion to an object can make the fact that the object is changing in other ways inconspicuous. In a phenomenon that Suchow and Alvarez (2011) dubbed "motion silencing", in certain circumstances, even when every dot in a display is changing color, those changes aren't noticeable if the color-changing dots are moving.

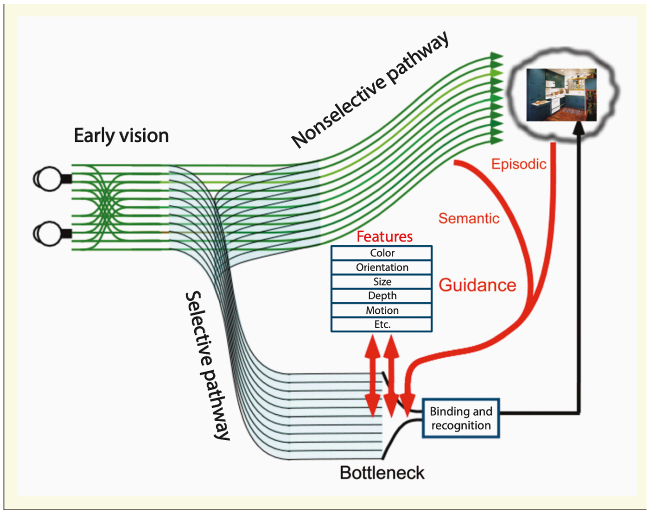


View the motion silencing animation here.

Jun Saiki & I showed that a key requirement for motion silencing to work is that the amount of each color stays approximately the same before and after the changes (Saiki and Holcombe, 2012). As an extreme example, if 80% of the dots changed to green, participants noticed that change. This shows that people continuously monitor, to some extent, the distribution of colors present, and a large change in this distribution is noticed. This is sometimes referred to as a representation of the scene statistics.

Another thing that can break change blindness are changes to the overall meaning of the scene. For example, if you were viewing a scene of a living room, and suddenly the furniture was re-arranged to resemble a lecture theatre, you'd be very likely to notice that. Like the distribution of colors, the "gist" or overall meaning of a scene seems to be continuously encoded and monitored for change.

To represent the existence of the less spatially-selective, global gist processing that occurs, Wolfe et al. (2011) created the figure below.



The non-selective pathway is more difficult to study and there's still a lot about it that researchers don't yet understand. However, we know that our understanding of the meaning or gist of a scene helps guide attention, for example if looking for keys in our car we will look around the ignition and around the seat, places the keys are most likely to end up.

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 splashes 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 for them to be fully processed. This is assisted by three abilities that we have:

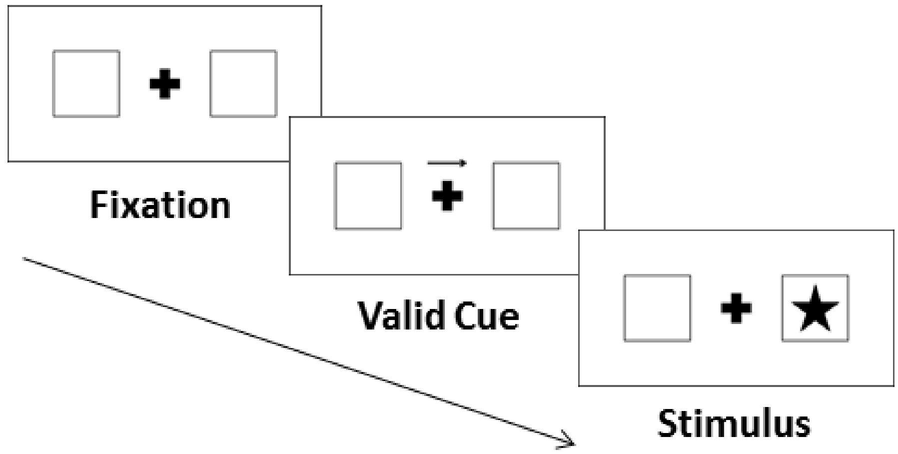
- Location selection
- Feature (e.g., color) selection
- Object selection and tracking

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 (overt attention).

But sometimes, people attend to multiple locations simultaneously. In that situation, their eyes have to point at one location, while they simultaneously attend to other locations. In previous units of study you've taken, location selection may have been the only kind of attention discussed. For example, in classic Posner 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.



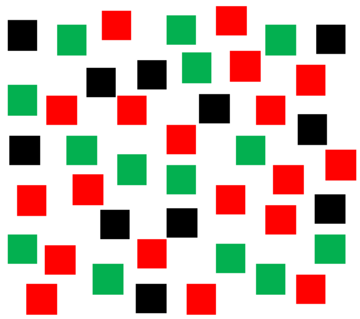
The schematic illustrates a shape identification experiment. The arrow cues the participant to covertly attend to the right while keeping their eyes on the center. The covert shift of attention results in participants identifying the shape on the cued side (the star) more quickly than if the arrow had directed the participant's attention somewhere else. We'll revisit location cuing in some new contexts in 12.

7.2 Feature selection

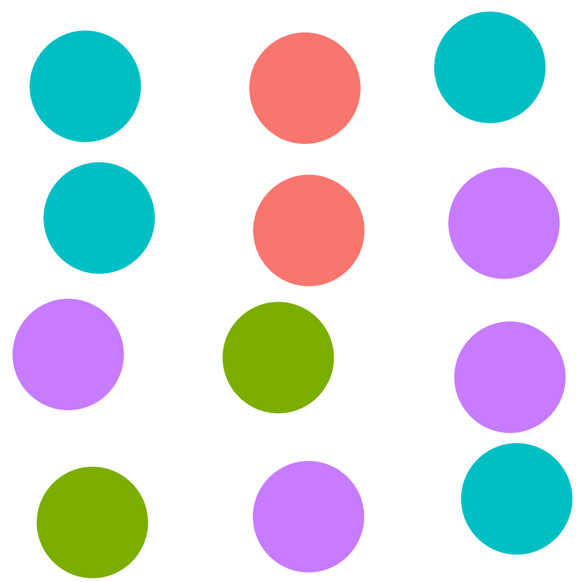
You might think that in order to attend to something you're looking for, you first have to find it by moving your attention all around (location selection) to evaluate whether each region of the scene contains your target. That is true in some cases, but not all, thanks to the power of **feature selection**. Feature selection allows your attention to go straight to certain sorts of targets, as we'll see.

Feature attention allows you to select an individual feature value, for certain features. One such feature is color.

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

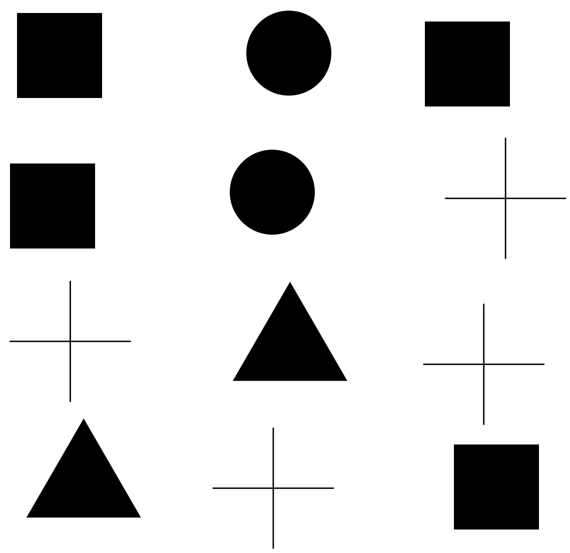
 Next, try concentrating on all the black objects. Again, you can use your power of feature selection to do so. Similarly with the green - try it.

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

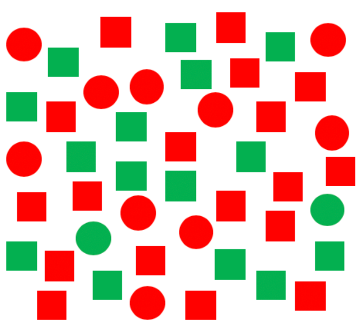


thinking of squares, circles, triangles, or pluses.

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

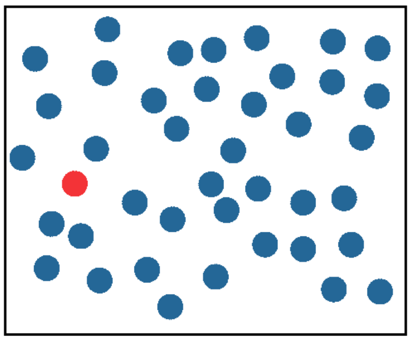
7.3 Feature combination selection?

However, selecting a combination of features does not work as well. You can think "green circles", but unless your attention lands on them right away due to luck, it will take longer to find the green circles.

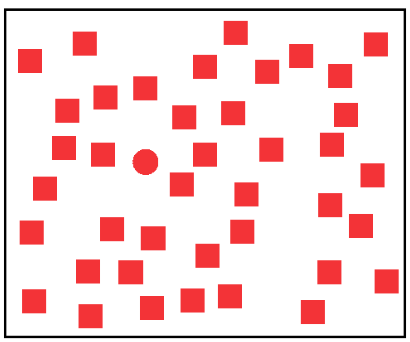


Your brain is unable to pick out a stimulus when finding it depends on a combination of features rather than a single feature.

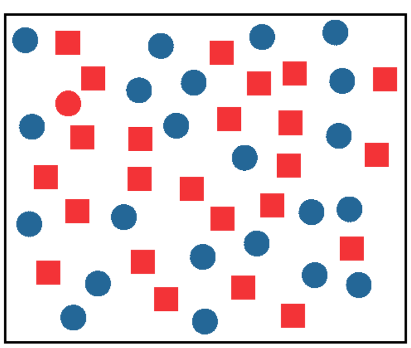
In the below three images, the target is a red circle in each case. In *b* it is easy to find the red circle, because it is defined by a single feature (color). in *d* it is also easy to find a red circle, because it is again defined by a single feature (simple shape). In *f*, however, it usually takes a longer to find the red circle. Selection of a combination of features is not effective.



(b)



(d)



(f)

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 to the instances of a feature that one decides to select. Researchers don't fully understand how this works, but as will be decribed in 9, a theory proposed by Anne Treisman claims 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.

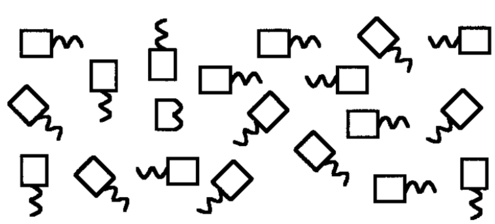
7.5 Complex shapes

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.

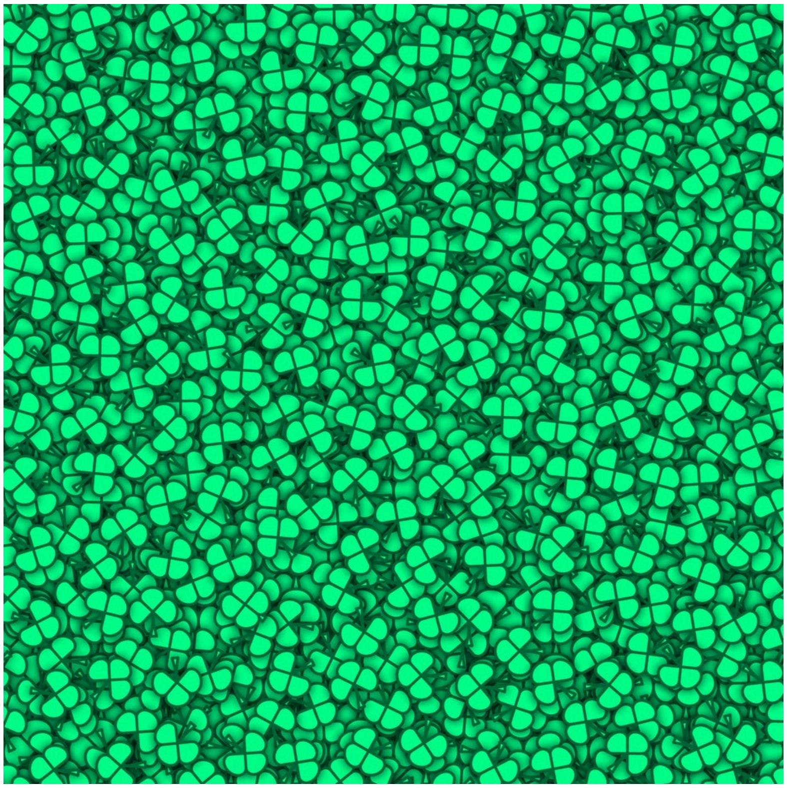
Figure 7.1: The target (left), which is tail-less, and a distractor (right).

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



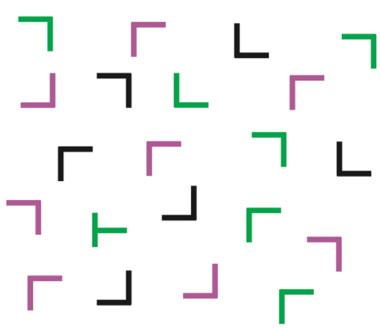
Finding the object without a squiggly tail takes, on average, much longer than in the simple search cases described in previous sections.

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 individual leaves).

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. 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

Identifying pairs of features

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.

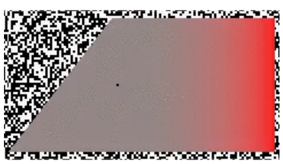


Figure 8.1: Try to perceive whether the red color is paired with leftward tilt or rightward tilt.

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

What's going on in the above 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.

Speeding up the alternation of the two frames can reveal the time it takes for our brain to combine features (Holcombe and Cavanagh, 2001).

If the movie above isn't viewable, watch it here. At this fast rate, note that 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.

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

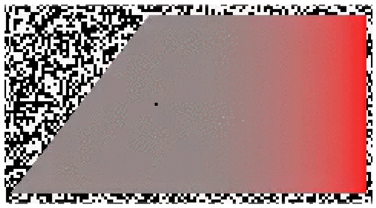


Figure 8.2: Try to perceive whether the red color is paired with leftward tilt or rightward title.

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 may not have enough time. If the animation doesn’t show, try watching it here.

Feature pairing 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’s difficult in the middle row, where the speed is slightly faster (Holcombe, 2009).

In the previous chapter (7), you learned that our brain has a process for attentional selection of an individual feature value (such as blue), but not for combinations of features. Consistent with this, the animations of this chapter illustrate that identification of feature combinations is time-consuming. This suggests that identifying features requires capacity-limited (post-bottleneck) processing. 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 does 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

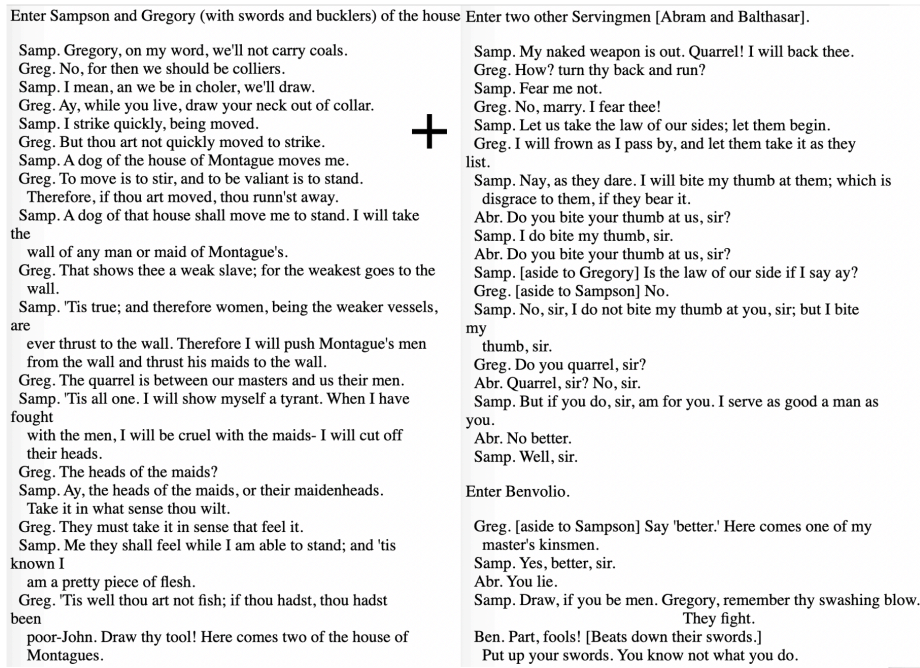
From time to time, we all need to find things. Rummaging through our closet for a particular shirt, or wandering about the house trying to find our keys, or for some of us, groping about for our spectacles that we know we put down *somewhere* around here.

Our search performance patterns can reveal aspects of the bottlenecks in mental processing. Slow search may indicate 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, if one is to evaluate all the rooms of the house for the presence of the keys, one has to visit each room.

Sometimes, even though something is right in front of our face and we have our glasses or contact lenses on, the sensory signals still 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.

Did you find 'wilt' yet? To find it, your eyes have to move back and forth. The task is impossible to do without moving your eyes ('wilt' is about 3/4 of the way down the left page). The main reason you have to move your eyes 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 experience 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*.

Thus, not only are overt attentional shifts (eye movements) often made when you want to attentionally select a region of space, but also sometimes *covert* attention isn't sufficient - you have to move your eyes to have any chance of seeing certain things well enough to know what they are.

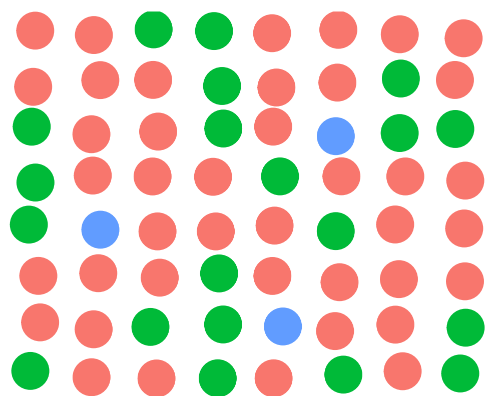
9.1 Information overload

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 them all just as quickly as when it is given only 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 well, then of course the brain wouldn't process it sufficiently even if it didn't have a bottleneck.

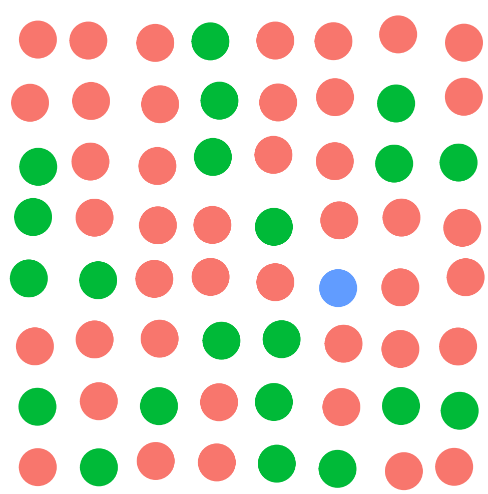
One of the tasks psychologists have used for this is called "visual search". In a visual search experiment, people are shown a display with a particular number of stimuli and asked to find a target. This is discussed this in first-year psychology. The next section is, in part, a review of that.

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 lots of object, but only one 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 major bottlenecks of the brain.

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

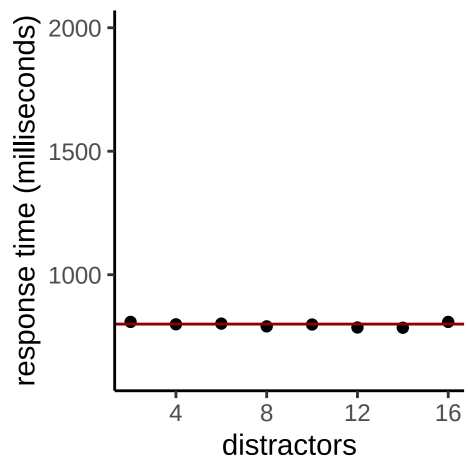


Figure 9.2: Time to find a lone blue circle among red and green circles

The diagram below provides a basic idea of the processing stages involved. 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.

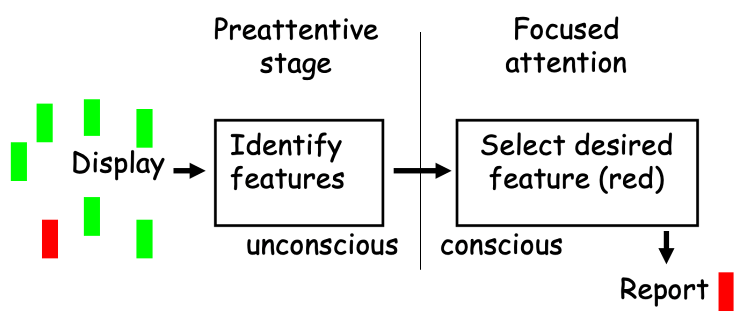


Figure 9.3: A schematic of the processing stages involved in searching for red.

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

Parallel search doesn't happen in most cases for *combinations* of individual features. Instead, there is a bottleneck. To put that in context, first let's remind ourselves of aspects of parallel versus serial processing.

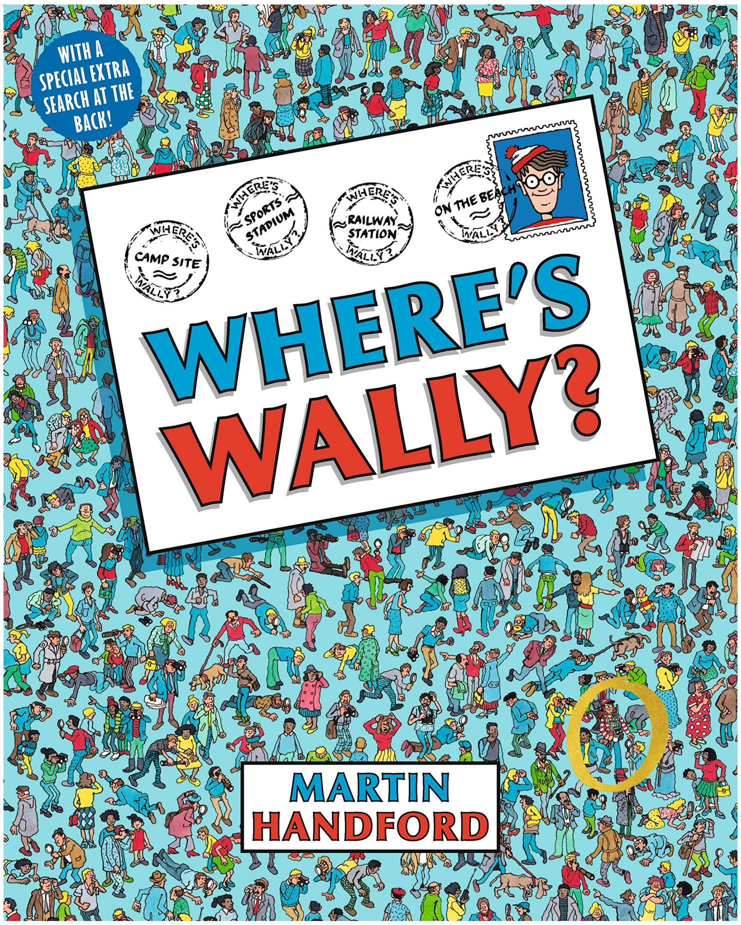
Imagine you were in an art installation where the artist had hung many speakers from the ceiling, and each speaker played a different person's voice, each telling a different story. That's pretty weird, but is precisely the situation I was in one day when I visited a museum in Havana, Cuba. What I heard sounded like an incoherent jumble. I couldn't follow 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 using overt attention.

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 that day in Havana, having to select individual locations to evaluate an aspect of what is present - specifically, the combination of features there.

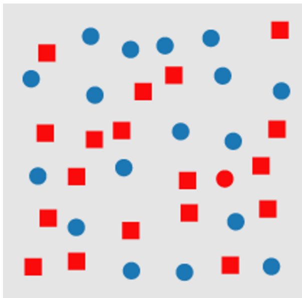
9.4 Combinations of features - serial search

As children, many of you will have had a book from the Where's Wally? series.



The game on each page was to search for the Wally character. Sometimes it probably took you a long time to find him. Why was he so hard to find? Wally wore a particular combination of clothing items that you could look for: a horizontal red-and-white striped shirt, a red and white beanie, and blue trousers. The difficulty stemmed from the fact that the artist gave those same features, including horizontal, red, white, and blue, to many other objects and characters in the scene, and human visual search for a combination of features is very time-consuming.

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



Instead of being able to rely on parallel processing to rapidly tell you where the target is, you have to bring limited 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.

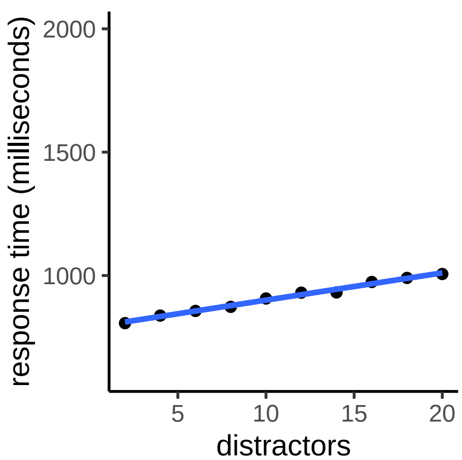


Figure 9.4: Searching for a lone red circle among blue circles and red squares. The average time it takes for a participant to find the target increases steadily with number of distractors.

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 should be very easy to find the red vertical item, simply by using feature selection for red.

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

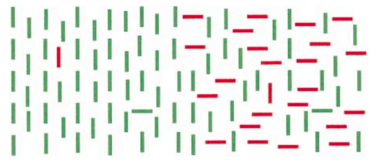


Figure 9.5: Two search displays. Left half: feature search. Right half: conjunction search

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).

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!

This search for a red vertical line among green vertical lines and red horizontal lines yields a very similar result for the pattern of response times as did the previous search, for a red circle among blue circles and red squares. Both are conjunction searches.

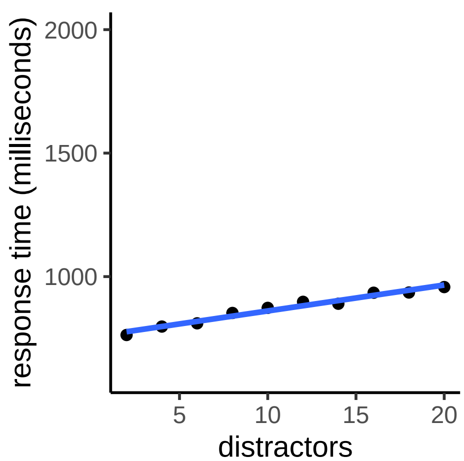


Figure 9.6: Searching for a vertical red line among vertical green lines and horizontal red lines.

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 name for this 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 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 do indeed occur serially.

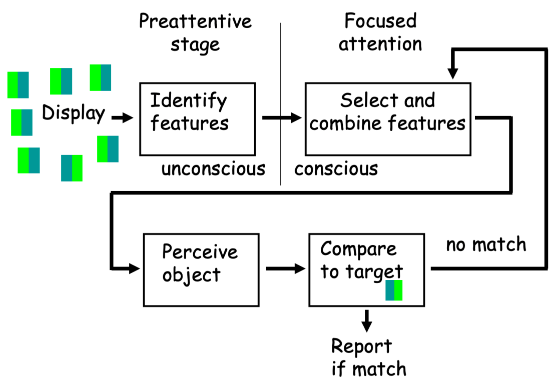


Figure 9.7: The mental processes that may be involved in serial search for a difficult target of dark green on its left, light green on its right. Combining the two features is required, which happens just one object at a time.

In 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. Think about it - you'd have to be very unlucky for you to not land on the target until you'd visited every other item in the display. This means that 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 thousands of a second, or 2 hundredths of a second, per item.

Twenty milliseconds per item is pretty fast! That's 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, or about 16 groups of three per second. That's still pretty fast, so it doesn't seem to be something you are very aware of - can you notice yourself moving your attention 16 times a second? I don't think so.

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

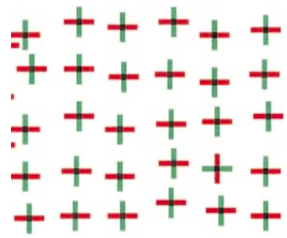


Figure 9.8: A conjunction search for red and vertical.

In the above display, vertical items and red items are interspersed throughout, making individual feature selection 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.

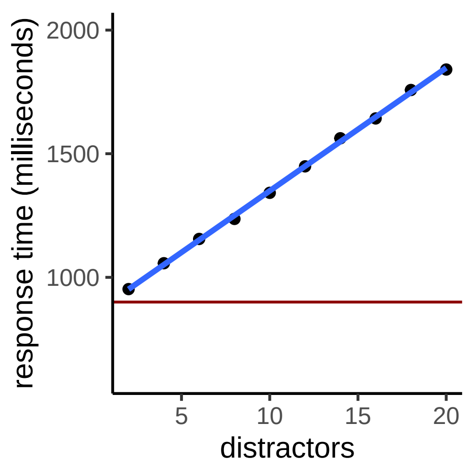


Figure 9.9: Searching for a vertical red line in the crosses display.

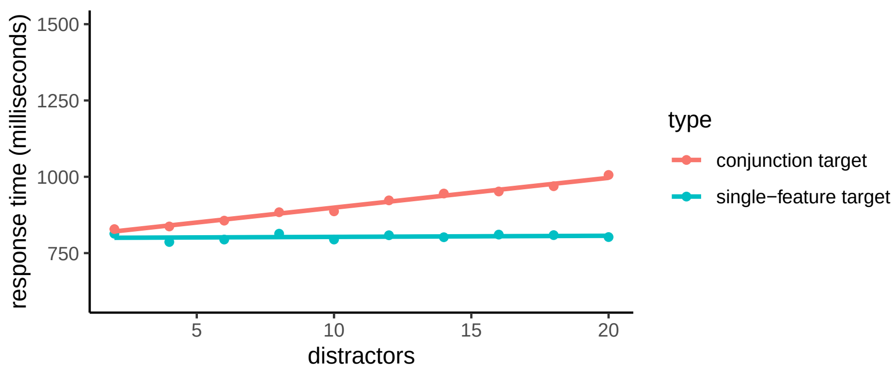
The slope of this graph is even steeper than that for the previous two searches. Here, each additional ten distractors increases response time by a full half second (500 ms). That is, the search rate is 50 ms per stimulus, or twenty per second.

To put this 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 2,660,000,000 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).
- In the 1980s, Anne Treisman 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.

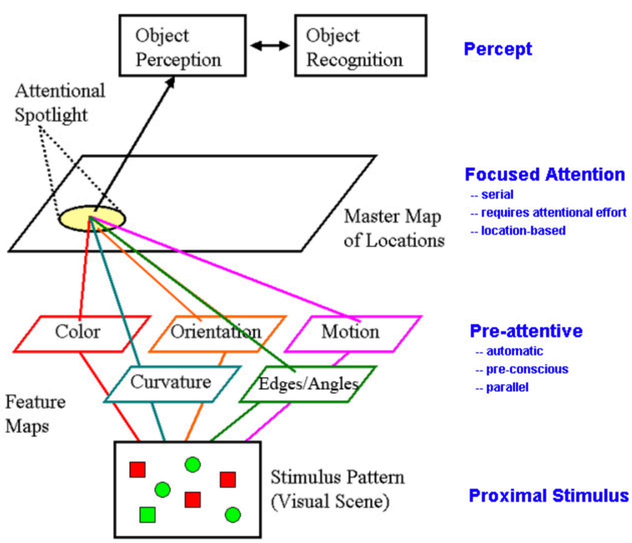
Now we can connect back to the early versus late selection question that was discussed in PSYC1. **Early selection** was the idea that there is an early bottleneck - that sensory information is not processed much before the bottleneck. If

Figure 9.10: Anne Treisman

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.

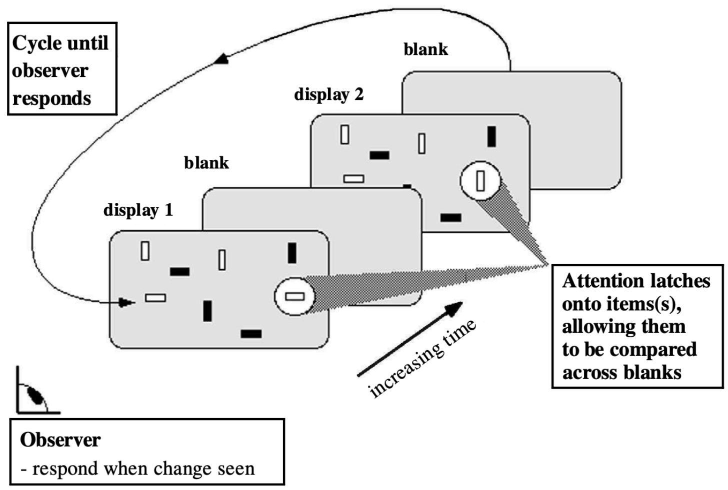
These visual search results suggest that some basic features get processed in parallel, but these features are not integrated into complex shapes or objects. To do that, selection is required. Anne Treisman's theory of the processing architecture is schematized here



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

Remember 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 natural 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).



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.

Recall from Chapter 6 that with the blank screen in the animation, the flicker/motion detectors provide no clue as to the location of the change. Do you think viewers could use their feature selection ability to help them find the change?

Because the changing object could be any of the objects in the scene, there is no particular feature people can use to find the changing object. All people are left with seems to be a more cognitive process of being able to note that an object one is attending to is no longer what it was - a change has occurred.

Rensink's hypothesis was that this kind of evaluation process exists after the bottleneck, in the realm of very limited-capacity cognition. So he expected 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 so that its features would make their way to cognition where a comparison process could be done over time.

In other words, one has to select the objects in a region with one's attention to store their appearance before and after the blank screen and then compare them to detect whether they changed. In that case, what do you predict should be the effect on number of items in the display on time to find the object?

Because only a limited number of objects can be simultaneously evaluated for change, the more objects that there are, the longer the task of finding the lone changing object should take. This was borne out by the results:

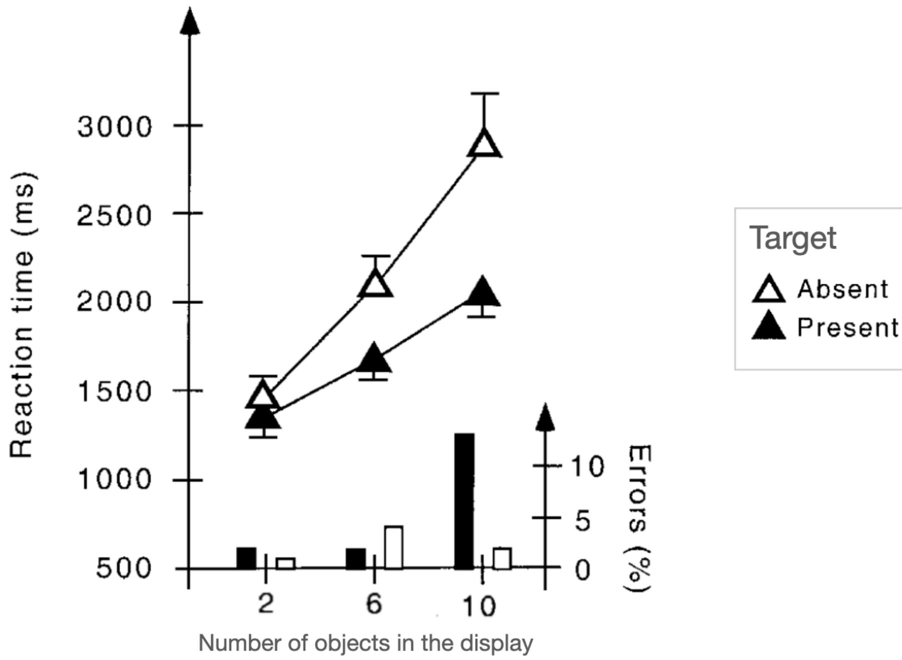


Figure 9.11: Results when searching for a lone changing object in a blank-screen sandwich, from Rensink (2000).

On some trials, the target was absent (the unfilled triangles). In that situation, the participants should not respond until they had evaluated every object in the display so they could be sure nothing was changing, hence the longer response times (on some occasions, some participants probably got tired of searching and responded prematurely).

9.7 Estimating the processing capacity of cognition's change detector

A stunning difference between Figure 9.11 and those of the earlier search results graphs (like 9.9) is how much longer the response times are. The response times are all well over a second! Why is that?

Consider what it takes to detect a change. The second picture is not presented until almost a second (920 ms) after the trial begins, so there is no change until then and no way to know what changed until then. The response time of about 1.3 seconds (1300 ms) with only two objects in the display is about as fast as you could expect people to respond - maybe one second before they can detect the change and three-tenths of a second (300 ms) to press the button to indicate they had found the changing stimulus.

Of most interest, then, is not the overall amount of time before a response, but rather the elevation in response time caused by adding more distractors (non-changing objects). If people were able to evaluate all the objects in a display (no capacity limit on change detection), then response times could be the same when there are 10 objects in the display as when there are just 2 objects in the display. Instead, the filled triangles show a steep increase in search time with number of objects in the display.

It's impossible to evaluate selected objects for change until they change (every 920 milliseconds in this blank screen sandwich). The data plot indicates that on average, 10 objects have to be added to the display to elevate the response time by 920 milliseconds. Because you can find the target on average by searching half the distractors (as mentioned above) this indicates that people could evaluate about 5 objects at a time for whether any one of them was changing. In other words, if on each cycle of the blank screen sandwich they attentionally selected 5 objects and were able to detect whether one of them was changing, this predicts the increase in response time observed in the results.

Let's sum things up. In Chapter 6 you were reminded of how long it can take people to find a changing object when the flicker/motion does not signal the changing object's location. The ingenious blank screen sandwich experiment of Rensink (2000) indicates that the reason it takes so long is that people only have the capacity to evaluate about five objects at a time for change. This reflects the very limited processing capacity of the cognitive processing we have to rely on when flicker/motion and feature selection doesn't help us. Note that Rensink (2000) used very simple objects - oriented lines. The capacity limit may be even worse for more complicated objects.

9.8 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

Social factors

Two thousand years ago, Aristotle wrote: “Man is by nature a social animal.” He was right — people are very interested in people, and in working out what they’re thinking and feeling. There are evolutionary reasons for this. We live in groups and like us, our ancestors depended on each other not only for mating opportunities, but also to survive.

This aspect of our nature shows up in what we choose to look at. In a previous chapter (4), you learned that most movements of attention are overt. If people are interested in something, they usually will look right at it.

Once researchers were able to record eye movements, they asked study participants to look at pictures of scenes. If there were people depicted in the scene, rather than looking mainly at parts of the scene that were salient due to bottom-up factors, such as colors or shapes that stood out by being unique, the participants tended to spend a lot of time looking at the people.

One scene researchers have collected data for is the dinner date change blindness scene. Specifically, the eye movements of people were recorded while they were viewing the animated blank screen sandwich, to see where people look.

The above image shows some eyetracking data for one participant. 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, and frequently did not occur in left-to-right reading order.

Remember that the task of the participants was to detect what was changing in the scene, which was presented in blank-screen sandwich format. The participants presumably understood that the change could be anywhere in the scene, but the data shows that they barely looked at the background, the bush in the foreground, or the railing (which was what was actually changing in the change

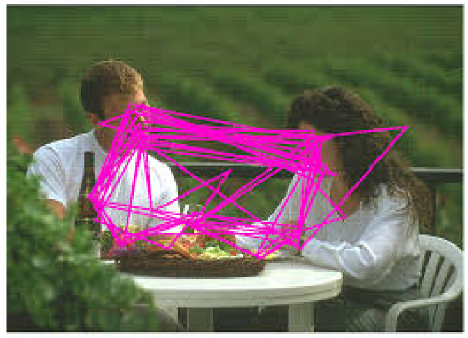


Figure 10.1: Pink lines indicate the trajectory of eye movements made by people searching for a change.

blindness animation). Rather than taking a systematic approach of moving their eyes across the screen bit-by-bit, this participant (and many others) gave in to their bias to focus on the people.

Things like people and bodies are sometimes referred to as a scene's **high-level properties**. The word "high-level" is used in part to indicate that those properties take more processing to extract, and thus are represented at later ("higher level") stages of the brain compared to, say, color. The same is true of objects like food and wine. They are not recognized as objects (as opposed to as meaningless collections of colors and edges) until the temporal lobe, after many stages of visual processing.

Many of the study participatnts likely were interested in understanding the "meaning" of the scene, which 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. If so, you might be like one of those participants (this is just an idea, NOT a validated psychological finding).

The above image shows the data for only 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 that was heavy on dialogue and characters (the *Andy Griffith* show). The soundtrack was turned off during the viewing.

The results, shown above, provide further support for the hypothesis that attention is biased towards faces. In another study, 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 social biases of attention are used by web app and advertisement designers who seek to control what you attend to. You may have noticed that many ads

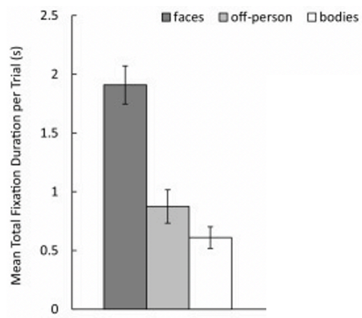


Figure 10.2: Average amount of time spent looking at different parts of the scene.

have a picture or animation of a person in them, even when this is completely unnecessary and superfluous to the information provided. It's kind of hack of your attentional system to get you to read or watch ads.

There is some evidence that the attention of many children with autism spectrum disorder (ASD) is less biased towards faces than is that of typically-developing children. The study whose results are plotted above was one study that investigated this, by also including among their participants a group of sixteen adults with autism spectrum disorder.

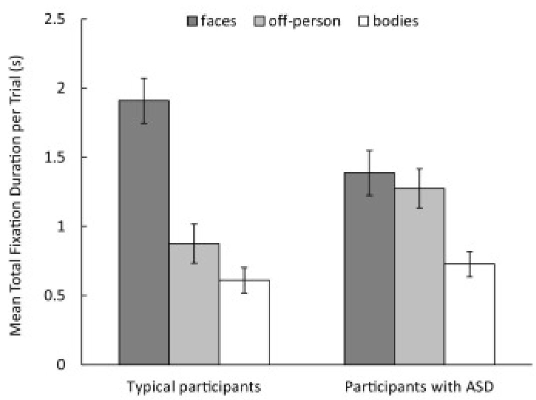


Figure 10.3: Average amount of time spent looking at different parts of the scene, in sixteen adults with (right) and without (left) autism spectrum disorder.

Based on this study, children with ASD don't spend as much time looking at faces. 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.

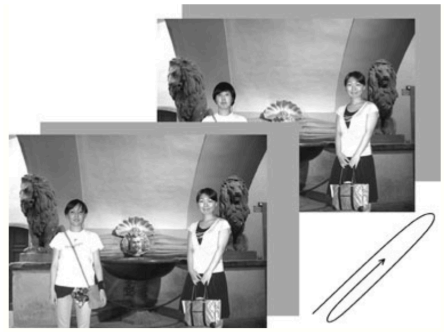


Figure 10.4: A schematic of one of the trials in the experiment. This trial is an example of the head change condition. The head of the person to the left was replaced by another head.

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. Shorter response times indicate that the person detected the change faster.

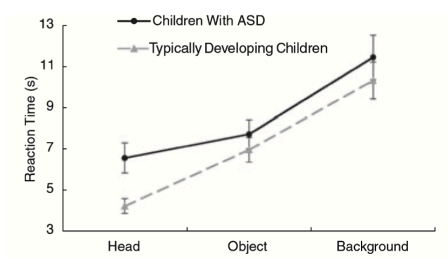


Figure 10.5: Mean correct response time for detecting a change. The black line represents children with ASD and the dashed gray line typically developing children. Error bars are one standard error.

10.1 Other factors

One shouldn't assume that people will always look primarily at the faces in a scene before looking at other things. When cycling or driving somewhere, there are times where it is important to keep our eyes on the road, such as when approaching a roundabout. At such times, most of us do not take our eyes off the road to linger on the faces of pedestrians or the drivers of other cars. At the same time, if we are bumper-to-bumper in a traffic jam, we might spend more time looking at the people in nearby vehicles.

Evidently our brains flexibly change visual priorities based on our interests and the demands of the task we are engaged in. Of course, this is not perfect, as we'll discuss in Chapter 16.

10.2 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 11

Real-world search

It was the use of certain technologies that helped spur psychologists to study capacity limits, including the problem of fighter pilots having to deal with both flying and monitoring things in the cockpit at the same time. In the aftermath of the September 11 attacks on the U.S., monitoring another aspect of airplanes became prominent: what goes on and off of them.

Figure 11.1: The hijackers of a plane that was flown into a building in New York carried knives in their hand baggage.

After 2001, tens of thousands of additional workers were hired at airports to search baggage and baggage scanner images for prohibited items.

Which type of response time pattern do

you think this kind of search would result in - parallel search or serial search? See chapter 9 for a reminder of what this means.

This kind of search involves additional complications beyond those described in the previous chapter (9), and the baggage inspections result in plenty of mistakes. More than once I've accidentally left a prohibited item in my checked baggage or hand baggage, but nevertheless sailed right through.

In 2015, a "sting" operation by the U.S. government reportedly found that baggage screening only caught 5% of bags with (fake) weapons that undercover agents had put in them. Two years later, a similar test found some improvement, to catching about 30% of bags with weapons inserted in them. Still, 30% is clearly a failing mark.

Why did the workers fail to detect such a high proportion of weapons? We don't know the details, because how the tests were done and the exact pattern of results are secrets. However, from considering the nature of the task more, you should be able to understand the basic issues.

In the searches described in 9, the target and distractors were chosen to be very different from each other. For example, when the target was a red vertical line, the distractors differed greatly on at least one feature, color or orientation. There were never any orange vertical lines, or red slightly-tilted lines. This was to ensure that the search would *not* be difficult due to the early visual system not being good enough to strongly distinguish between the target and distractor. By using displays that the early visual system can easily distinguish, any findings of poor performance could more confidently be attributed to a bottleneck. This situation was contrasted with searching for a particular word on a page full of text (9.1). When you looked at one part of the page, for most of the rest of the page your vision was too poor to distinguish between different words. There's a similar problem with looking at an image of a piece of luggage. This means that airport workers need to move their eyes around the image, re-doing the search for each region in the image that their eyes land upon.

Can you find the prohibited items in the bags above?

Another feature of the laboratory tasks is that participants only had to search for one kind of target at a time, for example for a red vertical line. But airport workers have to keep in mind many targets - scisssors, knives, and lighters as well as guns and explosives. This makes a big difference. Let's consider what it means for the use of feature selection. Feature selection really helps when the object you are looking for is a unique color, orientation, or simple shape. Unfortunately, prohibited items can be oriented in any direction in a bag and can be any color. The scanner technology does, however, portray metal objects in a single color (blue above), which is helpful because knives and guns are more likely to be metal than made of other materials. So, baggage inspectors can to some extent set their featural attention to blue. Unfortunately, however, even a single contraband category, such as guns, can assume a number of different shapes. The pistol in the bottom right suitcase, for example, has a surprising

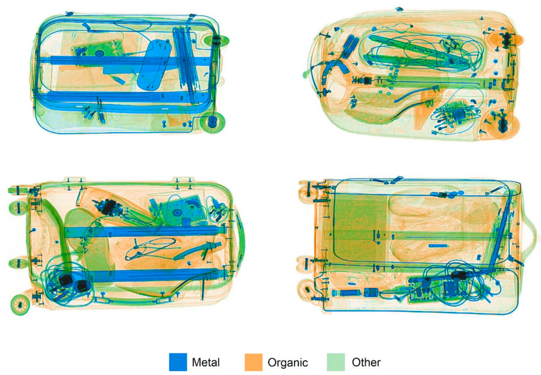


Figure 11.2: X-ray images of bags, two of which contain threats. From @donnellyUsingEyeMovements2019

shape because it is only partly metal and it is pushed against another metal object. So, shape selection ends up not being particularly useful.

Relatedly, even when you look directly at the pistol in the baggage scan, you may not recognise it as being a pistol. This testifies to the need to have expertise to be familiar with all the appearances that a prohibited item can take in a baggage scane.

Finally, laboratory studies have revealed that even in the best of conditions, people are not very good at searching for more than one kind of target at a time. Recall that the capacity limit on full processing of objects means that searches for complex shapes and combinations of features is slow. Something we haven't talked about is how the brain evaluates whether a fully-processed object representation is a target. To do so, you have to keep what the target is (e.g., "red vertical") in mind. But baggage inspectors have to keep a whole bunch of targets in mind. In that situation, limitations on **memory** can impair performance (Wolfe, 2012). What sort of memory is used for knowing what the targets are? At first, it may be short-term memory, but after one does the task over and over, it seems to become a form of long-term memory (Wolfe, 2020, page 13).

In summary, baggage scan searches are difficult for several reasons:

1. The targets and distractors can be very similar to each other.
2. The targets are not distinguished from the distractors by a single feature.
3. The targets do not have a consistent appearance.
4. Expertise is necessary to know all the possible ways a target might appear.
5. There are several different kinds of target, all of which have to be found.

With multiple factors involved, affecting different mental processes, psychologists' theories have not advanced enough to accurately predict the performance of people on the baggage scan task. This is true for most real-world tasks. They tend to involve a lot of variation that affects multiple mental processing stages, and our theories are too limited to make good quantitative predictions.

Appropriately engineering a system of humans and machines, then, for baggage inspection or anything else, requires a lot of trial and error. The principles of mental processing inferred from laboratory experiments can guide how a system is designed, but ultimately a lot of tests have to be done to validate those design decisions. For the case of airport baggage inspections, despite a lot of resources devoted to the problem, still the system scores a failing grade. The problem is just too difficult, at least for present scientific understanding and technology.

11.1 Practice and individual differences

Managers of airport security are need to know which people to hire and how to train them. The following, then, are some important questions:

1. With practice, do people improve much at detecting banned items?
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:

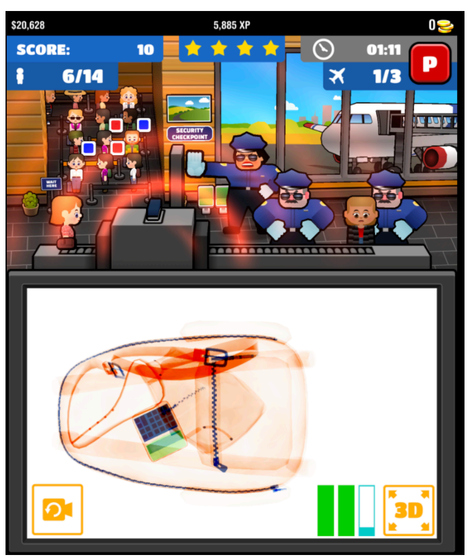


Figure 11.3: A screenshot from the game Airport Scanner.

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

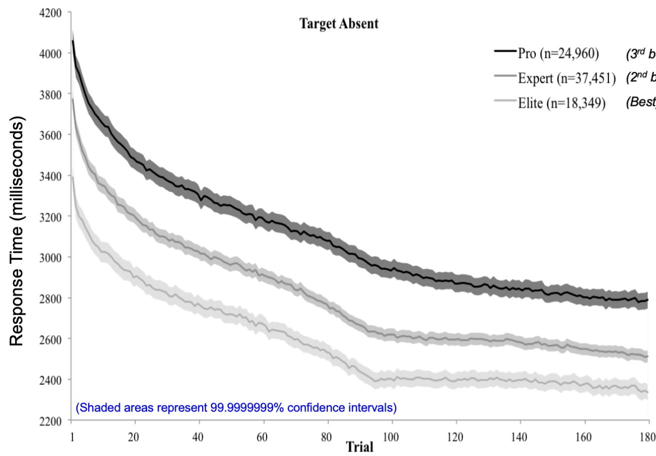


Figure 11.4: Response time when searching for a prohibited item versus trial number, for participants with 3 different levels of initial performance.

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?

11.2 Are running, cycling, and driving like change blindness experiments?

Cycling (private video shown during lecture)

Chapter 12

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.

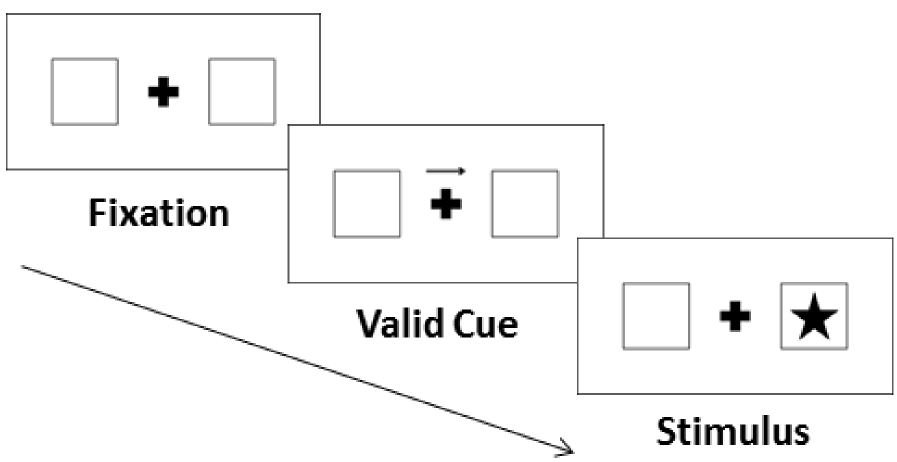


Figure 12.1: A schematic of a location cuing experiment (created by Local870, CC BY 3.0)

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 those things look like! They have to rely on bottom-up attention, so this 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.

12.1 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 can be 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.

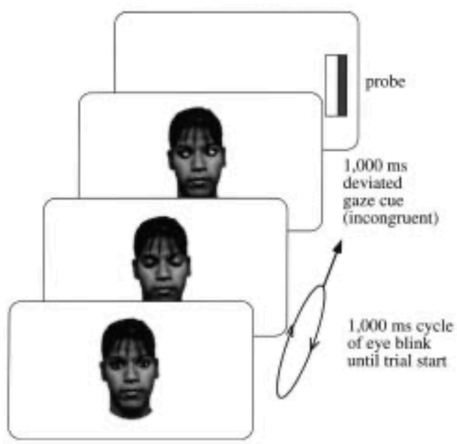


Figure 12.2: A schematic of the task used by Hood et al. (1998).

It might be a bit hard to see in the above image, but on the screen before the probe, the eyes of the face are looking to the left. In half of trials, they looked to the left like that, 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 study's 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 12.3: 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.

12.2 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:

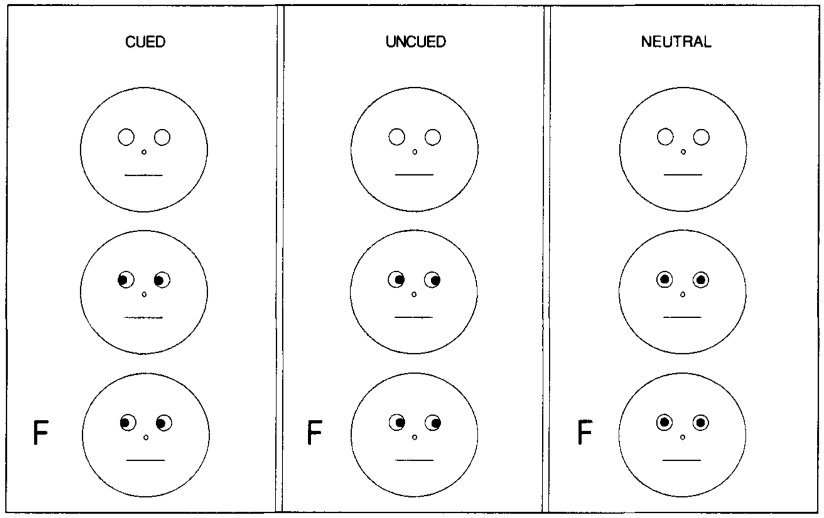


Figure 1. Examples of cued, uncued, and neutral trial sequences. Each trial began with the presentation of a face with blank eyes. After 680 msec, pupils appeared in the eyes, looking left, right, or straight ahead (the gaze cue). Then, after 105, 300, 600, or 1,005 msec, the letter F or T (the target) appeared to the left or the right of the face.

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.

- 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).

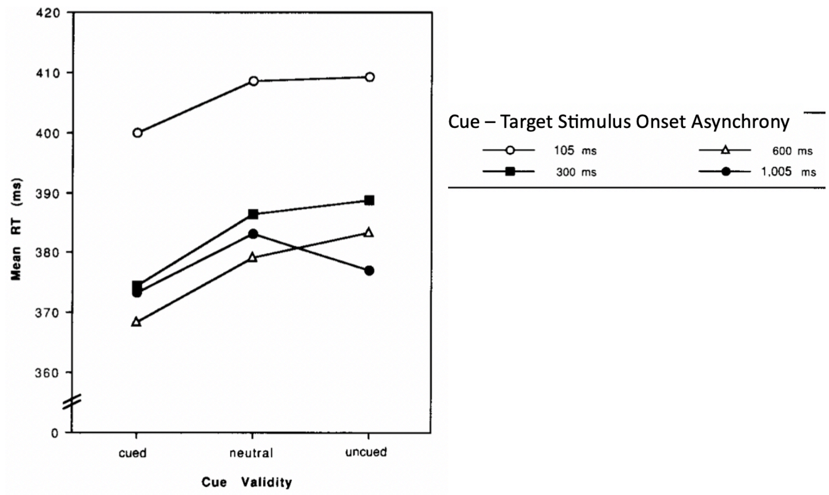


Figure 12.4: Response time against cue type, for four different stimulus onset asynchronies

12.3 Exercises

Eye gaze is an example of how social stimuli influence attention. What is an example from a previous chapter?

Chapter 13

The role of memory and expectation

13.1 The contents of memory can guide attention

13.1.1 Working memory

Working memory refers to information that you currently have in mind, often long-term memories that you are "working" with at the moment and thus may be present in consciousness. 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 an early 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 at the end of the trial they'd see a shape and be 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.

The digits to be memorized were presented on a background of three shapes, and critically, one of those three shapes had been presented previously as a prime shape to be memorized. On 10% of trials, that previously-memorized shape was also the shape to be memorized for the current 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.

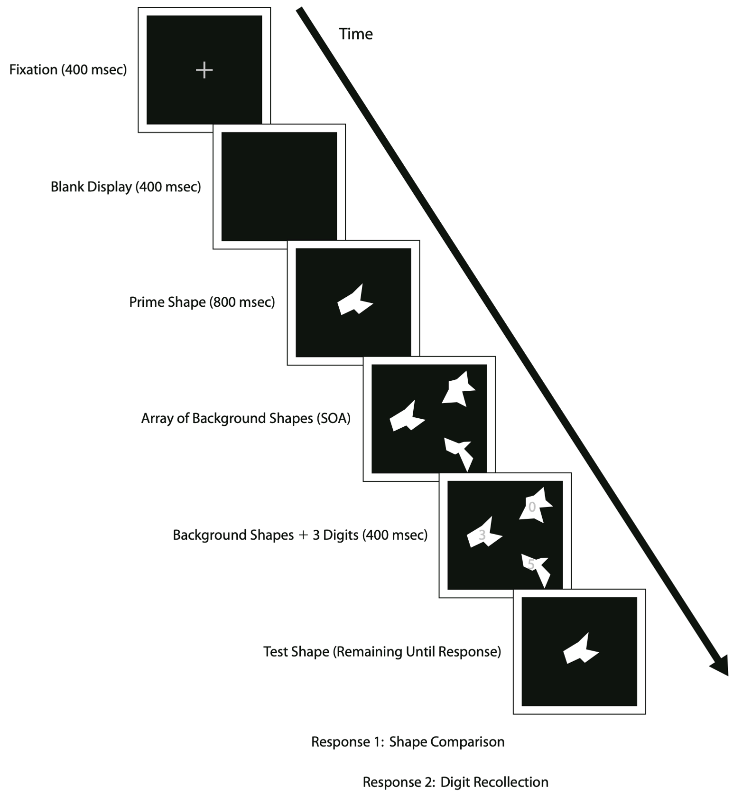


Figure 1. Event sequence of one trial in Experiment 1.

Figure 13.1: Schematic of the trial sequence in the Huang and Pashler (2007) study.

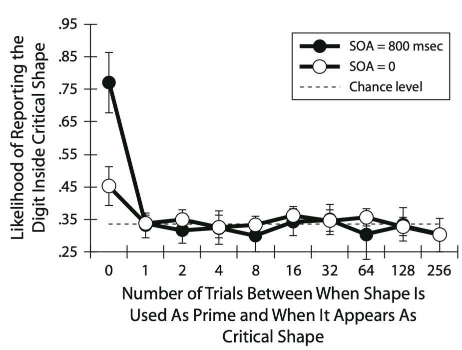


Figure 13.2: The leftmost data points represents performance when the digit was presented on that trial's prime shape. It is much higher than "chance level" - the rate of performance if participants merely guessed a digit (the dotted line). The advantage is bigger for the 800 ms SOA.

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

13.1.2 Long-term memory

That the contents of *long-term* memory may also 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 the Baader-Meinhof Gang, it's more likely that you'll pay attention to and thus remember any individual encounter with it.

13.2 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 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.

13.3 Four factors for managing attention

Apollo used the following four things to manage people's attention:

- 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.

13.4 Inattentional blindness

You might have already seen some of the inattentional blindness movies in this section, but now you are equipped with some new ideas to apply to them. Keep the four factors (13.3) 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. And 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 13.3: 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!

13.4.1 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. But 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 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.

13.4.2 The door swap

One of the funniest inattentional blindness demonstrations is the door swap. Watch this version created by Derren Brown.

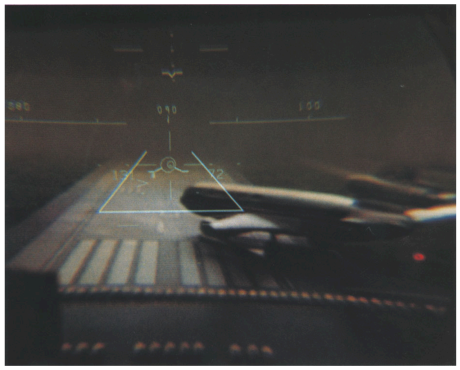
In the Derren Brown video, it seems clear that some people definitely didn't notice the swap, and it's pretty amazing if even only a minority of people didn't notice the swap. However, to make the video more amazing, the TV program producers may have edited the video to omit some cases of people who did notice and say something when the person changed. Because of such hijinks, you should always be suspicious of TV shows and 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 initially worried that perhaps those people did notice, but 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. The results of this and other studies supported the idea that many people don't notice such changes. 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?

13.4.3 In the air

Inattentional blindness also occurs in the real world. Landing a plane is a demanding task that involves monitoring several instruments, even if cockpit designs are better than in the 1960s (3.2). 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. Indeed, narrow task engagement is a factor that frequently contributes to accidents. This is the third factor mentioned in 13.3 that can be used to manage someone's attention.

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. Optional: read this article on pilots not hearing alarms.

13.4.4 Exercises

- Why do bottom-up attentional processes fail to reliably draw attention to the gorilla in the Simons and Chabris (1999) video?
- What is the key characteristic that differentiates inattentional blindness from change blindness?
- Chapter 5 mentioned that being touched can easily draw your attention. This is a problem for pickpockets and magicians. How does Apollo Robbins solve this problem?

Chapter 14

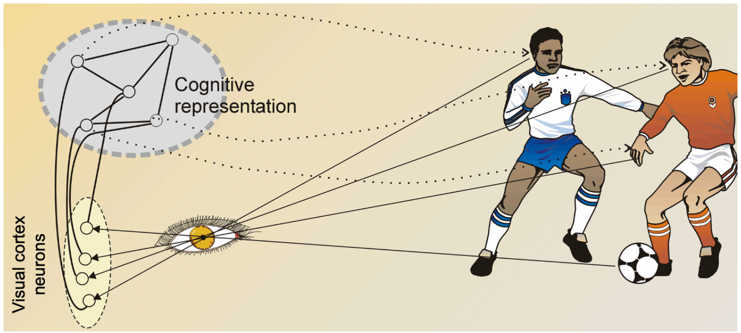
Divided attention and objects

You already know about selecting a location with attention. You can actually 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 locations or stimuli at one time.

You'll recall that higher-level processing is so limited that even identifying a simple feature of two objects yields much worse performance than identifying one (3). 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. This is consistent with the idea that any capacity limit on selection is not as severe as that on processing things for identification.

14.1 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 parts 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 locations 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 can be time-consuming (9). Even if it only took a third of a second to find the ball, for example, during that time a penalty kick traveling at 30 metres per second will have travelled a full ten metres.

Fortunately, attentional selection can easily follow a moving object. This is usually studied with the "multiple object tracking" procedure. First a bunch of identical objects appear on the screen.

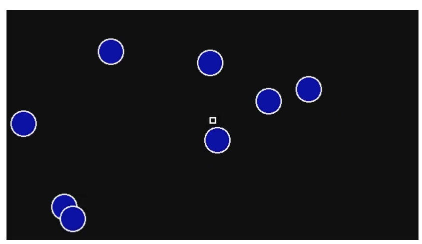


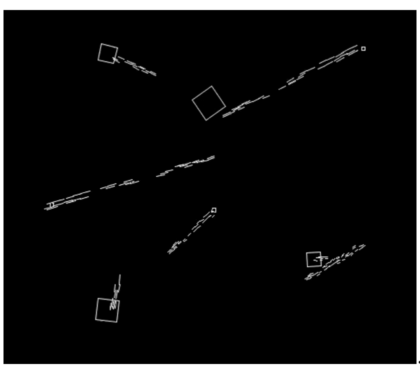
Figure 14.1: A snapshot of a multiple object tracking task

After the objects appear, 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.

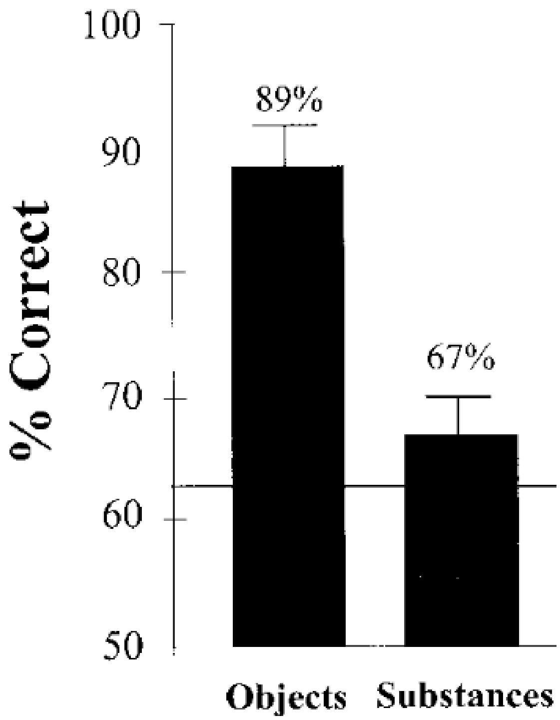
The task feels pretty natural. Once one selects an object with attention, if it starts moving, your attentional focus will tend to move along with it. What we have previously been referring to as location selection (7) may best be thought of as **object selection**, because if an object in a selected location starts moving, your attention naturally follows along rather than sticking to its original location. 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.

This fact that attention can select objects has important implications. For it to be possible for attention to select objects, some think the visual system must have first processed the image preattentively into objects, otherwise attention may not have been able to select an entire object. This is similar what Treisman used when arguing that processing of individual features occurred prior to the action of attention (*preattentively*) - without that pre-processing, 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.



Experiments like that of VanMarle and Scholl (2003) give an indication of how our visual system carves up the world into objects prior to the action of our attention. Our visual system not only carves up the world into objects, it subsequently groups objects based on how they move in relation to one another or are connected to each other.

The lines in the above animation fade in and out. When the lines are present, you may see the circles moving in individual revolving triplet groups. This is an instance of grouping by motion. When the large lines are present, the visual system unifies the discs of each color into a large circle structure.

14.2 The limits on tracking implicate slow, capacity-limited processing

People can only attend to a limited number of moving objects at a time. This is unsurprising, because the nature of attentional processing is that it is a *limited* resource for more extensive processing.

That limited resource also seems to be doing some of the work or computation required to keep attention on a moving object. One indication of this is that the maximum speed for tracking one target is faster than the maximum speed

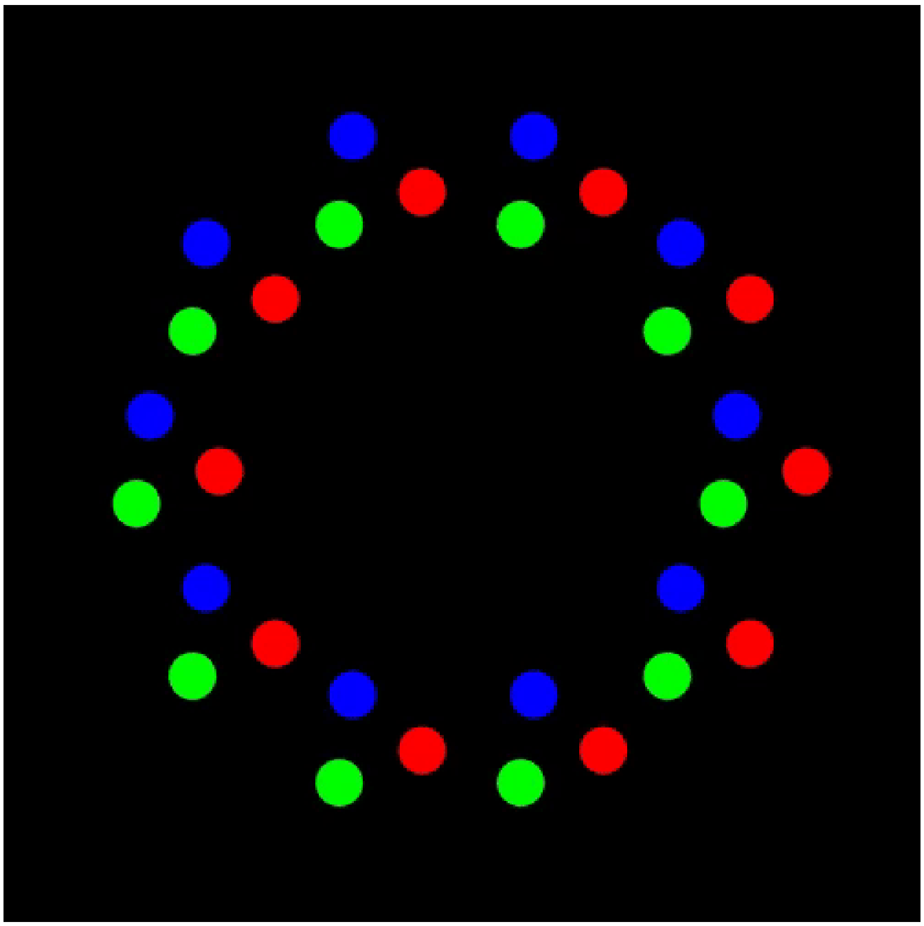


Figure 14.2: An animation by [beesandbombs](https://twitter.com/beesandbombs/status/791429723138224129).

for tracking two targets. Measure your speed limit here for one target and here for two targets. What were the two speeds, and which one was faster?

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). This is probably because tracking is mediated by high-level processes rather than the specialized processes that mediate motion perception.

14.3 Test yourself

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

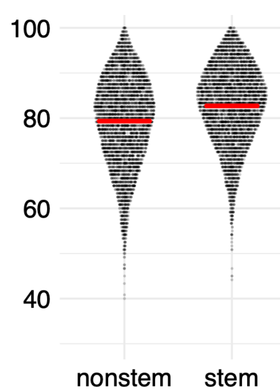
https://www.testmybrain.org/classroom/

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 (Treviño et al., 2021).



However, the difference is not large. If you took a random STEM person and a random non-STEM person, the overlap of the distributions above implies that a 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.

14.4 Exercises

- In 7, you learned about three kinds of attentional selection. How does this chapter change your interpretation of location selection?
- 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 15

Bottleneck - memory

What does it take to form a memory?

If I ask you about your memory, you'll probably assume that I'm asking you about things pretty far in the past.. at least a different day. But memory is required even for you to report something you saw a few seconds ago. In the laboratory, researchers have devised situations in which you will occasionally forget things that happened just two seconds ago.

One thing researchers do in their experiments is show people random, unrelated stimuli. That way, people can't use their knowledge of the way the world usually works to reconstruct what must have happened. In the real world, if I hear the doorbell ring and ten seconds later find myself holding a package addressed to me, I can reconstruct that probably what happened is I went to the door and got the package there. Standard sequences of events can help our brains reconstruct events in an act of remembering based on only fragments.

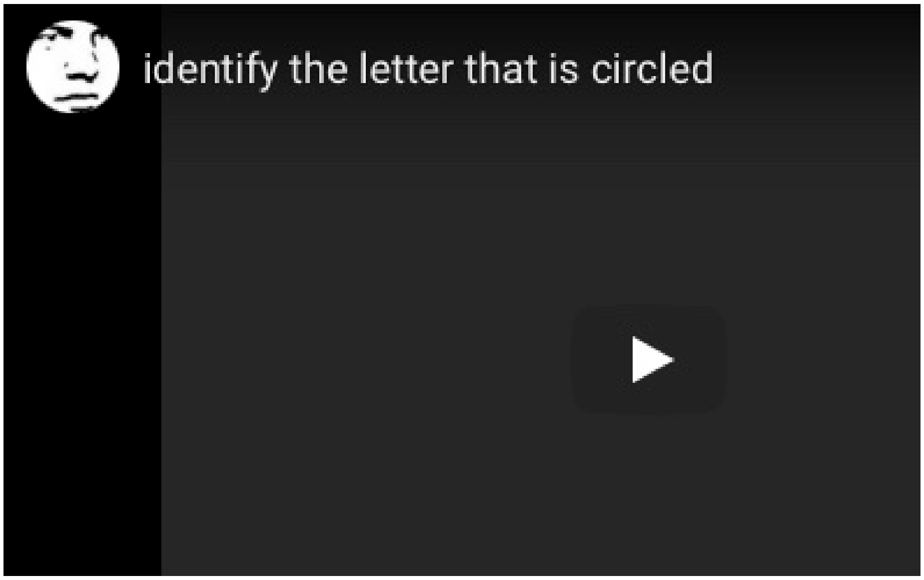
Besides showng people random unrelated stimuli in memory tests, another thing that researchers do is present stimuli quite briefly. Let's say I ask you to remember a photo like this one, of a moth I photographed in a rainforest in Panamá.

As you examine the moth, you'll notice various things. I can tell you that for me, the white and black pincer-shaped wing markings jumped out, as did the fuzzy white front edges of the feathery wings. You might start making various associations, possibly to an eagle's open talon for the wing markings, or to a fur coat for the white edges of the wings. The longer you look at it, the more details you notice and think about, which creates more and more parts of a memory that later can help you remember the image.

If you were shown the picture of the moth only very briefly, then you'd have less chance to form strong memories for it. Brief exposure, in combination with the use of unrelated stimuli, help researchers to probe how memory works even in the time available in a laboratory exeriment, testing for a memory only seconds after someone is shown a stimulus.

A third feature of memory experiments is that the stimuli tend to be much more boring than my photo of the moth I showed you above. This is partly for standardisation purposes, so that stimuli can easily be swapped and result in similar findings. So, sometimes researchers just use individual letters, as you'll see below.

Below is a movie that shows a rapid series of letters. While you watch it, look out for and try to remember the one letter that is circled.



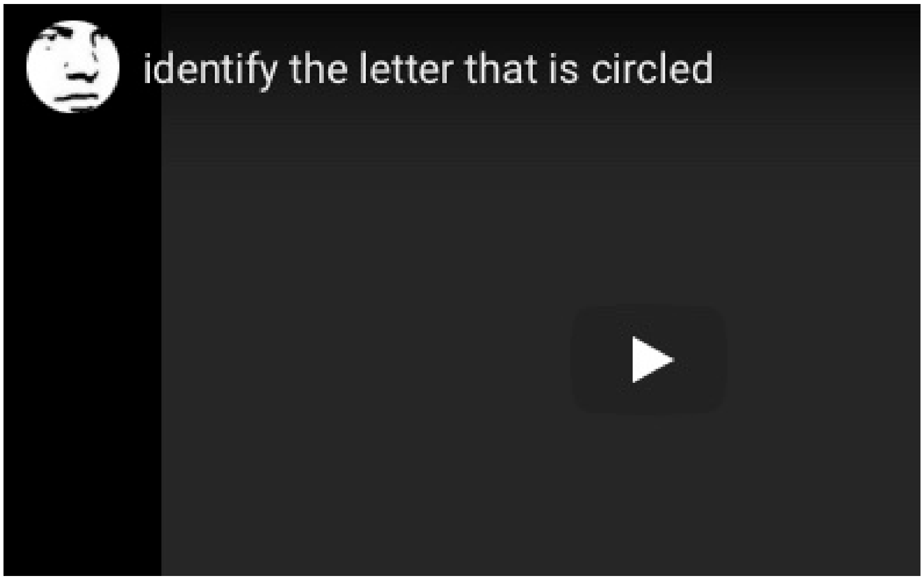
If the movie didn't work, you can watch it here.

Hopefully you were able to see which letter was circled. However, the act of trying to remember that letter can impair your memory of subsequent letters in the stream.

15.1 The attentional blink

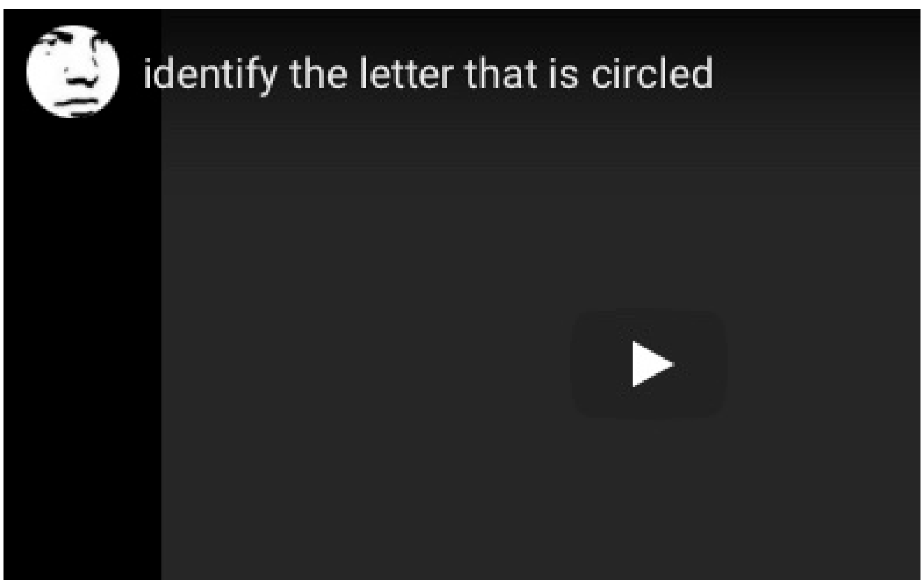
Putting together the three methodological features described above, researchers have investigated the limitations on memory processing by asking people to remember two images from a rapid series.

In the movie below, two letters are circled. Watch the movie and try to remember the circled letters.



If you can't see the movie above, you can watch it here.

How did you do? If you're good at the task, the above movie might have been too slow for you. In that case, try the one below.



If you can't see the movie above, you can watch it here.

When we investigated this in my laboratory (Goodbourn et al., 2016), we tested people's performance when the two circles were presented at different times relative to each other. In particular, the second circle was presented at different *lags* after the first circle.

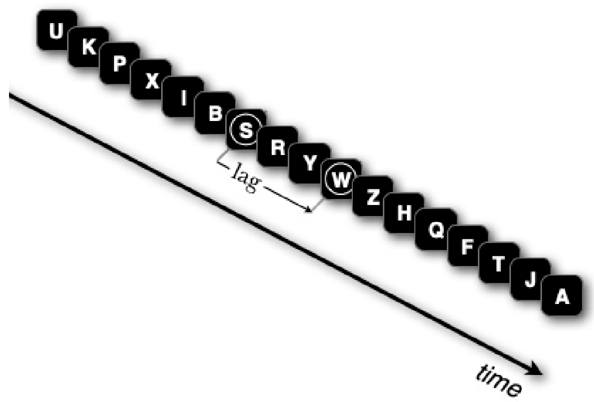


Figure 15.1: Schematic of a rapid sequence of letters, with a particular temporal lag between the two circled letters

At fast rates, most of the participants reported the identity of the first circled letter correctly about 60% of the time, regardless of the lag of the second circled letter. This is not too surprising - if a letter is presented too quickly, there often isn't enough time for the visual system to fully process it before the next letter overwrites it. More interesting from a memory perspective was the pattern of performance for identifying the second letter.

When the lag was long, so that the second circle was presented a long time after the first circle, performance was about the same for reporting the second letter as for reporting the first. But when the lag was shorter, performance reporting the second circled letter was quite bad.

Notice that when the second circled letter occurred about 200 ms after the first one, people only got it right about 10% of the time. Performance stayed lower than for the first circled letter until the lag was over 600 milliseconds, a bit more than half a second. This phenomenon is called the *attentional blink*. It's called the attentional blink because the researchers who documented it thought that the act of deploying attention to the first letter disrupted attention for several hundred milliseconds, preventing its deployment to the second letter.

In the years since its discovery, hundreds of experiments have been done on the attentional blink and these favor the idea that although attentional disruption may be one reason for it, a bigger factor is the time-consuming nature of memory consolidation. When the first letter is attended, your brain starts the processes of committing it to memory. This takes time, which mightn't be a problem if it could process multiple events at once, but it can't. That is, memory creation is subject to a bottleneck. And it's not just a problem for letters - it's a problem for anything you present to a person.

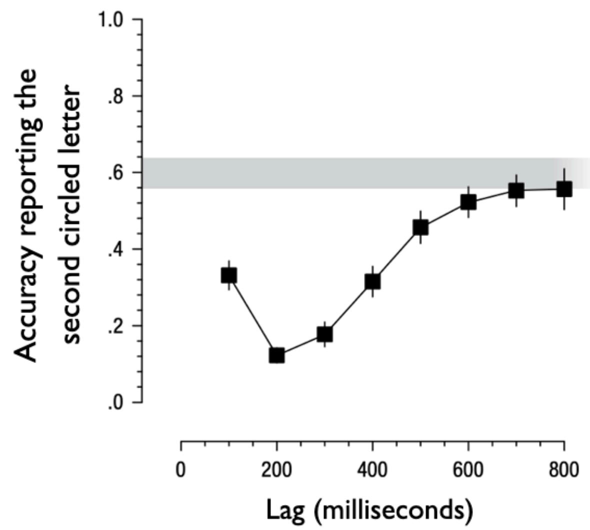


Figure 15.2: The vertical axis is the accuracy reporting the second circled letter, as a function of lag (only trials where the participants also got the first letter correct are included). They grey band shows the proportion of trials that average participants were correct for the first letter.

Previous chapters described the evidence for a bottleneck for judging two features of a single object (3.4 and 8).

The attentional blink is thought to reflect a bottleneck situated after objects are recognised. In the schematic below, the bottleneck for processing multiple features of an object would correspond to the "attentional bottleneck" and the bottleneck that causes the attentional blink would correspond to the "memory storage" process.

Memory involves three processes:

- Encoding - the sensory information must be successfully transferred into a durable representation.
- Storage - the encoded representation must be successfully stored until you need to retrieve it.

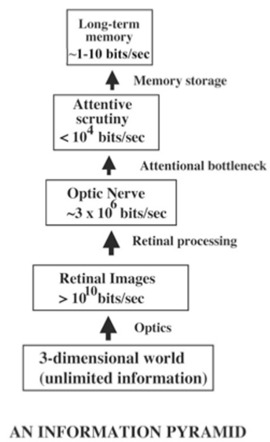


Figure 15.3: An estimate of the amount of information processed at successive stages of visual processing.

- Retrieval - the encoded representation must be successfully found in memory and retrieved.

Remembering something requires successful encoding, storage, and retrieval

In brains, the three processes are highly intertwined, but still they are somewhat separate. The inability to report the second letter in the attentional blink is thought to be a failure of memory encoding. The idea is that while the brain is busy encoding the first circled letter, it is unable to encode the second letter because it cannot easily encode two things at a time. In other words, memory encoding is subject to a bottleneck.

Fortunately, the world is not made up of an unrelated series of events, such as rapid series of random letters. Things typically change gradually in the real world, so we have more time to encode things than do participants in an attentional blink experiment. Moreover, when things do change, the changes are often somewhat predictable. The brain can pre-activate expected information and get a head start on encoding them into memory. This is one reason that in the real world our encoding limitations are not as noticeable as in a laboratory experiment.

Chapter 16

Distraction and learning

Figure 16.1: **The three wise monkeys from a Buddhist legend, as depicted in a wood carving in the 17th-century Tōshō-gū shrine in Nikkō, Japan. Photo: Ray in Manila**

Sometimes, we may feel a bit like the leftmost monkey above, 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 simply turn off our hearing or vision at times and thereby avoid distractions. As we discussed in 5, however, as creatures that evolved in a dangerous world, our brains are evolved to be continually vigilant.

This may be most obvious in prey animals. Horses are one example. As a result of them occupying a vulnerable ecological niche, horses evolved large, side-mounted eyes that give them almost 350° of vision, with only a small blind area spanning just 15°. This helped them detect their predators, such such as wolves and lions, that might otherwise sneak up on them.

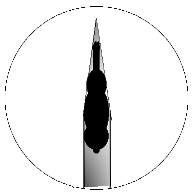


Figure 16.2: Schematic of a horse from above. The white areas depict the visual field, with the small blind region in grey.

The term “getting spooked” refers to when a horse becomes startled or afraid, which can be triggered by something they see or hear, and can cause them to bolt or engage in other evasive maneuvers. As prey animals, they are always vigilant for possible threats.

To avoid triggering these responses, some owners of working horses fit them with blinders to restrict their field of vision. Whether a horse’s owner is trying to get it to pull a carriage or to win a race, the blinders, also known as “blinkers”, help keep it on task by preventing anything to the side or behind from causing the horse to spook.

Humans may have not had as much to fear from other animals as horses did, but we still faced dangers in our ancestral environment, from other people as well as from lions. Like that of horses, our brains are constantly monitoring incoming signals from all our senses for things that might be a danger. And as social animals, we are also innately attuned to the presence of other people.

When I’m working at a café, the faces of the people around often distract me from my work. I sometimes literally blinker myself to avoid this. I’ll bring a baseball cap with me to the café, and pull the brim down so that I don’t see as many of the faces that are coming and going.

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. Not babies, though - they can be much harder to wake. They evolved to rely on adults for their defense, so they can afford to concentrate on brain development rather than monitoring the environment.

The top-down attention chapter (5) introduced the idea of *attentional capture*. The experiment discussed there, by Theeuwes et al., used these sorts of displays:

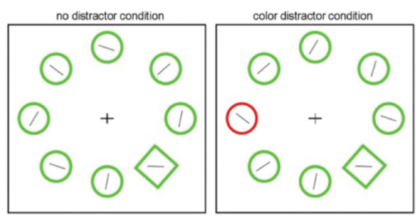
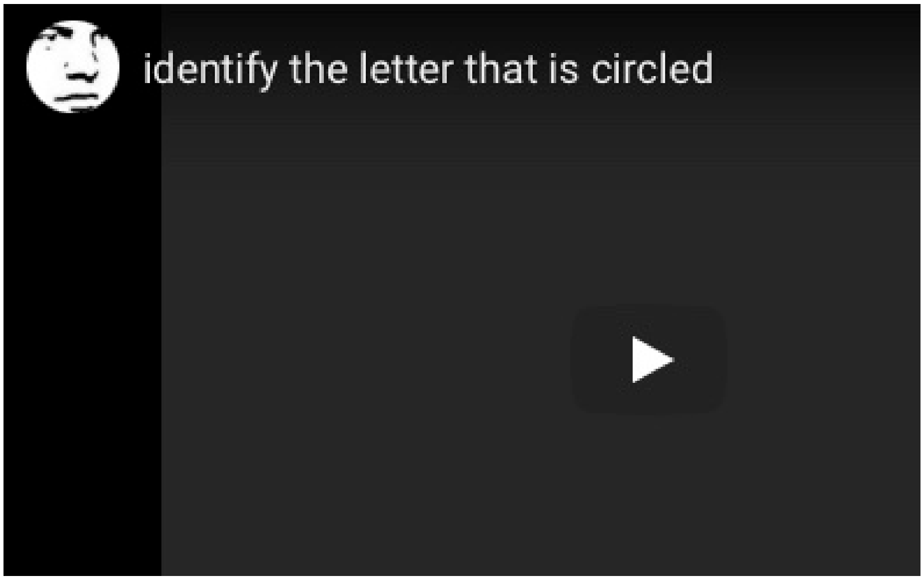


Figure 16.3: **Two displays, one with a distractor (right) and one without (left)**.

The task for participants in the experiment was to find the diamond and determine the orientation of the line within it. People took longer to do that in the presence of an odd-colored distractor like the red circle at right. It seems that on some trials, participants couldn't help but shift their attention to the red object, and only after that did they attend to the diamond. All sorts of sensory signals can be distracting, of course, not just odd-colored items.

Odd sounds, for example, are often distracting. One reason is that hearing is especially important for "interrupt signals" (5) because hearing provides an early warning system. With vision, you need to look around, moving your eyes about, to detect any threats around you. With hearing, in contrast, you can detect things anywhere around you without even moving your ears and your head. Indeed, most of us can't even move our ears, because we don't need to move our ears or head to listen.



Some people can move their ears, there's not much reason to.

Sudden sounds in an otherwise quiet scene, then, can distract our attention, taking us off task.

In such a situation, you can think of the sound as like an odd color in a scene that otherwise is uniformly colored. Similarly, if you feel a sudden touch against your body, this will also attract your attention. Indeed, preventing a touch from alerting people is the primary challenge faced by pickpocketers. The magician and pickpocketer Apollo Robbins discusssed in 14 addressed this by touching his victims throughout his routine, so that the touches corresponding to removing the victim's watch wouldn't be registered as anything different.

16.1 Bottleneck review

Previous chapters emphasized that many changes in a scene can go completely unnoticed. That's true even if the changes are highly unusual, as long as they don't stand out by being the only location associated with a flicker or motion signal. The reason for this is that because of the brain's bottlenecks, most objects in a scene may never be identified. So, most of the unusual aspects of a scene will never have a chance to attract attention because they aren't processed enough by the brain for the brain to determine whether they are unusual. If you show someone a cluttered scene and then tell them to find any unusual objects within, it will take them a long time to do so, just as it takes a long time to find Wally in the scenes of the Where's Wally books. One has to move attention around the scene to recognise the objects, possibly one-by-one. The only scene-wide processing for unusual things is limited to a few features, like color and orientation.

These issues are also illustrated by the *wimmelbilderbuch* (literally "teeming picture book") tradition of illustrations popular in the 16th century. Also known as a hidden picture book, the pages contained many disparate figures and objects.

Only by attending to each part of the image in turn could you recognise all the objects that were there. Just as with Where's Wally today, children found them fun to look at. In the painting above, the artist Pieter Brueghel depicted people enacting idioms that children were meant to learn, such as "swimming against the tide", "banging one's head against a brick wall", and "armed to the teeth" (Net, 2021).

Most surprising things in a scene, then, won't attract attention because they aren't processed extensively. A secondary question, which we will turn to now, is which sorts of things will *hold* attention if they actually *are* processed.

16.2 Avoiding a bottleneck to test potential distractors

A common way that researchers use to ensure that objects are processed extensively is by presenting them one-by-one, right where the study participant is looking. This is called the **rapid serial visual presentation** method and we already saw it used to reveal a memory bottleneck in the previous chapter (15).

Asplund et al. (2010) used this technique to investigate what attracted attention in conditions where many different objects were briefly presented, but the participant's task was to attend to just one. They used a rapid stream of letters, like those in the movies in 15. In one of the Asplund et al. (2010) experiments, participants had to identify one of the letters in the stream. On some trials, however, without warning the participants, they replaced one of the letters with a photograph of a face.

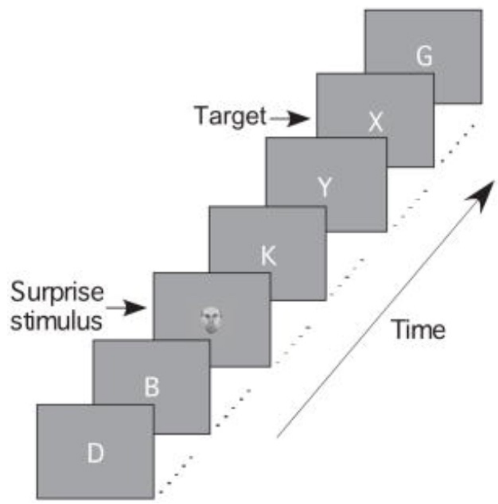


Figure 16.4: In the study schematised above, in 4 of 30 trials a surprise face stimulus popped up at various lags before a target letter.

The sudden appearance of the face was a surprise to the participants, at least at first. On those trials where it was included, performance identifying the letter often was worse. Specifically, if the letter was shown within about half a second after the face, accuracy reporting the letter was worse. The faces were proven to be effective distractors that held attention.

You might have guessed that faces would be potent distractors. After all, we are social beings, and as discussed in Section 10.5, given the chance, many people spend a lot of time looking at the people in a scene. However, Asplund et al. (2010) found that the disruption of performance by an unexpected image was not specific to faces.

In the experiment, half of participants were actually shown a stream of faces rather than a stream of letters.

The participants' task was to identify a target face, not a target letter, but in some trials, a letter was included in the stream. The results were quite similar to those for the participants who viewed a stream of letters rather than a stream of faces. That is, when the surprise letter was presented less than half a second

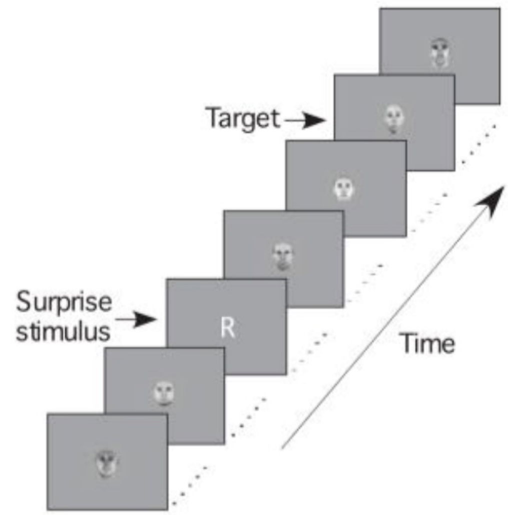
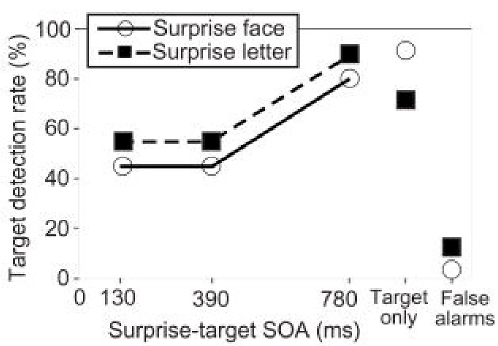
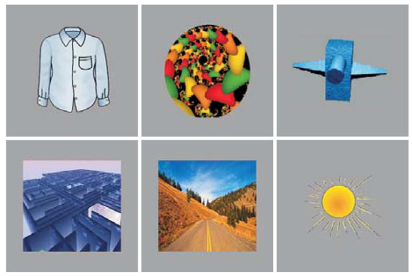


Figure 16.5: **Schematic of a stream of faces with a surprise letter included.**

before the target face, participants were less likely to report that they had seen the target face.



The black squares show the data for those participants. When in four of the trials a letter was included in the stream of faces, participants did poorly if that letter was presented 130 or 390 ms before the letter, but they did fine if it was presented 780 ms later. In summary, the odd stimulus disrupted performance for approximately half a second (500 ms). During that interval, participants missed the target about half the time.

The same pattern was seen for those participants who saw the stream of letters punctuated by the surprise presentation of a face, as shown by the white circles in the data plot. You will, however, notice that performance was slightly poorer for this condition, where the surprise was a face, which could be caused by faces being inherently more distracting than letters. To investigate this possibility, Asplund et al. (2010) tried using other images besides faces in the letter stream and compared their effects to those of the faces. 

The results yielded no significant difference between using any one of these images as a the surprise rather than the face, so although faces may have more distracting power (Langton et al., 2008), that didn't make much difference in these experiments. More important was simply that the stimulus presented was different rather than being a particular kind of stimulus.

When something distracts you, if it's something that interests you, you may simply choose to stop doing whatever you were in the process of doing. In that case, it's obvious that this would impair performance of the task you chose to stop doing. What's more interesting is when a distracter reduces your performance even though you are trying to completely ignore it. This is what was happening when the odd-colored items in the visual search experiment elevated response time (10.5) and in these Asplund et al. (2010) experiments with the images.

16.3 Surprise!

A major reason that attention was attracted to and held by the rare face presented among letters (Asplund et al., 2010), was because the face was simply different from the other stimuli presented in the trial. But another factor is the actual unexpectedness or surprise of seeing the face there's also another phenomenon contributing to the results. Remember that each participant saw the unusual stimulus in four of the thirty trials they participated in. In the first trial that the unusual stimulus appeared, it must have been very surprising, but not so surprising the third or fourth time. When Asplund et al. (2010) broke down the results in this way, they saw the results were different the third or fourth time.

Most of the deficit caused by the unusual stimulus was confined to the first one or two trials. After that, participants did much better. So a lot of the problem was the surprise itself. After participants began to expect the stimulus, the unusual stimulus had much less effect; participants were much less distracted by it.

A good deal of the effect of distractions, then, comes from them being unexpected. After a particular event has repeated many times, it has less effect

Figure 16.6: The first time one sees a jack-in-the-box, the sudden and unusual sensory event of the jack bursting out of the box can be highly surprising.

on reactions of many kinds, not just the kind of distraction assessed in these experiments. This is called *habituation.* Habituation, the decrease in response to a stimulus after repeated presentations, is a kind of learning. Neurons across much of the brain respond less to repetitions of an input than to the first exposure of an input. Habituation serves a variety of purposes for the brain, but here our point is just that it can reduce the orienting of attention to a stimulus.

Unfortunately, habituation is rarely complete. In the Asplund et al. (2010) work as well as the Theeuwes et al. experiment (5), even after many repetitions the average participant wasn't able to entirely prevent those salient stimuli from interfering with their performance.

In summary, there are two things going on in such cases:

- An unusual stimulus can attract attention, disrupting task performance for about half a second in these circumstances.
- If the stimulus is a surprise, the disruption is greater.

And don't forget, the stimulus has to be processed enough for it to be registered as unusual (which the researchers ensured by presenting them one-by-one).

Would it be possible to make a list of all the stimuli that, if processed, would attract attention? Perhaps not, because it may depend on the person. Have you ever been talking to a few people but then overheard someone else say your name in a separate conversation? When that's happened to me at a party, I've found it very difficult to keep concentrating on the conversation I was in, even if it was an important conversation with my boss. I might manage to keep my eyes on my boss, when he finished his sentence I'd realise that I didn't know what he had said.

Emotionally arousing stimuli are one stimulus class that may attract attention in most people. In the laboratory, one way that's been studied is by presenting a rapid series of images and testing people on which ones they remembered. By inserting different sorts of images in the stream, researchers can determine which ones are the most distracting. One study done by an honours student, Katherine Saunders, in my laboratory used that procedure. Watch this movie to see an example image sequence from Katherine's experiment.

After people were shown the movie, they were asked which of the below image(s) had been presented.

We found that participants were less likely to remember the couch than the other images. The reason for this is that it was placed soon after a very arousing image of a dead body.

Distractions like emotionally-charged images can cause people to completely miss pictures that are presented soon after them. Based on hundreds of experiments investigating this and related effects, researchers believe part of the reason is that consolidating sensory signals into memory takes time, as we described in the previous chapter (15). Once your attention is attracted by something, your mind will often process it for storage in memory, hampering your mind's ability to process other items into memory that occur around the same time. Unlike for the faces and letters used by Asplund et al. (2010), repeated presentation of such stimuli may not reduce their effect much (Onie et al., 2021); in other words, there may be very little habituation.

16.4 Learned and not learned

In the beginning of this chapter, and elsewhere (5), we referred to the importance of an alerting system in evolutionary history. Some of these alerts are probably hard-wired into our brains by evolutionary history, rather than learnt. Charles Darwin himself (Darwin, 2009) suggested this possibility, when writing about a trip he took to the zoo.

> I put my face close to the thick glass-plate in front of a puff-adder in the Zoological Gardens, with the firm determination of not starting back if the snake struck at me; but, as soon as the blow was struck, my resolution went for nothing, and I jumped a yard or two backwards with astonishing rapidity. My will and reason were powerless against the imagination of a danger which had never been experienced.

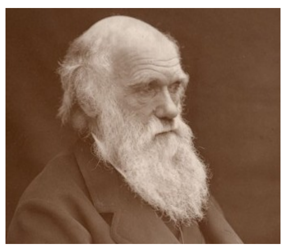


Figure 16.7: **Charles Darwin, soon after the 1872 publication of his book "The Expression of the Emotions in Man and Animals"**

Darwin believed that his fearful reaction to snakes was innate rather than learned. Subsequent evidence has provided some support for that idea, at least in other primates (Shibasaki and Kawai, 2009); it is difficult to confirm this in humans. Many of the things that attract our attention, however, must be learned. Evolution could not keep up with the changing array of important things in our ancestral environments, and our environment and associated lifestyles are changing even more rapidly today.

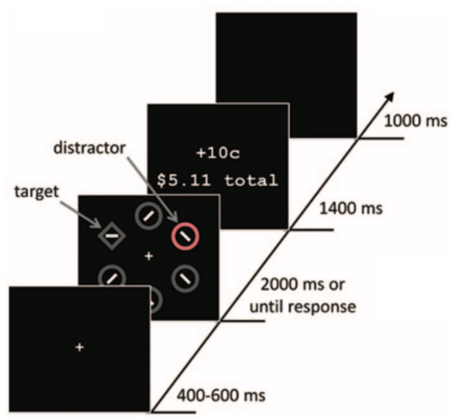
The habituation learning evident in the Asplund et al. (2010) work helps reduce the tendency for us to be continually distracted by things. However, the world today presents a constantly-changing set of distractions, so you will not habituate to everything, even if that were desirable.

16.5 Rewards make distractors more distracting

We have a dog in my household named Hugo. Most of the time, Hugo doesn't pay a lot of attention to me or to the other people in the house.

But when one of us starts shoveling wet dog food into Hugo's bowl, the impact of the spoon against the ceramic bowl makes a characteristic ringing sound, and Hugo will leap to attention. Hugo has learned to associate the ringing sound with the reward of food.

I think that even if Hugo were trying to concentrate on another task, the ringing sound would distract him. Something like that does happen in humans! For example, in a visual search experiment based on the study we discussed by Theeuwes (5), Le Pelley et al. (2015) asked participants to judge whether a little line they presented inside a diamond on each trial was horizontal or vertical. Le Pelley et al. (2015) also included an odd-colored object in each display, to distract the participants. There were two major changes compared to the Theeuwes study. First, the participants received money for each correct response. Second, they earned a greater amount (10¢) if the distractor were one particular color, while they made less money (1¢) if it were another.



The best strategy for participants to maximise their earnings was to completely ignore the distractor and its color. You already know from chapter 5 that people can't do that: the odd-colored distractor will slow them down. The new factor here was whether the distractor that indicated the reward would be 10¢ would affect people any differently than the distractor that indicated the reward for that trial would be just 1¢.

To end up with as much money as possible, if they can the participants ought to concentrate more on the target when the distractor indicates a large reward than when it indicates a small reward. However, this is not what happened! Instead, responses were *slower* when the distractor indicated the trial was high value than when it indicated the trial was low value.

Evidently, the participants attended more to the distractor that was associated with a stronger reward, even though this reduced how much money they received at the end of the experiment. Stimuli that are associated with greater rewards seem to attract more attention, even when that is completely counterproductive.

This brings us to the situation we all face with modern technology and devices.

16.6 Exercises

- What does the bottleneck on object identification have to do with what can distract you?
- What lesson did Darwin draw from his experience with a snake at the zoo?
- What do you find most distracting when you are trying to watch a lecture? What is most distracting when you are trying to study?

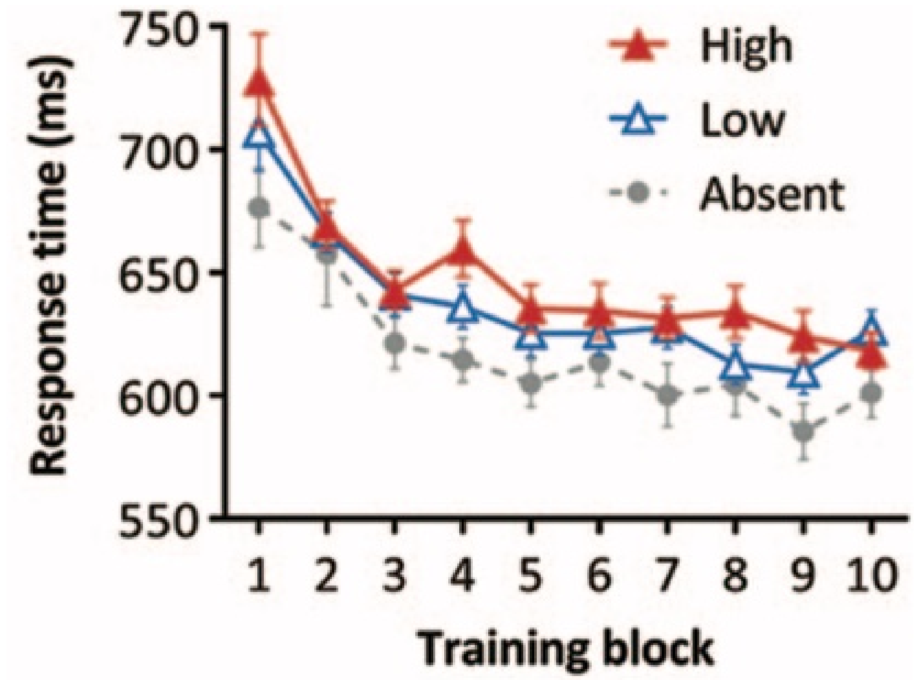


Figure 16.8: Response times were longest in the high-value condition, a bit faster in the low-value condition, and fastest when no distractor was presented.

Chapter 17

Distractions in the world

Humans aren't the only animals that get distracted.

17.1 Predators and prey

Various species of lizards can re-grow their tail if their tail gets cut off. Scientists are very interested in this - one reason is that it would be great if we could regenerate human limbs!

Figure 17.1: **This broad-tailed gecko was found in Sydney's northern suburbs. Like many other lizards, it can re-grow its tail.**

One curious fact is that some lizards are able to voluntarily release their tail when a predator is pouncing on them and, after the release, the tail starts flicking and even jumping around. So, the tail release mechanism seems to have evolved as a way for the lizards to save their ass, or actually lose their ass, but keep the rest of their body!

If a lizard is about to be completely eaten and it is able to detach its tail, leaving a predator with only the tail, as the rest of the lizard scurries away, the lizard's life has been saved. And the benefits of tail detachment may go beyond that. Lizards have often been witnessed to release their tail prior to the predator reaching them, whereupon predators have often been witnessed to go for the tail instead of the rest of the lizard.

Here are some candidate reasons why the predator goes for the detached tail:

- The tail itself is so worth eating that it's worth abandoning the uncertain chase to devour it. This is called the "consolation prize" hypothesis.
- The tail distracts the predator, but not because it's worth switching to and eating.

To investigate these two possibilities, two lizard researchers got on the case. By lizard researchers, I mean people who study lizards, not lizards who do research!

The researchers, Laura Naidenov and William Allen, reasoned that if the consolation prize hypothesis were true and the tail was considered by the predator a decent reward, then the longer the tail, the more likely the predator would be to attack the tail. The idea was that if the tail was considered desirable to eat, a bigger tail should be even more desirable.

If the tail simply distracts the predators from their target (the bulk of the lizard), the researchers reasoned that making the tail more visually conspicuous would increase the amount of engagement with the tail. To do so, the researchers varied how conspicuous the color of the fake lizard's tail was. One version of the tail was green, while the other version was a conspicuous blue. In nature, some lizards do in fact have tails that are brightly-colored and more conspicuous than the rest of their bodies.

Figure 17.2: **A Western skink, after it released its tail. Photo: William Leonard**

Naidenov and Allen created a fake lizard with a tail that could be detached by pulling a string. They brought their fake lizard to a local dog park, where they asked people whether they could borrow their dog for an experiment. If they said yes, they lured the dog into chasing the artificial lizard by using a string to pull it along the grass. Once the dog was about one meter from the fake lizard, the researchers released its tail. If at that point the dog lost interest in the chase, no data was recorded, but otherwise the researchers recorded whether the dog continued to chase and attack the body or alternatively attacked the tail.

The researchers ended up doing this to a total of thirty-four dogs, some of which participated in multiple conditions (short versus long fake tail length and blue versus green color). The results, as plotted in their paper, are below.

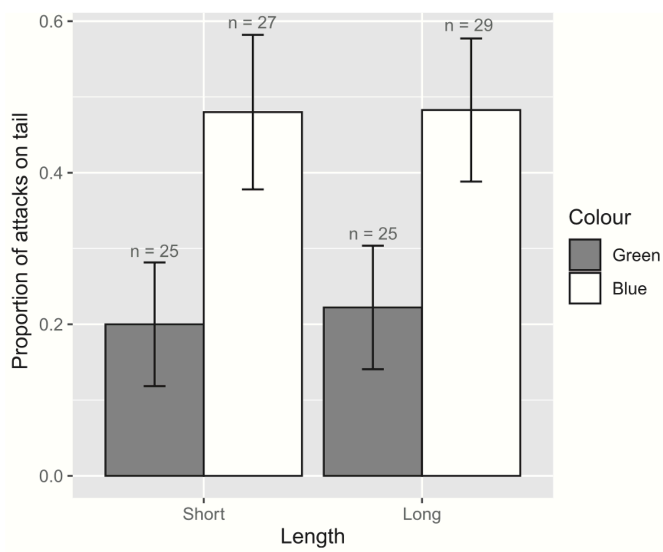


FIGURE 3 Proportion of attacks on the tail depending on the color and length of the tail. Error bars denote the standard error. Numbers above bars indicate the sample number

It doesn't look like length had an effect. Color, however, seemed to have a large effect. What can we conclude from this?

Many of us get a lot out of our phones and other devices. Whether it is scoring likes on Tik-tok or Instagram, seeing family updates on Facebook, or chatting with acquaintances on WhatsApp or Discord, or just getting an email from a friend, there is a constant stream of pleasant stimulation available to many of us. Thus, opening one of these apps can become strongly associated with a small feeling of reward. As the first president of Facebook put it, “It’s a social-validation feedback loop… Exactly the kind of thing a hacker like myself would come up with, because you’re exploiting a vulnerability in human psychology.” (Allen, 2017). The stream of small rewards keep us coming back to social media apps.

In 2003, a smartphone maker invented the push notification. Before then, you’d never know you had received a message or other update until you clicked on the associated app. But with the invention of the push notification, smartphone and other device developers added the capability for apps to pop up messages or sounds on your screen to indicate a new message. Later this was expanded to all sorts of updates, such as a “like” to one of your posts, or a share or retweets.

There was a tight association between these notifications and a small feeling of reward. This places the modern user of devices, then, in much the situation of those participating in the experiment by Le Pelley et al. (2015). When we are trying to concentrate on a task, a notification may pop up on our device. In part because of the association the notification has with reward, the notification is likely to distract us. Our performance on the task we are trying to concentrate on has been impaired. And that's the situation even when we don't click on the notification and go into the associated app.

Studies have shown that notifications and social media apps are very disruptive to people when they are trying to get things done. For that reason, psychologists and productivity gurus will tell you to turn off notifications in all your apps. Moreover, you should put your phone out of reach when you are working or studying or use an app to temporarily disable them if you have to use your phone. You can do the same thing on your computer - disable social media apps, Youtube, etc. with various programs that you can download to your computer.

At various times, I myself have found myself spending a lot of time on one social media site or app or another. It really helped to turn off all notifications to help me prioritise things. I still use social media, but now *I'm* the one controlling that rather than a tech company's algorithm.

Well, not entirely. I still find myself going to social media over and over again, lured by the possibility that a comment or 'like' might be there waiting for me. That is, while turning off notifications has prevented some bottom-up distractions (visual stimuli) from pulling me into activating a time-wasting app, it hasn't stopped my brain's top-down drive from continuing to push me in that direction.

Why does my brain have this drive? Social media site designers have structured things such that one often gets a small reward as soon as one opens their app. They have managed to get other users to frequently provide a little bit of social validation (a 'like'), or an interesting comment or post, that keeps me coming back for more, even if all the time I spend there doesn't add up to anything really.

17.2 Exercises

Chapter 18

Tutorial material

18.1 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?

18.2 Traffic lights

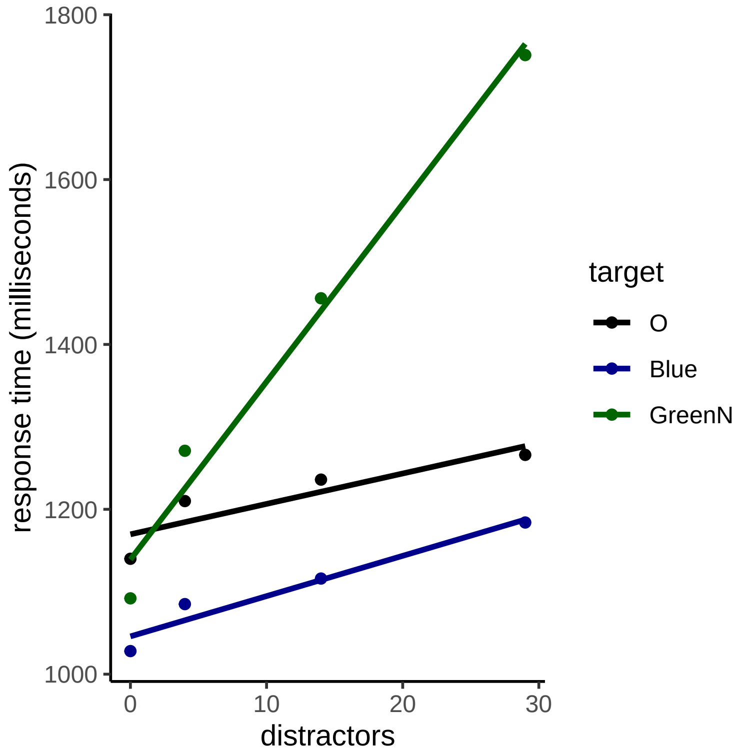


Figure 18.1: Results for a blue target (among differently-colored distractors), a green 'N' (among Ns of other colors and other letters that are sometimes green), and an 'O' (among other letters).

Figure 18.2: A Sydney suburban intersection. How many traffic lights do you see?

Bibliography

(2021). Netherlandish Proverbs. *Wikipedia*.

Allen, M. (2017). Sean Parker unloads on Facebook: "God only knows what it's doing to our children's brains". *Axios*.

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.

Asplund, C. L., Todd, J. J., Snyder, A. P., Gilbert, C. M., and Marois, R. (2010). Surprise-induced blindness: A stimulus-driven attentional limit to conscious perception. *Journal of Experimental Psychology: Human Perception and Performance*, 36(6):1372.

Block, N. (2011). Perceptual consciousness overflows cognitive access. *Trends in Cognitive Sciences*, 15(12):567–575.

Darwin, C. (2009). *The Expression of the Emotions in Man and Animals, Anniversary Edition*. OUP USA, Oxford ; New York, 4th edition edition.

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.

Goodbourn, P. T., Martini, P., Barnett-Cowan, M., Harris, I. M., Livesey, E. J., and Holcombe, A. O. (2016). Reconsidering Temporal Selection in the Attentional Blink. *Psychological Science*, 27(8):1146–1156.

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.

Langton, S. R., Law, A. S., Burton, A. M., and Schweinberger, S. R. (2008). Attention capture by faces. 107(1):330–342.

Le Pelley, M. E., Pearson, D., Griffiths, O., and Beesley, T. (2015). When goals conflict with values: Counterproductive attentional and oculomotor capture by reward-related stimuli. *Journal of Experimental Psychology: General*, 144(1):158.

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.

Onie, S., Donkin, C., and Most, S. (2021). Does a Single Session of Training Modify Emotion-Induced Blindness?

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.

Ransley, K., Goodbourn, P. T., Nguyen, E. H. L., Moustafa, A. A., and Holcombe, A. O. (2018). Reading direction influences lateral biases in letter processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

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.

Saiki, J. and Holcombe, A. O. (2012). Blindness to a simultaneous change of all elements in a scene, unless there is a change in summary statistics. *Journal of Vision*, 12:1–11.

Shibasaki, M. and Kawai, N. (2009). Rapid detection of snakes by Japanese monkeys (Macaca fuscata): An evolutionarily predisposed visual system. *Journal of Comparative Psychology*, 123(2):131.

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.

Suchow, J. W. and Alvarez, G. A. (2011). Motion silences awareness of visual change. *Current Biology*, 21(2):140–143.

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.

Treviño, M., Zhu, X., Lu, Y. Y., Scheuer, L. S., Passell, E., Huang, G. C., Germine, L. T., and Horowitz, T. S. (2021). How do we measure attention? Using factor analysis to establish construct validity of neuropsychological tests. *Cognitive Research: Principles and Implications*, 6(1):51.

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.

Wolfe, J. M. (2012). Saved by a log: How do humans perform hybrid visual and memory search? *Psychological Science*, 23(7):698–703.

Wolfe, J. M. (2020). Visual Search: How Do We Find What We Are Looking For? *Annual Review of Vision Science*, 6(1).

Wolfe, J. M., Võ, M. L.-H., Evans, K. K., and Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in cognitive sciences*, 15(2):77–84.