Week One

NOTES FROM LECTURE VIDEO

THE COGNITIVE REVOLUTION:

Back in the 1800s, scientists would show images and ask how a client was thinking and feeling. This method was used by Wilhelm Wundt, who established the first experimental psychology lab in 1879. And along with his student Edward Titchener, they studied conscious mental events, so things like feelings, thoughts, and perceptions. And the way that they studied these mental events was by asking people to report what they were experiencing. This technique is known as **introspection**.

There are a couple of problems with this idea. First of all, a lot of our mental activity occurs unconsciously. If these processes are unconscious, how are we supposed to tap into them when introspection is focusing on our conscious processes? And secondly, how people experience different events and what they report varies from person to person. Therefore, this is too subjective.

After introspective there was a bit of a shift to the opposite extreme. So instead of spending heaps of time prying into these mental processes and asking people about what they're thinking and feeling, the behaviourists (1900s) were led by John Watson. They only studied observable behaviours e.g. how people reacted in different situations.

One problem with this method is that we can't understand why someone's acted in a certain way unless we take into account their mental processes. So, if we see how someone responds, there might be a whole lot of stuff going on in their head that we are not accessing.

In the 1950s, computers were invented. Computers gave cognitive psychologists a measure to use for human information processing. So, we know that a computer encodes input or information. It stores this information and uses it, manipulates it, and produces some sort of output. The idea of this sort of information processing in stages was adopted by cognitive psychologists as a model of how human mental processes work. They replaced this computer in this diagram with a brain. They provided a useful metaphor for cognitive processes.
For example, if we gave people some input and it's a list of words (apple, book, chair). After giving them the list, we can ask them for an output. We can ask them to repeat those words back to us. When we look at memory, we will see that people are pretty good at repeating back a list of words, up until about 10 items. Without seeing these mental processes, we need to infer what we think might be happening. We think that people have some sort of storage space for the input that they receive. They can then hold these words there for a brief amount of time. We can make this inference based on the output that we get. This method of inferring what's happening was discovered in the 1700s by the philosopher Immanuel Kant. This is **Kant's Transcendental Method**. You start with observable facts (giving people a list of words) and looking at what they can remember about them and then inferring what's going on in the middle (some sort of memory storage space that holds onto those words for a brief amount of time) that leads to the output that we see (repeating those words back).

FORM PERCEPTION:

This is the Necker cube and it's an example of an ambiguous or reversible figure. It's ambiguous because it can be viewed in a couple of different ways. So first of all, you can see it with the blue panel at the front of the cube, so like in the middle picture.
The other way you can see it is with the blue panel at the back of the cube, like in the right-hand picture. If you look at the first cube in the series, you should be able to flick between those two different types of views.

Being able to see the cube in these two different ways tells us that our perception of the cube isn't neutral. In fact, you can perceive the cubes in different ways. If it was neutral, you would only be able to see it in one way. This means that our perception actually goes beyond the information that we've got available in the picture. Our perception allows us to specify an arrangement in depth, so whether it's with the blue panel on the front or the blue panel on the back.

These types of ambiguous and reversible figures are just one type of perceptual organisation that were identified by the Gestalt psychologists from Germany in the earliest 20th century. They also identified a number of other principles that influence how we perceive objects.

- The principle of similarity --> this suggests that we tend to group similar shapes.
- The principle of proximity --> we tend to group objects that are close together in space.
- The principle of continuity --> objects continue on. Are not broken up by other objects interrupting their sequence.
- The principle of closure --> intact or closed shapes.
- The principle of simplicity --> we perceive the simplest interpretation.

The Gestalt principles tell us that we organise information when we're perceiving it. We organise it in particular ways. When we see a particular stimulus, we detect its features so, for example, its straight lines, its corners, its curves etc. We then organise that information. We do this at the same time: perceive and organise. This is called **'in parallel'**.

Another feature of object recognition allows us to perceive perceptual constancy. When a door opens and we see it at different angles between open and closed, we still perceive it as the same object; the same door. Shape constancy, size constancy, brightness constancy, etc. We perceive these types of constancies because we're able to make unconscious calculations about what we expect to see. We make these observations based on our knowledge about the world. We take into account our viewing circumstances in a way that allows us to see the constant properties in our environments. Perceptual constancy can cause us to experience visual illusions.

OBJECT RECOGNITION:

Bottom-Up Processing:

- Looking at the features first → **Bottom-up processing**. Processing information from the smallest features to work out what this object is.
- **Top-down processing** gives you extra information. This occurs when your processing is influenced by your knowledge about the world.

- Knowledge about the world is driven by the context, affects how you perceive objects. These two types of processing occur in parallel.

WORD RECOGNITION:

To examine word recognition, experimenters have generally used what is called a tachistoscope procedure in which words flash very quickly on a computer screen.
Different factors affect how quickly we recognise words:

- Word frequency → words that appear more frequently and are more familiar to us, so things like words that we see every day, we actually recognise them quicker.
- Recency → how recently we've seen a particular word. Quite unsurprisingly, we're faster to recognise the word if we've seen the same word recently. If we saw an unrelated word, like 'pumpkin', before we saw the word 'elephant', we'll be slower to respond to 'elephant'. This is known as **repetition priming**, where the first time you see the word 'elephant', it primes your reaction to the second time that you see elephant. The first time you see the word 'elephant', your feature and your letter detectors get all excited, and your word detector gets activated. And because it's been activated, it has a level of activation just sitting there. It's been activated recently, so the next time you see it, it doesn't require much activation to fire. So that first sight of the word 'elephant' activates and primes your second sight of the word 'elephant', making you faster to respond. It occurs whenever we see one stimulus as a pause and then we see the same stimulus again. We are usually faster the second time we see the same thing.
- Word superiority effect → We're better at recognising letters when they're presented in words than when they are presented by themselves. When the letter is presented in a particular context such as a word, it makes it easier for us to recognise it later. We don't even need to have a real word to produce this effect. It just needs to be word-like or something like a word. However, this does not occur if the non-words aren't easy to pronounce or well-grouped. Here, **well-formedness** refers to how well a letter sequence conforms to the usual spelling patterns of English. E.g. FIKE or LAFE --> Produce the word superiority effect. HZYE or SBNE --> do not produce the word superiority effect.

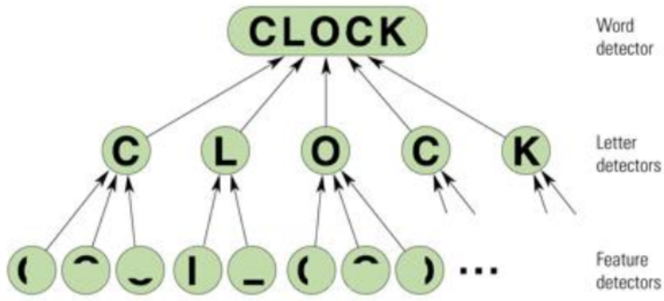
FEATURE NETS:

Feature nets are basically networks of detectors that are arranged in layers. Feature detector layer is how the model gets its name. The next layer up is concerned with larger-scale objects (letters). And then we move up again to even larger objects (words), to the final layer which is the word detector layer.
For example, the word "clock" is flashed extremely briefly on a computer screen, and then a mask is applied. A person would briefly see the features of that word, which would activate the feature detectors on the bottom layer. When these detectors have received enough activation, so when they get a strong signal when they get a good view of that particular word, they'll then fire and spread activation up to the next layer of the feature net. The letter detectors in turn fire and send activation up to the top level, the word detector, until a person recognises what the word is.

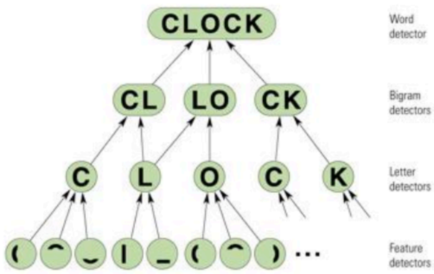
The amount of activation that's required for a detector to fire depends on the amount of activation that it has initially. When detectors have been fired recently, they have a higher activation level. Like they've already warmed up and they're ready to go again. Also, when detectors are fired frequently, they also have a higher activation level. This can help us explain **repetition priming**. We're faster and more accurate at recognising words we've just seen.

The first time we see that word, the activation starts at the feature detector, spreads up to the letter detector, and hits the word detector, we recognise the word. The second time we see that word, those detectors are already primed and ready to go from that first time we saw the word. So, the second time we see the word, we don't need as much activation to make it spread up the layers for us to detect the word.

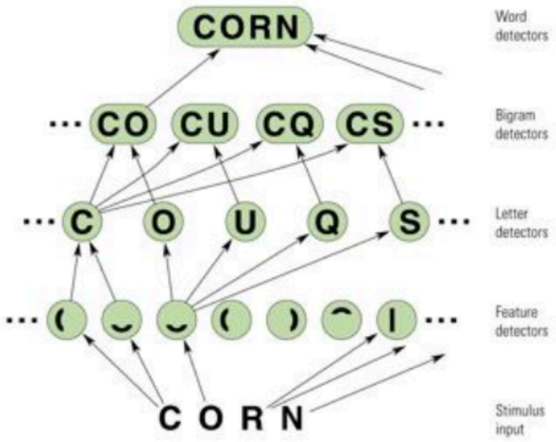
We can also explain the **word frequency effect** as we're faster to recognise words and we're more accurate at recognising words that we encounter more frequently than words that we encounter less frequently. This is because our detectors are warmed up. Faster to recognise frequently-occurring words.



To explain the effects of well-formedness, we actually need to add another layer to our feature net model. This model has actually an extra layer of bigram detectors, which is between the letter detectors and the word detectors. Bigram is simply two letters. When we add this layer, it helps us to explain why we can recognise English-like non-words, but not other types of non-words. When we get a familiar non-word like LAFE flashed on the screen, activation still flows up the feature net because we detect the features, we detect the letters, we detect the bigrams, and then we are able to detect the non-word. For the unfamiliar non-word HZYE, this requires a lot more activation because those letter combinations are so unfamiliar to us. So, we still get the feature detectors going off, we get the letter detectors going off, but then the bigram detectors need a lot more activation because we don't encounter the bigrams of HZ and YE often at all. Therefore, it takes a very strong signal for us to be able to recognise a non-English-like non-word.



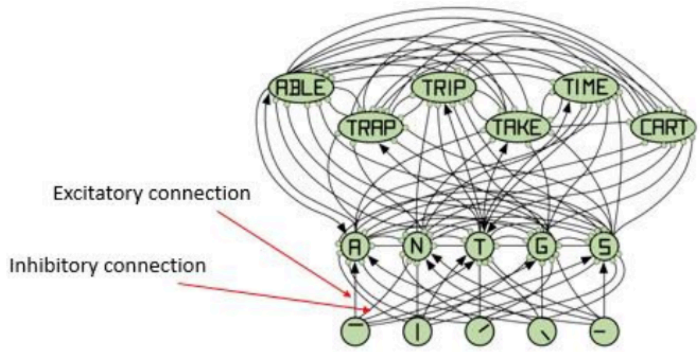
Let's say the word corn is briefly flashed and a mask is applied, so that not all of its features are detected. However, the feature detectors are activated by some of the curves and lines from the letters. E.g. the C's bottom curve and side curve have activated the feature detectors. For the O, the bottom curve has activated the feature detectors etc. These feature detectors fire and spread activation up to the letter detectors. Here, because the word was flashed so briefly, there may be some letters that weren't actually seen. So, the bottom curve of the O, for example, has activated not only the O in terms of the letters, but also the U, the Q, and the S letter detectors because they all have a similar lower curve. Now it's at the level of the bigram detectors where this issue of extra letters being detected or activated gets sorted out. A number of bigrams have been activated here. So, we've got the CO, the CU, the CQ, and the CS bigrams all being activated. However, the bigram detectors that then fire and send activation to the next level of word detectors are those that have been well primed. This is basically the CO detector because CO is most frequently seen in the English language. The CU detector is less well primed because it occurs less frequently. The CQ bigram detector shouldn't really exist because I don't think there are any words with CQ, and so that bigram detector won't fire at all. The CS one may fire weakly because occasionally, we see a word with CS in it. But the CO one should give the strongest signal and send the most activation up to the next level of word detectors, and the word "corn" is able to be recognised.



But this recognition at the bigram level can also lead to recognition errors, and in this example, for a non-word letter string. Imagine that the letter string, CQRN is briefly flashed. So again, not all the features are detected, but some of the feature detectors are activated by some of the curves in the lines of the letters, and they fire and spread activation up to the letter detectors. So again, let's say the bottom curve of the Q has been activated, and that sends activation up to the O letter detector, the U, the Q, and the S. However, when activation goes up to the bigram detectors, we get the same bigrams being activated as before. We get the CO, the CU, the CQ, and the CS. Now this is when the problem exists because activation is going to be sent up to the word detectors from the most frequently used bigram in the English language, which is the CO detector. The CU detector may weakly fire. The CQ detector is not going to fire at all. The CS bigram detector may weakly fire. The CO detector sends the activation up, and the person recognises the word "corn". In this case, that's an **error in recognition**. It's occurred because of the bigram detectors firing with the most frequently used bigram.

Another type of feature net is **McClelland and Rumelhart's Model**. One major difference between this model and the other one mentioned is that activation in this model not only spreads from one set of detectors to another, but the activation of a set of detectors can also inhibit or decrease the activation of another set. This wasn't present in the last model.

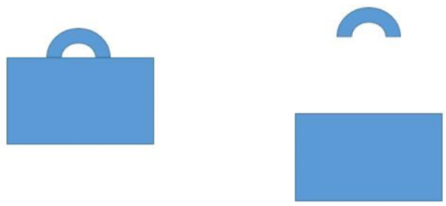
The current model also doesn't have a layer of bigram detectors. Some stimulus has been presented and five different feature detectors have been activated. So those are the five circles at the bottom of this model. There's a horizontal line, a vertical line, a couple of diagonal lines, and a shorter horizontal line. Now the first horizontal line will activate the letter detectors that have a horizontal line. And for example, if we look at the one on the left, this has activated the letter A. You can see the arrow pointing straight up to the A. It will also spread activation to the T, and the G, and the S because these also have straight lines. You can see the arrows going to these. These are excitatory connections. The activation will spread through to these letters. However, that horizontal line will not send activation to the N because this doesn't have a horizontal line. And in fact, it's going to decrease any activation in the N, and that's shown in the diagram by the line with the small circle. That's called an **inhibitory connection**. If we look at the fourth feature along, this will activate only one letter. That diagonal line will only activate the N due to the arrow only pointing at the N. It will then inhibit all of the other letters, which is shown by the lines with the small circles on them. The other difference with this model compared to the last one is that the flow of information in this model goes both ways. It goes from the lower-level detectors up to the higher levels, but it also comes down from the higher-level detectors to the lower levels. Takes into account the bottom-up processing and the top-down processing that occur in parallel.



McClelland and Rumelhart Model

Another type of network model is the **recognition-by-components model**. This one is a network model for recognising three-dimensional objects. The lowest level in this model is feature detectors, and then the next level up is geon detectors. So first we detect the features, e.g. the straight lines, the curves etc. and then we detect geons. These are simple shapes that make up all three-dimensional objects. An example here is that a suitcase is made up of a rectangle geon and a semicircle geon. And the highest level is an object model, so that's when you're recognising the word at the top of the previous feature nets.

It's important to note with this model that you can recognise objects from any angle. So, you can identify a chair from any angle etc. using the geons that make up that object. This type of recognition is what we call **viewpoint independent**. It doesn't matter what angle you're seeing the object from. You're able to detect the features, get activation going to the geons, which spreads up to the object level.



This next model is **viewpoint dependent**. And the recognition-by-multiple-views model suggests that we've got lots and lots of images stored in our memories, and these images show different objects from various angles. For example, let's pretend you have these 9 images of the teapot stored in your memory. If you see the teapot from another angle, then you actually need to mentally rotate one of these images that you've got stored in your mind to see whether it fits the object in front of you. If it does, then you can recognise that particular object. So, this multiple-view network model is viewpoint dependent.

FACIAL RECOGNITION:

There's evidence that suggests the way in which we recognise faces is different from the way in which we recognise words and other objects. So, faces have a different recognition system from what we've been talking about for words and objects.

People who suffer from **prosopagnosia**, which is the inability to see faces, can't recognise their parents, their friends, their partners, their children, or anyone. It might occur due to brain damage, or it might have been present since birth. But we know these people can recognise objects and words. So that suggests there are different systems for faces from objects and words.

The second line of evidence that we have for a different facial recognition system comes from what we call the **inversion effect**. Faces are much more difficult to recognise when they're inverted - upside down - than other objects are. Facial recognition is strongly dependent on orientation, because people are much worse when that face is upside down, whereas the recognition of other objects is not so strongly dependent on orientation.

It is thought that **faces are processed holistically**. So that means as a whole, which is quite different from the feature-based processing that we've been talking about. So instead of recognising someone just based on their nose, and then their eyes, and their mouth, and their eyebrows, it's their face as a whole that is recognised. This includes the spacing between features like their eyes, and the spacing between the nose and the mouth, etc. It's all those features taken together. Some of the evidence from this idea comes from the composite effect.

NOTES FROM CHAPTER ONE OF TEXTBOOK

THE BROAD ROLE FOR MEMORY:

The relevance of cognitive psychology is quite broad - thanks to the fact that a huge range of your actions, thoughts, and feelings depend on your cognition.

AMNESIA AND MEMORY LOSS:

Clinical amnesia includes cases in which someone, because of brain damage, has lost the ability to remember certain materials.

THE COGNITIVE REVOLUTION:

The science of psychology went through a succession of changes in the 1950s and 1960s that are often referred to as psychology's "cognitive revolution". This "revolution" involved a new style of research, aimed initially at questions we've already met; questions about memory, decision making, etc.

The cognitive revolution centred on two key ideas. One idea is that the science of psychology cannot study the mental world directly. A second idea is that the science of psychology *must* study the mental world if we're going to understand behaviour.

The Limits of Introspection:

In Wundt's and Titchener's view, psychology needed to focus largely on the study of conscious mental events - feelings, thoughts, perceptions, and recollections.

Wundt, Titchener, and their colleagues concluded, therefore, that the only way to study thoughts is through introspections, or "looking within", to observe and record the content of our own mental lives and the sequence of our own experiences.

Wundt and Titchener insisted, though, that this introspection could not be casual. Instead, introspectors had to be meticulously trained: they were given a vocabulary to describe what they observed; they were taught to be careful and as complete as possible; and above all, they were trained simply to report on their experiences, with a minimum of interpretation.

Psychologists gradually became disenchanted with this style of research. As one concern, these investigators soon had to acknowledge that some thoughts are unconscious, which meant that introspection was limited as a research tool. After all, by its very nature introspection is the study of conscious experiences, so of course it can tell us nothing about unconscious events.

The Years of Behaviorism:

Decided that psychology needs objective data, and that meant data out in the open for all to observe. First, an organism's behaviours are observable in the right way: you can watch my actions, and so can anyone else who is appropriately positioned. Therefore, data concerned with behaviour are objective data. Likewise, stimuli in the world are in the same objective category: these are measurable, recordable, physical events.

A person's beliefs, wishes, goals, preferences, hopes, and expectations cannot be directly observed, cannot be objectively recorded. These "mentalistic" notions can be observed only via introspection, and introspection has little value as a scientific tool. Therefore, a scientific psychology needs to avoid these invisible internal entities.

This perspective led to the behaviourist movement. The movement was in many ways successful and uncovered a range of principles concerned with how behaviour changes in response to various stimuli (including the stimuli we call "rewards" and "punishments"). By the late 1950s, however, psychologists were convinced that a lot of our behaviour could not be explained in these terms. The reason, basically, is that the ways people act, and the ways they feel, are guided by how they understand or interpret the situation, and not by the objective situation itself. Therefore, if we follow the behaviourists' instruction and focus only on the objective situation, we will often misunderstand why people are doing what they are doing and make the wrong predictions about how they'll behave in the future. The behaviourist perspective demands that we not talk about mental entities such as beliefs, memories, etc. because these things cannot be studied directly and so cannot be studied scientifically. Yet it seems that these subjective entities play a pivotal role in guiding behaviour, and so we must consider them if we want to understand behaviour.

The Intellectual Foundations of The Cognitive Revolution:
To use Immanuel Kant's (1724-1804) transcendental method, you begin with the observable facts and then work backward from these observations. This method, sometimes called "inference to best explanation," is at the heart of most modern science., Physicists, for example, routinely use this method to study objects or events that cannot be observed directly.

The Path from Behaviourism to the Cognitive Revolution:
In setting after setting, cognitive psychologists have applied the Kantian logic to explain how people remember, make decisions, pay attention, or solve problems. In each case, we begin with a particular performance and then hypothesise a series of unseen mental events that made the performance possible. We also ask whether some other, perhaps simpler, sequence of events might explain the data. We do more than ask how the data came about; we seek the best way to think about the data.

There have been some concerns about classical behaviourism, and some of those were voiced by Edward Tolman (1886-1959) - a researcher who can be counted both as a behaviourist and as one of the forerunners of cognitive psychology. Prior to Tolman, most behaviourists argued that learning could be understood simply as a change in behaviour. Tolman argued, however, that learning involved something more abstract: the acquisition of new knowledge.

In one of Tolman's studies, rats were placed in a maze day after day. For the initial 10 days, no food was available anywhere in the maze, and the rats wandered around with no pattern to their behaviours. Across these days, there was no change in behaviour - and so, according to the conventional view, no learning. But, in fact, there was learning, because the rats were learning the layout of the maze. That became clear on the 11th day of testing, when food was introduced into the maze in a particular location. The next day, the rats, placed back in the maze, ran immediately to that location. Indeed, their behaviour was essentially identical to the behaviour of rats who had many days of training with food in the maze (Tolman, 1948; Gleitman, 1963).

Across the initial 10 days, rats were acquiring what Tolman called a **"cognitive map"** of the maze. In the early days of the procedure, however, the rats had no motivation to use this knowledge. On Days 11 and 12, the rats gained a reason to use what they knew, and at that point they revealed their knowledge. Therefore, we need to talk about (invisible) mental processes (e.g., the formation of cognitive maps) if we want to explain behaviour.

B.F. Skinner (1904-1990) was an influential American behaviourist, and in 1957 he applied his style of analysis to humans' ability to learn and use language, arguing that language use could be understood in terms of behaviours and rewards (Skinner, 1957).

Two years later, the linguist Noam Chomsky (1928-present) published a ferocious rebuttal to Skinner's proposal, and convinced many psychologists that an entirely different approach was needed or explaining language learning and language use, and perhaps for other achievements as well.

European Roots of the Cognitive Revolution:
Overall, the Gestalt psychologists argued that behaviours, ideas, and perceptions are organised in a way that could not be understood through a part-by-part, element-by-element, analysis of the world. Instead, they claimed the elements take on meaning only as part of the whole - and therefore psychology needed to understand the nature of the "whole". This position had many implications, including an emphasis on the role of the perceiver in organising his/her experience.

Another crucial figure was British psychologist Frederic Bartlett (1886-1969). Although he was working in a very different tradition from the Gestalt psychologists, Bartlett also emphasised the ways in which each of us shapes and organises our experience. Bartlett claimed that people spontaneously fit their experiences into a mental framework, or "schema", and rely on this schema both to interpret the experience as it happens and to aid memory later on.

RESEARCH IN COGNITIVE PSYCHOLOGY - THE DIVERSITY OF METHODS:
In some settings, we ask how well people perform a particular task. For example, in tests of memory we might ask how complete someone's memory is (does the person remember all of the objects in view in a picture?) and also how accurate the memory is (does the person perhaps remember seeing a banana when, in truth, no banana was in view?). We can also ask how performance changes if we change the "input" (how well does the person remember a story, rather than a picture?), and we can change the person's circumstances (how is memory changed if the person is happy, or afraid, when hearing the story?). We can also manipulate the person's plans or strategies (what happens if we teach the person some sort of memorisation technique?), and we can compare different people (children vs. adults; novices at a task vs. experts; people with normal vision vs. people who have been blind since birth).

A different approach relies on measurements of speed. The idea here is that mental operations are fast but do take a measurable amount of time, and by examining the response time (RT) - that is, how long someone needs to make a particular response - we can often gain important insights into what's going on in the mind.

We can also gain insights from observations focused on the brain and nervous system. Over the last few decades, cognitive psychology has formed a productive partnership with the field of cognitive neuroscience, the effort toward understanding humans' mental functioning through close study of the brain and nervous system.

Information about healthy brains comes from neuroimaging techniques, which enable us, with some methods, to scrutinise the precise structure of the brain and, with other methods, to track the moment-by-moment pattern of activation within someone's brain.

NOTES FROM CHAPTER THREE OF TEXTBOOK (PAGES 80-92)

The reliance on parallel processing creates a problem of reuniting the various elements of a scene so that these elements are perceived in an integrated way. This is the **binding problem**. One key in solving this problem lies in the fact that different brain systems are organised in terms of maps, so that spatial position can be used as a framework for reuniting the separately analysed aspects of the visual scene.

Visual perception requires more than the "pick-up" features. Those features must be organised into **wholes** - a process apparently governed by the Gestalt principles. The visual system also must interpret the input, a point that is especially evident with reversible figures. Crucially, though, these interpretive steps aren't separate from, and occurring after, the pickup of elementary features, because the features themselves are shaped by the perceiver's organisation of the input.

The active nature of perception is also evident in perceptual constancy. We achieve constancy through a process of unconscious inference, taking one aspect of the input (e.g., the distance to the target) into account in interpreting another aspect (e.g., the target's size). This process is usually quite accurate, but it can produce illusions.

NOTES FROM CHAPTER FOUR OF TEXTBOOK

In some cases, patients suffer from **apperceptive agnosia** - they seem able to see an object's shape and colour and position, but they can't put these elements together to perceive the entire object. Other patients suffer from **associative agnosia**. They can see but cannot link what they see to their basic visual knowledge.

RECOGNITION - SOME EARLY CONSIDERATIONS:

Processes directly shaped by the stimulus are sometimes called data driven but are more commonly said to involve bottom-up processing. The effect of context, however, reminds us that recognition is also influenced by one's knowledge and expectations. This sort of influence - relying on knowledge - is sometimes called "**concept-driven**", and processes shaped by knowledge are said to involve **top-down processing**.

The Importance of Features:
Recognition might begin with the identification of visual features in the input pattern - the vertical lines, curves, diagonals, etc. With these features appropriately catalogued, you can start assembling the larger units.

The importance of features is also evident in data from visual search tasks - tasks in which participants are asked to examine a display and to judge whether a particular target is present in the display or not. This search is remarkably efficient when someone is searching for a target defined by a simple feature - for example, finding a vertical segment in a field of horizontals or a green shape in a field of red shapes. But people are generally slower in searching for a target defined as a combination of features.

Integrative agnosia derives from damage to the parietal lobe. Patients with this disorder appear relatively normal in tasks requiring them simply to detect features in a display, but they are markedly impaired in tasks that require them to judge how the features are bound together to form complex objects.

WORD RECOGNITION:

Several lines of evidence indicate object recognition does begin with the detection of simple features. Then, once this detection has occurred, separate mechanisms are needed to put the features together, assembling them into complete objects.

Factors Influencing Recognition:
Recency of view. If participants view a word and then, a little later, view it again, they will recognise the word more readily the second time around. The first exposure primes the participant for the second exposure; more specifically, this is a case of repetition priming.

The Word-Superiority Effect:
The advantage for perceiving letters-in-context is called the word-superiority effect (WSE). The WSE is demonstrated with a "two-alternative, forced-choice" procedure. For example, in some trials we might present a single - let's say K - followed by a post-stimulus mask, and follow that with a question: "which of these was in the display: An E or a K?" In other trials, we might present a word - let's say "DARK" - followed by a mask, followed by a question: "Which of these was in the display: An E or a K?"

For the word stimulus, both of the letters asked about are plausible endings for the stimulus; either ending would create a common word - DARE or DARK. Therefore, participants who saw only part of the display, perhaps DAR, couldn't use their knowledge of the language to figure out the display's final letter. In order to choose between E and K, therefore, participants really need to have seen the relevant letter - and that is exactly what we want.

Degree of Well-Formedness:
The term "word-superiority effect" may be misleading. If we present participants with letter strings like "FIKE" or "LAFE", although they are not English words and are not familiar, they look like English strings and are easy to pronounce. And, crucially, strings like these produce a context effect, with the result that letters in these contexts are easier to identify than letters alone.

Making Errors:
First, it seems that a letter will be easier to recognise if it appears in a well-formed sequence, but not if it appears in a random sequence. Second, well-formed strings are, overall, easier to perceive than ill-formed strings; this advantage remains even if the well-formed strings are made up ones that you've never seen before. All of these facts suggest that you somehow are using your knowledge of spelling patterns when you look at, and recognise, the words you encounter.

Misspelled words, partial words, or nonwords are read in a way that brings them into line with normal spelling. In effect, people perceive the input as being more regular than it actually is. Once again, therefore, our recognition seems to be guided by some knowledge of spelling patterns.

FEATURE NETS AND WORD RECOGNITION:

The Design of a Feature Net:
The idea is that there could be a network of detectors, organised in layers. The "bottom" layer is concerned with features, and that is why networks of this sort are often called feature nets. As we

move "upward" in the network, each subsequent layer is concerned with larger-scale objects; the flow of information would be bottom-up - from the lower levels toward the upper levels.

At any point in time, each detector in the network has a particular activation level, which reflects the status of the detector at that moment - roughly, how energised the detector is. When a detector receives some input, its activation level increases. A strong input will increase the activation level by a lot, and so will a series of waker inputs. In either case, the activation level will eventually reach the detector's response threshold, and at that point the detector will fire - that is, send its signal to the other detectors to which it is connected.

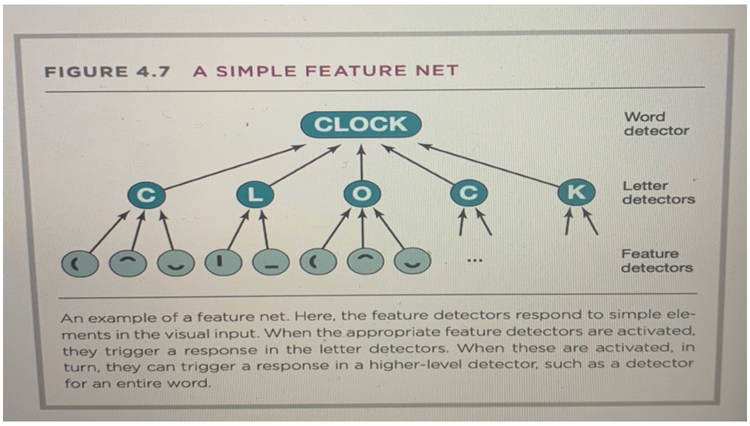


FIGURE 4.7 A SIMPLE FEATURE NET

An example of a feature net. Here, the feature detectors respond to simple elements in the visual input. When the appropriate feature detectors are activated, they trigger a response in the letter detectors. When these are activated, in turn, they can trigger a response in a higher-level detector, such as a detector for an entire word.

Within the net, some detectors will be easier to activate than others - that is, some will require a strong input to make them fire, while others will fire even with a weak input. This difference is created in part by how activated each detector is to begin with. If the detector is moderately activated at the start, then only a little input is needed to raise the activation level to threshold, and so it will be easy to make this detector fire. If a detector is not at all activated at the start, then a strong input is needed to bring the detector to threshold, and so it will be more difficult to make this detector fire.

The Feature Net and Well-Formedness:
In this situation, therefore, the network has partial information at the feature level (because only one of the O's features was detected), and this leads to confusion at the letter level: Too many letter detectors are firing (because the now-activated bottom-curve detector is wired to all of them). And, roughly speaking, all of these letter detectors are firing in a fashion that signals uncertainty, because they're each receiving input from only one of their unusual feature detectors.

The network's "knowledge" is not locally represented anywhere; it isn't stored in a particular location or built into a specific process. As a result, we cannot look just at the level of priming in the CO detector and conclude that this detector represents a frequently seen bigram. Instead, we need to look at the relationship between these priming levels, and we also need to look at how this relationship will lead to one detector being more influential than the other. In this way, the knowledge about bigram frequencies is contained within the network via a distributed representation; its knowledge, in other words, that's represented by a pattern of activations that's distributed across the network and detectable only if we consider how the entire network functions.

Efficiency versus Accuracy:
To maximise accuracy, you could, in principle, scrutinise every character on the page. That way, if a character were missing or misprinted, you would be sure to detect it. But the cost associated with this strategy would be intolerable. Reading would be unspeakably slow (partly because the speed with which you move your eyes is relatively slow - no more than 4-5 eye movements per second). In contrast, it's possible to make inferences about the rest. And for the most part, those inferences are

safe - thanks to the simple fact that our language contains some redundancies, so that one doesn't need every letter to identify what a word is; often the missing letter is perfectly predictable from the context, virtually guaranteeing that inferences will be correct.

DESCENDANTS OF THE FEATURE NET:

The McClelland and Rumelhart Model:
In the network proposal considered thus far, activation of one detector serves to activate other detectors. Other models involve a mechanism through which detectors can inhibit one another, so that the activation of one detector can decrease the activation in other detectors. The network by M & R is better able to identify well-formed strings than irregular strings; this net is also more efficient in identifying characters in context as opposed to characters in isolation. However, several attributes of this net make it possible to accomplish all this without bigram detectors.

Excitatory connections --> connections that allow one detector to activate its neighbours.
Inhibitory connections --> are shown by dots to deactivate its neighbours.

This model allows for more complicated signalling than we've used so far. We have assumed that lower-level detectors trigger upper-level detectors, but not the reverse. The flow of information was a one-way street, as presumed. But in this model, higher-level detectors (word detectors) can influence lower-level detectors, and detectors at any level can also influence other detectors at the same level (e.g. letter detectors can inhibit other letter detectors etc.)

Recognition by Components:
The McClelland and Rumelhart model was designed initially as an account of how people recognise printed language. But of course, we recognise many objects other than print, including 3D objects that fill our world - chairs, laps etc. A network theory was developed to recognise this feature network.

The recognition by components (RBC) model includes several important innovations, one of which is the inclusion of an intermediate level of detectors, sensitive to geons (short for "geometric ions"). The idea is that geons might serve as the basic building blocks of all the objects we recognise - geons are, in essence, the alphabet from which all objects are constructed.

Geons are simple shapes, such as cylinders, cones, and blocks, and according to Biederman (1987, 1990), we only need 30 or so different geons to describe every object in the world, just as 26 letters are all we need to spell all the words of English. These geons can be combined in various ways - in a top-of relation, or a side-connected relation, and so on - to create all the objects we perceive.

The RBC model also uses a hierarchy of detectors. The lowest-level detectors are feature detectors, which respond to edges, curves, angles, and so on. These detectors in turn activate the geon detectors. Higher levels of detectors are then sensitive to combinations of geons. More precisely, geons are assembled into complex arrangements called "geon assemblies," which explicitly represent the relations between geons (e.g. top-of or side-connected). These assemblies, finally, activate the object model, a representation of the complete, recognised object.

The presence of the geon and geon-assembly levels within this hierarchy offers several advantages. For one, geons can be identified from virtually any angle of view. As a result, recognition based on geons is viewpoint-independent. Thus, no matter what your position is relative to a cat, you'll be able to identify its geons and identify the cat.

Recognition via Multiple Views:

A number of researchers have offered a different approach to object recognition. They propose that people have stored in memory a number of different views of each object they can recognise. According to this perspective, you'll recognise Felix as a cat only if you can match your current view of Felix with one of these remembered views. But the number of views in memory is limited - maybe a half dozen or so - and so, in many cases, your current view won't line up with any of the available images. In that situation, you'll need to rotate the current view to bring it into alignment with one of the views in memory, and this mental rotation will cause a slight delay in the recognition.

The key, then, is that recognition sometimes requires mental rotation, and as a result it will be slower from some viewpoints than from others. In other words, the speed of recognition will be viewpoint-dependent. Data indicate that recognition is faster from some angles than others, in a way that's consistent with this multiple-views proposal.

In this proposal, there is a hierarchy of detectors, with each successive layer within the network concerned with more complex aspects of the whole. Thus, low level detectors respond to lines at certain orientations; higher-level detectors respond to corners and notches. At the top of the hierarchy are detectors that respond to the sight of whole objects. It is important that these detectors each represent what the object looks like from a particular vantage point, and so the detectors fire when there is a match to one of these view-turned representations.

There has been a lively debate between advocates of the RBC approach (with its claim that recognition is largely viewpoint-independent) and the multiple-views approach (with its argument that recognition is viewpoint-dependent). And this may be a case in which both sides are right - with some brain tissue being sensitive to viewpoint, and some brain tissue not being sensitive. Moreover, the perceiver's task may be crucial. Some neuroscience data suggest that categorisation tasks ("Is this a cup?") may rely on viewpoint-independent processing in the brain, while identification tasks ("Is this the cup I showed you before?") may rely on viewpoint-dependent processing.

FACE RECOGNITION:

Faces Are Special:

Damage to the visual system can produce a disorder known as agnosia - an inability to recognise certain stimuli - one type of agnosia specifically involves the perception of faces. People who suffer from prosopagnosia generally have normal vision. They can look at a photograph and correctly say whether the photo shows a face or something else, whether the person is a man or a woman, and generally their age as well. However, they can't recognise individual faces - not even their own parents or children, whether in photographs or right in front of them. They can't even recognise themselves (and so they sometimes think they’re looking through a window at a stranger when they're actually looking at themselves in a mirror).

Often this condition is the result of brain damage, but in some people, it appears to be present from birth, without any detectable brain damage. Prosopagnosia seems to imply the existence of special neural structures involves almost exclusively in the recognition and discrimination of faces. Presumably, prosopagnosia results from some problem or limitation in the functioning of this brain tissue.

The special nature of face recognition is also suggested by a pattern that is the opposite of prosopagnosia. Some people seem to be "super-recognisers" and are magnificently accurate in facial recognition, even though they have no special advantage in other perceptual or memory tasks. These people are consistently able to remember and recognise faces that they viewed only briefly at some distant point in the past, and they're also more successful in tasks that require face matching - that is, judging whether two different views of a face actually show the same person.

There are some advantages to being a "super-recogniser" such as being able to remember faces for a politician or a sales person; super-recognisers also seem to be much more accurate as eyewitnesses (e.g. In selecting a culprit from a police line-up). On the downside, however, being a super-recogniser can produce some social awkwardness. Imagine approaching someone and cheerfully announcing, "I know you! You used to work at the grocery store on Main Street." The other person (who, say, did work in that grocery store eight years earlier) might find this puzzling, perhaps creepy, and maybe even alarming.

Face recognition is strongly dependent on orientation, and so it shows a powerful **inversion effect**. Moreover, with non-faces, the (relatively small) effect of inversion becomes even smaller with practice; with faces, the effect of inversion remains in place even after practice.

Holistic Recognition:
Face recognition does not depend on an inventory of a face's parts; instead, this process seems to depend on holistic perception of the face. In other words, face recognition depends on the face's overall configuration - the spacing of the eyes relative to the length of the nose, the height of the forehead relative to the width of the face, etc.

Evidence suggests that in recognising familiar faces, you rely more heavily on the relationships among the internal features of the face; for unfamiliar faces, you may be more influenced by the face's outer parts such as the hair and the overall shape of the head.