# 10

# Is the whole greater than the sum of its parts? From gestalt psychology to artificial intelligence

The question addressed in this chapter, the last chapter in this book, is in some ways the most important question in psychology. It is important because an answer ‘yes’ can be used to justify the existence of psychology as a discipline separate from physiology. It is important because the difference between those assuming yes and those assuming no fuels a controversy that runs throughout the history of psychology. It is important because artificial intelligence and therefore the future of humans is linked to the assumption of ‘yes.’ It is important because artificial intelligence can act as heuristic for a new type of theory in psychology, a type of theory that would create a paradigm shift from the current neurocognitive paradigm – assuming that this *is* the current paradigm, which by now the reader will realise is not assumed by all psychologists!

There are two ways of understanding the world: analysis and synthesis. The process of analysis examines the different parts of something. After analysing something into its separate ‘bits,’ the whole is found by adding the separate bits together to make a whole. In the physical sciences this often involves micro-analysis – i.e., finding the smaller and smaller parts, for example, the genes in biology or fundamental particles in physics.

The process synthesis shows how the different parts of something interact and combine such that the whole cannot be understood by the separate contributions of the individual parts. Throughout the history of science, and in several different disciplines, there is a tension between those who want to analyse yet smaller and smaller parts of the jigsaw that is knowledge, and those who want to see the meaning of the jigsaw as a whole.

The coalescence versus brick wall hypothesis was introduced in Chapter 2. If the taste of lemonade can be understood as the combination of two separate tastes, lemon and sugar, then the sensation of lemonade can be analysed into its separate components. If, on the other hand, lemonade is a unique taste, different from lemon and sugar, then lemonade is a synthesis of sugar and lemon and must be understood as such. The brick wall versus coalescence controversy was a debate amongst philosophers in the 19th century, and it is a precursor to a much more important movement in psychology, the gestalt movement.

## The gestalt movement

The gestalt movement was never associated exclusively with one university or person, nor was it associated with a particular type of psychology. Instead, the gestalt movement represents an idea that predates the beginning of academic psychology and continues to be important (Ellis, 1999). The assumption underlying the gestalt movement can be summed up in the phrase:

> The whole is greater than the sum of its parts.

The word *Gestalt* is a German noun meaning shape or configuration – it does not translate easily into English. The use of this noun reflects the idea that psychological shapes or structures are more than the sum of their parts.

Phenomena that are now described as gestalt phenomena were discovered in the middle of the 19th century, long before the gestalt movement was recognised. The Necker cube was drawn by Necker in 1832 and the old woman/young girl illusion appeared on a German postcard in 1888. The significance of the gestalt movement is that an explanation was provided for phenomena that were previously interesting illusions. The gestalt movement provided theory.

The beginning of the gestalt movement is attributed to a chance observation made by Max Wertheimer (1880–1943). In 1910, Wertheimer was travelling by train from Vienna while on holiday, and noticed how an illusion of movement was created by telegraph wires along the track. This illusion is familiar to many train travellers – the wires seem to move up and down. The idea of apparent movement was not new. There was a 19th-century toy called a stroboscope that exploited this phenomenon, and Lumière invented the Cinématographe in 1895. Wertheimer realised that the phenomenon of apparent movement could not be explained by the then current theory of movement perception that had been proposed by Wundt (see Chapter 2) (Wertheimer, 1912). Wundt thought that apparent motion was caused by the summation of eye movements – i.e., you perceive movement because the eyes move. Wertheimer felt and was eventually able to show that movement perception was something that 'more than the sum of the parts' of any elements observed by the observer (Ash, 1995).

Wertheimer started to investigate this apparent movement, or the *Phi* phenomenon as it was called, using himself and two colleagues as observers. Along with Wertheimer, these two colleagues, Wolfgang Köhler (1887–1967) and Kurt Koffka (1886–1941), are looked on as the founders of the gestalt movement.

Wertheimer used an instrument called a tachistoscope that enabled images to be presented on a screen at precisely defined times and for precise durations. Tachistoscopes were standard psychology experimental equipment before the introduction of computers. As an example of an experiment done by Wertheimer, imagine lines placed at slightly different angles, and which we can call line A and line B, and which can be shown independently on a tachistoscope screen or on a computer (Figure 10.1).

First, line A is shown on the screen. Then line A disappears, and line B appears. Then line B disappears and line A reappears again, and so on.

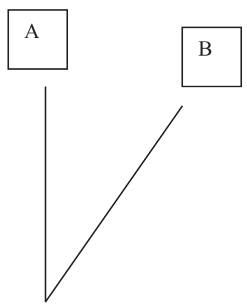


**FIGURE 10.1** Line A appears and then disappears and is replaced by line B, then vice versa.

Wertheimer was able to vary the rate of alternation between line A and B. He found that as the rate increased, there came a point when it appeared that a line was moving between A and B and *something* was actually visible between them. The line appears to be rotating around the point where the two lines meet. In fact, the same experiment is more easily demonstrated by two lights in a dark room that are separated by a short distance. If the light alternates between A and B, it appears that the light is moving between A and B and *something* is visible between A and B. The demonstration of apparent motion is interesting but was not novel. Wertheimer's important contribution, which was published in 1912 (Wertheimer, 1912), was a modification of this type of *phi* experiment.

Wertheimer used three lines, two of which are called A, and one B (Figure 10.2).

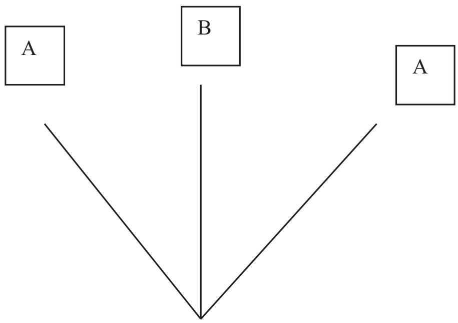


**FIGURE 10.2** Line B appears then disappears and is followed by lines A and vice versa.

As before, the A and B lines alternate on the screen, but in this case the two A lines appear at the same time. The B line appears, then the two A lines appear as the B line disappears, then the A lines appear again and the B line disappears. This arrangement creates the peculiar impression that the B line moves in two directions at once – it looks as though the B line splits and moves outwards. Wertheimer realised that this impression could not be caused by eye movement, because *the eyes cannot move in two directions at the same time*. His demonstration therefore showed that Wundt's explanation of how we perceive motion was false. Motion is not perceived because of the way the eye moves, but is due the overall pattern of what we see. Why do we see movement from the overall pattern? The answer is that perception follows certain rules, so that when a pattern of a particular kind appears, then we perceive according to that rule.

The idea that perception follows rules or principles led over a short period of time to the discovery of several gestalt principles. Students studying psychology will already be familiar with these principles from the study of visual perception. They include:

Figure-ground: Several examples of these exist, including the face-vase illusion and young-old woman illusion.

Continuity: There is an assumption that lines that are interrupted are continuous (a principle exploited in the 'three rope trick' (search online and you will find videos showing you how to do this).
Proximity: Objects close together are assumed to be connected.
Similarity: Objects that are similar are assumed to be connected.

Gestalt principles also are used in colour perception. The colours of the rainbow are well known: red, orange, yellow, green, blue, and violet. So where is brown? Where is gold? Where is silver? Gold and silver are the easiest ones to explain. Gold is yellow and silver is light grey with particularly bright highlights (highlights are the reflections that come off shiny objects). If you examine a painting of a gold shield, you will see that from a distance the shield really looks gold, but when you come up close it is simply yellow. The trick is that artists puts in more highlights than are in the surrounding objects so that you assume that the object must be shiny gold. Colours such as gold and silver show that people infer colour not only from the colour impinging on the eye but from the whole of the visual array.

Brown is a little more difficult to understand. If you have paints that are red, orange, yellow, green, and blue next to each other, they look like those colours – i.e., red orange, yellow, etc. However, if you take a colour such as yellow, and reduce the amount of light that is emerging *relative to its surroundings*, then this 'yellow' looks brown. Brown (there are lots of different browns) is the colour that we infer from knowing that relatively little light is being reflected – of course, when no light is reflected, then we assume the colour is black. Colour perception shows that colour is not just a matter of an addition of signals from the cones in the eye but is an inference that is made from the total array of the visual image.

## Gestalt learning phenomena and other gestalt psychologists

Although much of the early research and phenomena labelled as 'gestalt' was associated with perception, the idea of a gestalt obviously has a wider application than perception. The perceptual focus of the early gestalt psychologists was because most early psychology was perceptual, and the gestalt psychologists were proposing a new, holistic way of understanding the psychology of the time. Gestalt psychologists were also interested in learning.

Wolfgang Köhler was one of Wertheimer's collaborators in his early experiments (Henle, 1971). Köhler's later work focussed on additional topics including human values (Köhler, 1947, 1959, 1966). Köhler investigated insight learning in chimpanzees (Köhler, 1925), working on the island of Tenerife at the German Primate Research Centre. He conducted a famous study with chimpanzees. The study took various forms, but in one form there is banana outside the chimpanzee's cage which is just out of arm's reach. A stick is placed in the cage with the chimpanzee. After a time, the chimpanzee will pick up the stick and hook the banana inside. Repeated study shows, for example, that the chimpanzee picks up the stick faster if the stick is between the chimpanzee and the banana rather than at the other end of the cage. The significance of Köhler's studies was to show that learning happened in an 'all-or-nothing' fashion. At one point in time the chimpanzee did not know what to do and at the next point it did. This finding was important because it contradicted the more atomistic theories of learning that were being developed by animal psychologists working within the behaviourist tradition (see Chapter 4) where learning was believed to be a gradual process of the strengthening of stimulus-response bonds. Köhler's research showed that learning could be all or nothing.

## America

Koffka moved to America in 1924. Wertheimer, who was Jewish, moved to America in 1933 when the German National Socialists dismissed all Jewish professors from German Universities – including Nobel prize winners such as Albert Einstein. Köhler published an article in 1933 strongly criticising German discrimination against Jews (Henle, 1978). Köhler left Germany for America in 1935. The three men were friends throughout their academic lives (Brett & Wertheimer, 2005).

Kurt Koffka moved to America before his German friends. He was proficient in English before he left (Koffka, 1922, 1924, 1935) and his proficiency in English enabled him to promote gestalt principles in the English-speaking world (Gibson, 1979). He influenced Tolman (see Chapter 4) and befriended James J. Gibson (1904–1979), whose book *The Ecological Approach to Visual Perception*, published in 1979, is viewed as an important culmination of the gestalt approach. One of Gibson's many ideas was that, when people look at a visual array, they immediately register what objects are *for*. So, for example, if a person sees something that has the structure of a path, that person will know that the path is for walking on. The shape of a hammer will indicate that the object is for hitting things with. Gibson described these 'what things are for' as affordances.

> A *path* affords pedestrian locomotion from one place to another. . . . An *obstacle* can be defined as an animal-sized object that affords collision and possible injury.
>
> (Gibson, 1979, p. 36)

So an affordance is a gestalt – it is the whole that is perceived directly.

## Field theories

Many theories in psychology take the logical form of A causes B. The experimental method is suited for assessing such causal statements. Causal statements are important because they have practical implications. Does spaced learning help a student revise more than massed learning? Do adverse childhood experiences affect behaviour in later life? Does drinking alcohol on a regular basis lead to greater happiness? Causal theories of this kind take the form of

S → O → R

Or more generally

Situation/stimulus/environment → person variable/mechanism/processing → response/behaviour/mental state

However, any causal relationship is just one of many, many different causal relations that are happening at any one time. The assumption of the gestalt movement is that the whole is greater than the sum of its parts. That being the case, it also makes sense to consider how all the many, many different causal relations that are happening at one time interact to produce some kind of outcome. Not only is perception and learning holistic, it also may be that one needs holistic theories. Field theories are based on the assumption that theories should explore how all the different causal relations occur together. In this section two field theories are examined, the field theory of Lewin and inter-behaviourism.

Lewin (1890–1947) was Jewish and like Kofka moved to America in 1933, as did other Jews who had university careers and had to leave Germany. Lewin was familiar with gestalt psychology and refers to this in his field theory. The following quote sums up the essence of Lewin's field theory:

> Whether or not a certain type of behavior occurs depends not on the presence or absence of one fact or of a number of facts as viewed in isolation but upon the constellation (structure and forces) of the specific field as a whole.
>
> (Lewin, 1939, p. 889)

Lewin suggests that behaviour is the result of different forces that result from a person's "life space" where life space consists of all the different experiences and motives that make up a person. Lewin suggests that behaviour can be considered a path through this life space, and he gives a name to this space, *hodological space*. Lewin was influenced by newly developed mathematics of topology and he invented symbols to represent the various concepts in hodological space. In his most detailed publication of his theory (Lewin, 1938), he provides four pages of symbols that are used in the diagrams which he created to represent hodological space. This type of representation looks very scientific!

Although field theory naturally stems from gestalt principles, field theory was also developed by a behaviourist, Jacob Kantor (1888–1984). Kantor had worked with Skinner (see Chapter 4) whom he admired, but felt that the interactional nature of behaviourism was not sufficiently developed (Morris, 1984). Kantor proposed a form of behaviourism called inter-behaviourism (Kantor, 1970). Kantor suggested there was a bidirectional causal connection between stimulus and response, instead of the single-directional, stimulus-causes-response approach taken by other behaviourists. Kantor's approach is described as ecological behaviourism. He suggested that the stimulus-response interaction takes place in a context and it is the context that is neglected in other behaviourist approaches to understanding behaviour.

What these different field theories have in common is the idea that it is not possible to isolate the organism and study it separately from its natural environment. It is the interaction with the natural environment that is important. The idea of understanding behaviour in terms of a number of simultaneous causal interactions is found not only in field theories of psychology, it is the most basic assumption in the study of ecology. Ecologists study the mutual interactions of species in the total environment. Some argue that experimental studies fail to provide a true picture of the mechanism's underlying behaviour because they lack ecological validity. The term *ecological validity* is based on the assumption that a valid – i.e., true – account of human behaviour requires investigation in its natural setting, not in the laboratory

Field theories can be considered an extension of ideas first proposed by the gestalt psychologists. Sharps and Wertheimer sum up the contribution of the gestalt movement to modern psychology as follows:

> Several points for modern psychology emerge from the Gestalt perspective. Phenomena should be studied within their full context; there is a need to acknowledge the domain specificity of principles in experimental psychology; it is wise to study phenomena that either exist in the real world or have close real-world analogues; psychology must recognize interchanges between organisms and surroundings as determinants of behaviour.
>
> (Sharps & Wertheimer, 2000, p. 315)

Field theories in psychology are intuitively plausible. So why are they practically unknown? Students reading this book could ask fellow students whether they have heard of either Lewin or Kantor. Lewin and Kantor suggested *how* psychology theory should be constructed. The problem is they never demonstrated that this type of theory was able to solve problems that could not be solved by other existing theories – despite Lewin being committed to the practical use of psychology. The history of psychology is one where different people have advocated different types of psychology. If all that is done is to recommend a particular type of theory, then this advocacy eventually disappears. What is needed is a demonstration that the particular approach or theory can provide an answer not provided elsewhere. That is the reason this book has given examples of how particular approaches can be used in practice (for example, that of conversational analysis in the previous chapter). Field theories were never used in practice, despite their intuitive plausibility. They are important because the problem of multiple simultaneous causes was eventually solved with the development of connectionist psychology and artificial intelligence.

## Why do emergent properties occur?

What is the difference between a living organism and a dead one? This question is at the root of why emergent properties occur. Several explanations were provided for the difference between living and inanimate objects during the 19th and early 20th centuries. A popular one was that living organisms have a 'vital force' that leaves them when dead. This idea of a vital force was used by Mary Shelley in her book *Frankenstein*. Shelley proposed that this vital force came from a newly discovered force, the force of electricity.

The psychologist William McDougall (1871–1938) had a number of interests outside psychology. In his last book, *The Riddle of Life* (McDougall, 1938), McDougall reviews several explanations for the difference between living and inanimate matter. These explanations included animistic ones (such as that used by Mary Shelley) which McDougall dismisses as being unscientific. One of the explanations favoured by McDougall is that put forward by Max Loewenthal in his book *Life and Soul*, which was published in 1934. Although Loewenthal strays into animistic theories, which McDougall rejects, Loewenthal makes the prescient suggestion that living organisms are different from non-living matter because of organisation. The reason for this hypothesis is that snails can be frozen to –120 °C when they appear dead but, if carefully thawed, come back to life. As a 'life force' or any form of energy cannot exist at such low temperature, Loewenthal suggests that life must have something to do with structure, a structure that is independent of temperature. The nature of this structure was not known.

It is now known that it is a particular type of structure that makes life possible. It is the same network structure that forms the basis for artificial intelligence (AI) and a type of structure that gives rise to a parallel distributed processing (PDP) system. A PDP system is one where processing is occurring over the whole network at the same time. It is the existence of network structure and parallel distributed processing that makes emergent properties possible, including the emergent property of life. Research and understanding of networks can therefore be seen as the solution to a problem that the gestalt psychologists failed to answer: How exactly are patterns recognised as a whole?

Artificial intelligence.

## Artificial intelligence (AI)

There are a number of key steps in the discovery of AI, and two are described here, the work of the behavioural psychologist, Donal Hebb (1904–1985) and the collaboration between the neuropsychiatrist Warren McCulloch (1898–1969) and mathematician Walter Pitts (1923–1969).

Donald Hebb published his book, *Organization of Behaviour: A Neuropsychological Theory* in 1948. Hebb was interested in discovering what happened in the brain when learning took place, in particular the type of learning known as classical conditioning. Hebb knew that the brain was a network of inter-connected neurones, and came up with a simple proposition, a proposition now known as the Hebbian learning rule. The Hebbian learning rule is that if two neurones fire simultaneously, then the connection between those two neurones strengthens. The rule is sometimes stated by the slogan:

> Neurones that fire together get wired together.

Hebb showed that networks can adapt or learn if they follow rules. Several other rules were later developed to show how networks could learn.

**Different types of leaning rule**

A distinction is made between unsupervised and supervised learning rules. The Hebbian learning rule is unsupervised in that there is no external agent giving feedback. Other rules can explain 'supervised learning' where the network learns to recognise patterns on the basis of feedback from a 'teacher.' Supervised learning explanations are used in several applications of AI, including pattern recognition and behavioural adaptation.

McCulloch and Pitts developed their ideas through the inter-disciplinary collaboration of neuroscience and mathematics (Abraham, 2002). Inter-disciplinary developments can be some of the most important in the development of science. The McCulloch-Pitts model is based a simple assumption (McCulloch & Pitts, 1943). Suppose there is a network of neurones that are either 'on' or 'off.' Then assume that this on-off state of a neurone corresponds to true-false in a logical argument. McCulloch and Pitts showed that variation in the connection strengths between the neurones in a network could produce on-off or true-false states that followed the rules of logic. The contribution of their approach was to show that networks could *solve problems*, simply by the connection strengths of the neurones.

The two pieces of the jigsaw, from Hebb and from McCulloch-Pitts, were put together in what become known as connectionist psychology (Ellis & Humphreys, 1999). A network system can learn and in doing so solve problems. In adapting to their inputs, networks can learn to function better in their environment. Connectionist psychology and artificial intelligence are based on the same assumption. The assumption is that in order to understand or mimic the intelligence of a human, it is necessary to have a structure of a network where multiple, simultaneous causal connections occur. Pattern recognition is one of the achievements of AI, and the way an AI system recognises a pattern provides an answer to the question that the gestalt psychologists were never able to answer, how patterns are recognised.

How is a pattern recognised? How do we recognise the letter A when handwriting differs so much between people? One way would be to have a series of templates, of different kinds of A. For example, the following are the letter A using different fonts

Any new letter could be compared with a storage bank of letters A to see whether or not the letter was an A. The problem with the template approach is that it will recognise the letter A only if it has been seen in that format before. The templates here would fail to recognise a letter A that has never been written before, but which you, the student, will recognise as a letter A.

Humans – and machines that mimic humans – are able to recognise the letter A in a format that has not been seen before. Pattern recognition requires an understanding of the relationship between the elements of the pattern – the relationship between the lines that make up a letter A, for example.

Pattern recognition devices, i.e., devices capable of recognising a letter even if that form of the letter has not been seen before – use neural network structure similar to that shown in the Figure 10.3.

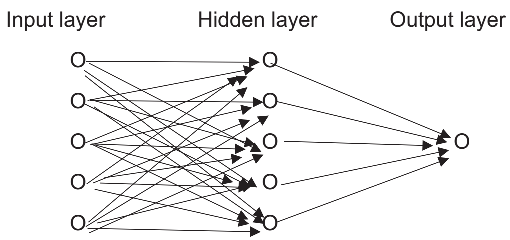


**FIGURE 10.3** In this figure, there are three layers of neurones, input, hidden, and output.

The input layer receives inputs from the external environment. Each input node is either on or off depending on its input from the external environment. Figure 10.3 shows only five input nodes, but there can be many more. The input nodes could also be arranged in a two-dimensional array connected to detectors that detect whether a point in a pattern is either black (i.e., a line passes through that point) or white (a line does not pass through that point). Just one input layer is shown in the figure, but additional hidden layers can be added one after another, all connecting with each other.

The neurones in the output layer – just one is shown in the Figure 10.3, but there could be more – provide the answer, for example whether the pattern is an A or not. Figure 10.3 shows that information is processed in parallel – hence a parallel distributed processing system, rather than a linear system where one logical process is followed by another. The perceptive student will ask, "but how *exactly* does this parallel distributed processing system provide the correct answer?" It does it by getting the correct strengths of connection between the all three layers, input, hidden, and output. The strengths of connection can vary and once they are *exactly* right, then the system will correctly recognise patterns such as the letter A. The perceptive student will point out, "yes, but how do you get the connection strengths *exactly* correct? The diagram looks incredibly complicated and it looks like an impossible task." The answer is that a human does not work it out. Instead the answer is provided from a type of machine learning, called supervised learning.

Supervised learning occurs when a network adapts on the basis of feedback. First, the network needs to be trained. Training is achieved by presenting the network with letters of the alphabet in different fonts, some of them being As, and providing the network with information about whether the letter is an A or not. Then, if the network adapts using a 'supervised learning rule' or algorithm, then the connection strengths will gradually change so that eventually they are just right to be able to detect whether a letter is an A or not. There are several different possible supervised learning rules, but they all provide a way of gradually changing the connection weights until they provide the correct solution. So long as the network adapts – i.e., the connection strengths change – in accordance with the learning rule, then, after training, the network will be able to correctly identify the letter A, whatever form that A takes. This form of learning is also referred to as 'deep learning.'

Supervised learning or deep learning shows an important feature of artificial intelligence. After training, a machine will be able to perform tasks that it has not been programmed to perform. The human does not have to write the rules for task performance. The machine learns those rules by itself. The machine is capable of learning something that the human has not been able to make. The science of artificial intelligence was motivated by a desire to create a machine to do what humans are capable of doing. However, in creating artificial intelligence, scientists have been able to demonstrate *how* humans do what they are capable of doing.

### Example of a learning rule

Supervised learning rules are based on mathematical rules and are beyond the scope of this book. Here is one example, called the back-propagation method. First, calculate the difference between the actual and desired output and multiply this by some function of the inputs to that output unit. This calculation creates an error term. Then, using that error term, adjust the weights of the connections from the penultimate layer in the network to the output unit. Repeat this procedure between the penultimate units and the layer behind them, and then repeat back through all the other network layers back to the input layer.

The human ability to recognise patterns, an ability demonstrated by the gestalt psychologists, is possible because the brain uses a network structure. Information is encoded in the connections between the nodes in the network – as Hebb originally suggested. When people see a hammer, they realise that it is used for hitting, and when they see a road, they realise that it is used for walking because actions – i.e., affordances – are just some of the many connections in the network of information that is encoded in the brain. It is the network structure of the brain that gives humans (and other animals) the capacity to process, understand, and respond to their environment in the way they do.

## Parallel versus sequential processing machines

Artificial intelligence is now part of everyday life and will become more and more important as time passes. Phones and tablets provide predictive text, where each device learns its own user's use of language. Predictive text is made possible because of a network architecture. Life is made possible because of a network architecture. However, phones and tablets are not alive. As everyone knows, iPhones do not grow up and become iPads. What is going on?

Computers, tablets, and phones are sequential processing machines. They have the same basic structure that they have had since the very first computers. They perform tasks one after another. However, modern computers are very, very fast. Because of their incredible speed, they can do the calculations that are needed to *simulate* a parallel processing machine. Parallel processing is complex, but is simply a matter of mathematics. The sequential processing machine is sufficiently fast that it can give the appearance of doing the same tasks as a parallel processing machine.

Your phone, tablet, and computer are not alive, and if robots of the future are simply extensions of the same processing structure of modern computers, then the robots of the future will not be alive. If, however, it becomes possible to make a truly parallel processing machine in the way that humans are parallel processing, then one can certainly raise the question as to whether the machine is alive and should be granted the rights of 'wet' organisms. Who knows what the future will bring? As a student watching episodes of *Star Trek*, I thought the communication devices were an impossibility. Mobile phones are now taken for granted.

## Modular versus connectionist psychology

Cognitive psychology, like any other part of psychology, developed on the basis of assumptions. The assumption of cognitive psychology is that the information is processed in the brain in the same way that it is processed in a computer – a sequential processing computer. This assumption forms the basis of Fodor's modularity approach to the mind. Information is processed in modules that have different functions (see Chapter 6). Each process is completed in a module before the next one takes place.

Connectionist psychology provides a challenge to the modularity theory of mind (Massaro, 1990). Examination of this challenge falls outside the remit of this book, but here is an example why connectionist psychology is a new conceptual approach to information processing.

Suppose I ask you the question, "what is big, red, and found in London?" You may guess straight away, "a London bus." How do you do this so quickly? What you do *not* do is search through a module of all big objects, a module of all red objects, a module of all things in London, and find the one object that is common to all three. A computer could do it this way because a computer is incredibly fast, but it would take a human too long. People immediately realise "a London bus" because information isn't organised in modules. Instead, information is organised as a network of concepts, where concepts are nodes connected to other concept nodes with connections of varying strengths. The word *red* is a node connected to everything that is red. The word *big* is connected to everything that is big, and the word *London* is connected with everything in London. When I ask you "what is big, red, and found in London?" what happens is that the connections from all three nodes become activated and the *combined activations* all link to one concept, that of a London bus.

Modern pattern recognition devices use the same logic. They look for connections not for templates in modules.

## Networks, emergent properties, and the reduction of psychology to physiology

Networks are intuitively difficult to understand because so many causal relations are occurring at the same time. It is easy to understand A causes B causes C. But it is not easy to understand simultaneous cause between multiple nodes in a network. One way of getting an intuitive understanding of networks is to consider a flock of birds. Birds form flocks that fly through the air, often in beautiful patterns. Have you ever watched a flock of birds and wondered how it happens? How is it that different species of birds form flocks in different shapes? It looks magical. Magic is simply a mechanism that is not understood. If you don't want the magic taken out of flocks of birds, skip the next paragraph.

The shape of a flock of birds is an emergent property. It is achieved because each bird follows a rule. The rules are slightly different between species, but they all involve some kind of copying of the behaviour of neighbouring birds. The shape and behaviour of the flock are achieved by all the elements of the network following the same simple rule. The shape of the flock of the birds is an emergent property. It cannot be predicted from the behaviour of a single bird, only from the behaviour of the flock as a whole. If the brain (and body) is a network system, then there will be emergent properties that cannot be reduced to individual neurones. The network structure and its consequent emergent properties therefore provides an explanation for why psychology cannot be reduced to physiology. Psychology depends on physiology, in the same way that a flock of birds depends on the individual birds themselves, but there are psychological properties that cannot be predicted from physiology – assuming, of course, that the underlying structure is one of a network.

**Clocks versus flocks**

A mechanical clock (the sort with six cogwheels) is a sequential processing system. A flock of birds is a parallel processing system. If one cog in the clock is damaged or missing, then the entire clock stops functioning. If one bird in a flock of birds drops dead, then the flock continues without interruption. Sequential processing systems are sensitive to local error. Network systems are not. The fact that the brain has some, but not infinite, plasticity when damaged might suggest that there is some modularity and some connectionism in the way the brain works.

## Robotics and psychology

There was a fundamental weakness of the field theories described earlier. These theories presented an intuitively correct idea about simultaneous multi-causality but did not develop useful theories, i.e., theories that were able to solve applied problems (see Chapter 3). What about psychology and robotics? Robots are particularly interesting because, unlike the computer that sits on your desk, robots can engage in spontaneous behaviour. Robots may be driven by AI and deep learning, but their ability to behave makes them more human-like than a computer. Psychologists have already made contributions to the understanding of human-robot interactions. The 'uncanny valley' hypothesis proposes that there is a non-linear relationship between the human realism of a robot and likeability (Mathur & Reichling, 2016). People prefer robots that are recognised as non-human or robots that can be mistaken for humans. Robots that are almost human (i.e., imperfect humans) are considered creepy.

Psychologists can contribute to the development and use of robotics, but what about the other way round? Can an understanding of robotics help in the development of psychology? The adaptive network theory is an example of how AI and robotics acts as a heuristic for a new type of theory, with practical application in helping people with somatic and psychological symptoms.

## The adaptive network theory: an example of how AI can help theory development in psychology

The adaptive network theory is based on the assumption that symptoms are the result of biological symptom-causing mechanisms, but the mechanisms form a network of nodes with connections between the nodes. This network of nodes has emergent properties and so should be understood not in terms of biology but in terms of information in a network (Hyland, 2011, 2017). The information network is an adaptive system and receives inputs from information that is psychologically mediated and biologically mediated (as suggested by biopsychosocial interactionism, see Chapter 6).

In addition to having the structure of a theory consistent with biopsychosocial interactionism, the adaptive network theory assumes that symptoms and emotional states have the function of guiding behaviour. This additional assumption is not novel. William James suggested it over a hundred years ago (see Chapter 3). Symptoms adapt the body to the current environment and one of their functions is to inhibit behaviour. The following is a list of symptoms and how they are adaptive in inhibiting behaviour.

Pain – prevents injury.
Fatigue caused by infection or injury – enhances healing through resting.
Sadness – prevents continuation of unsuccessful activity.
Anxiety – prevents exposure to danger.

The novel part of the adaptive network theory is the proposal that the body adapts when symptoms fail to have their intended function, and it adapts by increasing the intensity of the symptom. If a person doesn't listen and respond to what the body is trying to tell them, then the body shouts louder. A similar idea had been presented some 30 years previously. The control theory of depression suggests that failing to disengage from unattainable goals leads to depression (Carver & Scheier, 1990; Hyland, 1987). The weakness of the control theory explanation of depression is that it did not explain *why* persistence in unattainable goals leads to depression. The science of AI shows how this is possible. Developments in AI show that if a control loop is part of an intelligent, adaptive system, then this is precisely what will happen. The parameters of the control loop will adapt if control is not being achieved.

The adaptive network theory provides a new paradigm for understanding symptoms. Instead of viewing symptoms as the result of a fault that is either psychological or biological, the theory presents symptoms as an adaptation which, from the patient's perspective, has gone wrong. The theory proposes that non-adaptive psychological and biological states creates 'stop signals' in an information network. The stop signals create symptoms that inhibit recurrence of the non-adaptive states, for example, by creating short-term symptoms of pain, fatigue, sadness, or state anxiety. The stop signal encodes the instruction to 'stop doing whatever you are doing' and the stop signal and symptom disappears once the behaviour stops. However, when stop signals and associated symptoms repeatedly fail to stop behaviour, doing so over a period of time, then the stop signals potentiate and become fixed to form a stop programme. The stop programme encodes the instruction 'whatever you are doing is unhelpful so stop doing things,' and the stop programme creates potentiated and persisting symptoms such as unexplained pain, severe fatigue, depression, and trait anxiety. The body has adapted in a way that inhibits any behaviour. A schematic representation of the theory is shown in Figure 10.4.

There are several reasons why people will persist in activities despite stop signals and symptoms that inhibit behaviour. People can 'keep going' despite

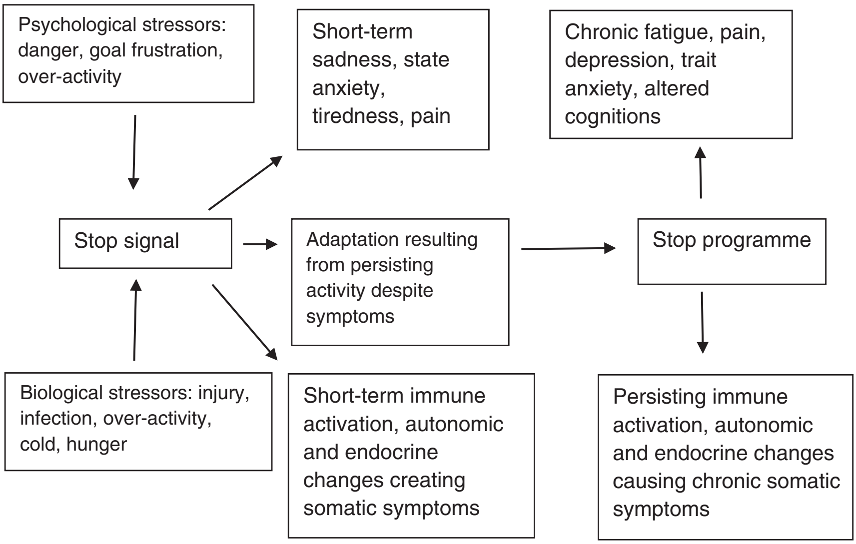


**FIGURE 10.4** Schematic representation of the adaptive network theory. Stop signals encode the instruction to 'stop doing whatever it is that you are doing at the moment.' The stop programme encodes the instruction to 'stop everything that you are doing.'

pain and fatigue because of social obligations, because of work obligations, because of educational goals, because they find something very interesting, or because they find themselves in a situation where persistence is the only reasonable response to a difficult situation. The message to students is the same one that arises from the theory of reactive inhibition (see Chapter 4). Listen to what your bodies are telling you. Don't work too hard!

The theory can inform the treatment of any of the symptoms shown in Figure 10.4 (symptoms that tend to covary), but is currently being used only to treat unexplained symptoms of functional disorders (Hyland, 2017) where biological and psychological interventions have limited success (Henningsen, Zipfel, & Herzog, 2007).

**Treatment**

'Body reprogramming' is the name given to the educational intervention based on the adaptive network theory (Hyland et al., 2016). The aim is to explain to patients why they became ill and how they can help themselves by engaging in behaviours that do not create stop signals, see www.bodyreprogramming.org. Patients are taught how to teach their bodies that the world is a safe and rewarding place using evidence-based lifestyle change such as relaxation, stress avoidance, positive experiences, exercise, and nutrition.

The adaptive network theory provides a historical account of the development of symptoms in terms of past events, and therefore adds to existing models of illness. For example, Beck's theory of depression (Beck, 1967) is based on the hypothesis that depression is the consequence of erroneous cognitions, but does not explain the historical antecedents to those erroneous cognitions. The adaptive network theory shows how erroneous cognitions can be one of the outputs from a network that has adapted to a particular lifestyle. The adaptive network theory adds to the kindling hypothesis of depression and other stress-health theories as it predicts that the effects of stressors depend on how those stressors are interpreted within an intelligent system. Consequently, predictions based only on severity and frequency of stressors may be poor predictors of the health consequences of stress (Monroe & Harkness, 2005).

The adaptive network theory can be used to illustrate many of the principles covered in this book. First, science proceeds best by bold conjectures or hypotheses, but the hypotheses must be tested (see Chapter 1), because bold conjectures may be false. It is the lack of testing that was a problem with psychoanalytic theories (see Chapter 5). A therapy can be effective for any one of many possible reasons, so evidence that therapy is effective provides only weak evidence to support a theory on which the therapy is based (see Chapter 5).

There is evidence consistent with the predictions of the adaptive network theory that is separate from effectiveness of therapy (Hyland, Lanario, Wei, Jones, & Masoli, 2019; Melidis, Denham, & Hyland, 2018), but evidence merely corroborates and never proves a theory. The existing data may be explainable by another theory, and future data may falsify the theory. The adaptive network theory is an example of a theory being used to solve an applied problem (see Chapter 3), but its relative usefulness compared to other theories is yet to be evaluated. The theory combines the idea of adaptation and information processing (see Chapter 4) and provides a paradigm where humans are understood in terms of a mechanism similar to that found in robots. Some may argue that this paradigm is a travesty of what it means to be human, where humans have spiritual needs (Chapter 8) and where deep learning is no substitute for deep meaning (Chapter 9). Others may argue that in the long run developments in neuroscience will provide a better understanding (in what sense better?) of the phenomena explained by the adaptive network theory. The theory is a biopsychosocial theory in that symptoms result from an interaction between have psychological and biological causes. That interaction taking place in an information level consistent with the mechanistic explanations of cognitive psychology (see Chapters 4 and 6), but the usefulness of this approach in contrast to a purely physiological or purely psychological one may prove in the long term to be deceptive. The original hypothesis was informed by observation and qualitative information from patients (see Chapter 8), but failure to do a proper qualitative analysis is a shortcoming in any understanding of the phenomenon of functional disorders. Finally, any theory, however ardently supported by its inventor (as this one is) has a good potential for being wrong. Science advances only when conjectures are found to be false and new conjectures proposed.

### Exam anxiety

Many students experience exam anxiety. According to the adaptive network theory, exam anxiety is caused by a history of exam taking. Each time a child takes an exam, they are engaging in a behaviour that creates anxiety. Their bodies are telling them to stop doing things that signal danger, but they keep on putting themselves in (what the body interprets as) danger by taking exams. Social and other pressures cause people to ignore what their bodies are telling them to do. The adaptive network theory predicts that the degree of exam anxiety in students should be correlated with the number of exams taken previously – a prediction consistent with observation (McDonald, 2001).

The history of psychology shows that new technologies can provide the opportunity for new paradigms in psychology. The original aim of AI was to create machines capable of simulating the intelligence of humans. Those machines can now do more than simulate what humans can do. The adaptive network theory uses AI to develop a new way of thinking about how the body functions as an adaptive system that is simultaneously biological and psychological. The original aim of AI is reversed. Instead of modelling machines on humans, theories to explain human behaviour can be modelled on what AI machines do. Whether this approach will lead to a new paradigm in the future is impossible to tell. One things that the history of science shows for certain: Paradigms do not last forever. Science changes as it advances. Psychology will change.

## Summary

The idea that 'the whole is greater than the sum of its parts' was suggested by gestalt psychologists using perception as an example of this phenomenon. Later the idea was applied to learning to show that learning was (or could be) all or none rather than incremental. This idea of holism influences field theorists and underpins the idea of ecological validity. What happens in the laboratory may not happen in real life. Finally, the emergence of network theory in the 1940s and 1950s led to connectionist psychology in the 1980s and the start of the new technological revolution, that of artificial intelligence.

The development of factory production lines coincided with rise of behaviourism. The sequential processing computer coincided with the rise of cognitive psychology. Parallel processing, the basis of AI, is a new technology that could be more revolutionary than the other two technological revolutions. Connectionist psychology is based on the assumption that concepts form the nodes of a network and the psychological processes reflect an underlying network architecture. The adaptive network theory is based on the assumption that the body is a network with emergent properties and that particular patterns of lifestyle lead to adaptations that create symptoms.

The discipline of psychology has changed over the 150 years of its history, and will continue to change. When historians of psychology look back in 50 years' time, what will they see?

## Essay questions

1 **What were the main achievements of the gestalt movement?**
2 **What is machine learning and how has machine learning and artificial intelligence contributed to psychology?**
3 **Scientific paradigms change. Behaviourism lasted about 50 years, and cognitive psychology and humanistic psychology about the same length of time. What do you think the discipline of psychology will be like 50 years from now?**