PART I

SETTING THE STAGE

1

Talkin' Bout a Revolution: A Paradigm Shift

The all-powerful king Louis XVI of France, grandson of the sun king of "L'état, c'est moi" ("The state, that's me") fame, held absolute autocratic power unbounded by a constitution or any other laws curbing his might. That all changed on 14 July 1789, a day that changed the course of world history. Events unfolded that gave Louis XVI his nickname of Louis The Last. The skyline of Paris was dominated by the eight high towers of the Bastille, a symbol for the complete and unquestioned authority of the French monarchy. The royal fortress and prison held only seven prisoners but the angry crowd that had gathered before the building had not come for them. They were after one thing: gun powder. When the prison governor refused to hand it over, the mob stormed the building and took it by force. The governor was pulled away, insulted and spat on. After periodic beatings, his body was showered with daggers and bayonets, riddled with pistol shots, his head sawed off with a pocket knife, and then paraded on a pike round the streets of the royal capital. Louis XVI and his wife Marie Antoinette were eventually guillotined in Paris in the bloody years that followed Bastille Day, now the national day of France.

The old order had collapsed. Europe's monarchs watched in horror and panicked. The unimaginable had happened. The French revolution is remembered for the tens of thousands of people who were executed or murdered, for the hundreds of thousands of people who died in the Napoleonic wars that followed it, but also for the radical political and social transformations that it gave rise to. The French revolution is widely celebrated for ringing the death bells for feudalism in Europe, for the ideals of liberty, equality, and justice it inspired, and for the 'Declaration of the Rights of Man and of the Citizen' granting what we now call basic human rights to (some) common people. There is a consensus among historians that a complex combination of economic, social, and political factors provided the breeding ground for the revolution. Yet no one had seen it coming.

The sudden rupture of the French revolution altered how intellectuals thought about historical change. The word revolution before the dramatic events in Paris was restricted to refer to the cycles of celestial bodies such as the stars and planets seen from Earth. A revolution denoted a cycling back to a previous situation or condition. After the momentous events in Paris, the term revolution acquired the new meaning of a radical, typically sudden, and unexpected departure from an old order in favor of a new system, approach, and world view[1].

The understanding that change in science is also rarely anticipated and rather sudden came a whopping 150 years later. American physicist and philosopher Thomas Kuhn was the first to notice that science does not progress on a linear and continuous trajectory the way lay people typically imagine the scientific accumulation of knowledge to occur. His book '*The structure of scientific revolutions*'[2] has become one of the most widely cited works in the social sciences. Kuhn's insight was to describe scientific progress as a process of periodic paradigm shifts during which an established conceptual framework about a scientific problem is abruptly abandoned in favor of a new one: A scientific revolution occurs when, during a relatively short period of time, large numbers of researchers adopt a new way of thinking.

Behaviorism and the cognitive revolution
A classic example for a Kuhnian paradigm shift in psychology, the study of human behavior, is the cognitive revolution which took place in the 1950s. The dominant theoretical approach in psychology in the first half of the 20th century, especially in the United States, was behaviorism. On a fundamental level, behaviorism was a very good and rigorous psychological science. It was all about pairing a stimulus with a behavioral response, with the idea that only an immediately observable response to an observable stimulus could safely be interpreted. Behaviorism[3] rejected dualism, the notion that the physical body and mental phenomena, including a soul, could be separated. The soul, but also any mental phenomena, behaviorists argued, could not be directly observed and as a consequence should not be open to scientific investigation. Behaviorists did not go so far as to claim that things that cannot be observed directly did not exist, but considered them highly suspect as a topic of scientific study. Behaviorism made many important contributions to psychology because of this directness to the object of study, especially with regard to the development of rigorous experimental methods.

As the decades passed, however, the short-comings of behaviorist approaches became more and more apparent. During the heyday of behaviorism in America, Europe explored a very different approach to understand the human mind. Sigmund Freud, trying to treat patients with neurotic disorders, developed in Vienna his theory of psychoanalysis. Freud drew attention to the possibility that a lot of human behavior is the consequence of instinctual drives hidden in the unconscious mind. Mental disorders began to be understood as having their roots in unconscious mental processes rather than simple stimulus-response processes. Behaviorism had little to nothing to say about these ideas that very quickly caught public imagination.

Another reason for the demise of behaviorism as the dominant approach in psychology was the development of information theory and the invention of the computer. New technologies tend to captivate scientists influenced by the metaphor of the brain as a machine. The history of the scientific study of the mind and brain in fact reveals that the new technologies of the time quickly became the favorite analogies of the time for how the brain might work[4]. In the middle ages mechanical clocks and hydraulic systems were the most advanced technologies and scholars likened the functioning and structure of the brain to clockwork and hydraulic power. The invention of the telegraph and the discovery that nerve cells respond to electrical stimulation in the 19th century led scientists to consider the idea that the brain works like a telegraph network. The discovery that neural networks consist of neurons and synapses led to thinking about the brain as a flexible telephone switchboard. Most recently, the invention of the Internet and cloud computing saw the rise of the notion of the brain as a massively distributed computer system.

The invention of computing led to an enormous paradigm shift in psychology and brain science during the cognitive revolution of the 1950 and 60s. Scientists became aware of the many apparent similarities between minds, brains, and the new machines. The metaphor of the mind as some kind of computer became the dominant overarching theoretical framework in psychology and related sciences. The most important foundation stone of cognitive psychology is the notion that minds and brains, just like computers, transmit information. One of the founding fathers of information theory was the mathematician Claude Shannon. Shannon's discovery was that electrical switches could solve Boolean algebra, a form of algebra in which the truth values *true* and *false*, typically expressed as 1 and 0, are the values of the variables. Boolean algebra makes use of the operations *and*, the conjunction, *or*, the disjunction, and *not*, the negation. Shannon thus showed that electrical switches could implement a formalism for binary logical operations. Thus, electrical circuits could be described in terms of symbols which formsthe underpinning bedrock of all digital computers. Incredibly, Shannon did not reach this insight as a culmination of a life's work, this was his master's thesis at MIT in 1937. The psychologist Howard Gardner quite rightly called Shannon's master's thesis the possibly most important, and also most noted, master's thesis of the 20th century[5].

At about the same time when Shannon developed information theory, an idea took hold in researchers studying the nervous system that at the time was completely novel. This idea was that neurons in the brain contain some kind of code. The currency of this postulated neural code was information, similar to the succession of dots in Morse code[6]. We now know that neurons are not digital in the way computers are. Important characteristics of neural functioning, such as an increasing rate of neural firing corresponding to the strength of stimuli, are analogue. Yet, the conception that minds and brains, just like computers, process information became the central dogma of the new fields of cognitive psychology and cognitive neuroscience. The computer metaphor caused an important shift in the sciences of the mind: a shift away from trying to understand human behavior as a response to something happening in the surroundings of the human being, the stimulus, and towards understanding human behavior as a consequence of human information processing. Brain science became all about how brains receive informational input, process that information, and deliver an informational output. Psychology became all about understanding how people encode information from the environment, store, retrieve, process and manipulate it. Cognitive science was meant to explain all of mind and behavior as a consequence of human information processing.

The predictive revolution

The beginning of the third millennium, starting in the early noughties and increasing in strength throughout the 2010s, has seen another paradigm shift in the mind sciences. In what might be called the predictive revolution or the predictive turn, many researchers in the psychological and brain sciences have come to consider the human mind a 'predictive engine' or 'prediction machine'[7]. Like its predecessor, the cognitive revolution, more than half a century before, the predictive revolution is grand in ambition. It tries to explain all of mind and brain functioning within one common framework. Prediction is on the verge of replacing cognitivism as the grand unified theory of the mind. In this unified theory, the functioning of mind and brain are no longer best explained as simple information processors, mind and brain have become prediction systems.

The view that some form of prediction explains how the human mind works is not as new as the current hype may suggest. The notion has been around for about a thousand years, but it was the German physiologist Hermann von Helmholtz who first articulated clearly that the human brain makes use of predictions to understand the world. In 1867 von Helmholtz coined the term 'unbewusster Schluss', which translates to unconscious inference, to describe how humans perceive things. Von Helmholtz came to this conclusion, when he studied the human visual system and noticed that the human eye was quite ill-equipped to produce high quality images of a stimulus. Von Helmholtz thus proposed that perceptual sensations are not direct copies of the stimuli but constructed correspondences with the stimuli that the human observer arrives at through repeatedly learned unconscious inferences. Von Helmholtz investigated several characteristics of human vision that were consistent with this conclusion. Stereoscopic vision, our ability to use the two overlapping visual images of our two eyes to create three-dimensional vision, told him that the two different images projected to the retinas of the left and right eye are resolved into one image by the brain. The brain also has to adjust for size distortion caused by perspective, for example in forced perspective photos when objects appear to be closer, further away, smaller or larger than they actually are such as an elephant on top of a person's hand or someone holding the leaning tower of Pisa. Von Helmholtz also noticed that we learn the spatial arrangement of objects by moving our fingers over the object. All this taken together, von Helmholtz argued, suggests that perception is not an inborn ability like breathing, it is not a direct perception of a stimulus, it is learned prediction. This learned predictive perception is determined by the properties of our sensory organs, for example that we see the world with two eyes rather than one, as well as our past experience because inferences require prior exposure to the world. In other words, when we perceive something, we rapidly make unconscious guesses about what it might be rather than relying only on our sensations. Von Helmholtz admitted in this regard that it is not straightforward to distinguish what perceptual content might be contributed by predictive processes and what by the sensory input[8].

The importance of predictive processing for mind and brain however has until recently not been strongly advocated by most scholars in the mind and brain sciences despite von Helmholtz's influential views about inferences in perception. One reason for that appears to be that, although cognitivism strongly rejected behaviorist views, it nevertheless carried over some substantial baggage from the earlier behaviorist frameworks about a linear progression from stimulus to response during human behavior. Cognitive theories, especially in the early days, often hypothesized a similar serial progression of processing. Information flow was typically described as unidirectional from lower to higher levels of processing[9].

Information processing during face recognition, for example, was postulated to proceed in distinct, successive, and sequentially organized stages from low level sensory input to eventual high level face recognition without any possibility of feedback from higher to lower levels[10]. Such a sequence of processing, starting from the incoming sensory stimuli and working strictly upwards for the brain to interpret these signals to finally consciously recognize a face, is often called bottom-up processing. During bottom-up processing, the sensory input at the retina of the eye of the observer is the entry point, with the brain putting together all the incoming input and transforming the increasingly complex information in successive stages in systematic ways until there is some form of match with a face stored in memory and the now recognized face is outputted. This strict input to output order is a reflection of the view of humans as information processors. It is a hallmark of the computer analogy of the human mind in cognitivism: inputting, storing, retrieving, and outputting data. The bottom-up process can thus be described as 'what you see is what you get', or to use the computer metaphor of the cognitive revolution, purely data-driven.

The predictive revolution, however, suggests a different way of how faces might be recognized, making use of what is often called top-down processing, namely relying on existing knowledge and prediction to aid recognition of a face. A striking example that prior knowledge and expectations can shape what we perceive is a phenomenon called pareidolia, the tendency to see an object or perceive a meaningful pattern where there is in fact none. Examples of pareidolia include when people believe they see a man in the moon or the face of Jesus on a piece of toast. In line with von Helmholtz's notion of unconscious inference, pareidolia shows that our typically subconscious expectations can affect what we see. It is called top-down processing because such phenomena are typically interpreted as the brain sending 'down' general information to the sensory system analyzing the stimulus about what a perceived stimulus is likely to be. This process is usually framed as higher levels not waiting for bottom-up processes to be completed. In other words, low level analysis of the basic visual features of the stimulus has not yet been completed, but higher levels already 'impose' the interpretation that it is the face of Jesus on the toast in front of us. Pareidolia is not restricted to vision, other examples are when someone believes to hear hidden messages in music recordings played in the reverse, or voices in random noise. In accordance with the predictive revolution in the mind sciences, phenomena such as pareidolia are no longer interpreted as purely abnormal psychological processing. Rather, they are seen as consequences, albeit striking ones, of how the mind anticipates ongoing input.

What are the reasons for the predictive revolution?

It is an interesting question why the paradigm shift towards explaining the mind in terms of prediction has occurred in the last few years. Why has prediction recently become such a promising and viable candidate principle that may explain the functioning of the mind and brain? What has changed in the mind sciences that provided such a fertile ground for the predictive revolution to occur? One reason has been the growing realization in the mind sciences that there are severe processing bottlenecks during human information processing. Our brains get bombarded constantly with such a huge amount of information from our surroundings. Even the hundred billion neurons and hundred trillion synaptic connections in the human brain cannot handle this all at the same time[11]. Prediction is a good potential solution to this bottleneck problem because it can do the heavy-lifting of reducing the onslaught of information to surprising input. Prediction may allow the mind to focus on unexpected information, thereby reducing the severity of the bottleneck or avoiding it altogether.

A second realization contributing to the predictive revolution has been the recognition that perception ultimately supports action. In much of traditional cognitive science, perception had been assumed to be a rather passive, stimulus-determined process in which a visual object or scene, for example, is recognized in a step-by-step bottom-up fashion to build a rich visual model of the surrounding environment. In contrast, researchers increasingly noticed that perception is rarely a passive process, but instead active and often related to the tasks people are engaged in[12]. Perception tends to happen in the service of a specific action rather than the construction of a 'visual copy' of the outside world in our heads. Prediction here again offers the potential to concentrate on the most task-relevant perceptual information which in turn can result in quicker and more efficient actions.

The third reason is the apparent theoretical elegance of predictive mind and brain frameworks in that they may overcome the short-comings of both behaviorism and cognitivism. In hindsight, at least, it is obvious that the simple stimulus-response coupling explanations put forward by behaviorism do not do justice to the complexities of human mental life and behavior. The strict focus of cognitivism on often rather passive human information processing, however, falls also short as a plausible complete account of human psychology. The mind as a computer analogy has reached its limits. The predictive revolution promises to reconcile cognition and behavior as the intrinsically connected two sides of the same coin serving human interactions with the environment.

References

[1] Hatto, A. (1949). " Revolution": An Enquiry Into the Usefulness of an Historical Term. *Mind*, *58*(232), 495-517.

[2] Kuhn, T. S. (2012). *The structure of scientific revolutions*. University of Chicago Press.

[3] Watson, J. B. (1914). *Behavior: An introduction to comparative psychology*. New York: Holt.

[4] Cobb, M. (2020). *The idea of the brain: The past and future of neuroscience*. Hachette UK.

[5] Gardner, H. (1987). *The mind's new science: A history of the cognitive revolution*. Basic books.

[6] Adrian, E. D. (1932). *The mechanism of nervous action*. University of Pennsylvania Press.

[7] Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*(3), 181-204.

[8] Helmholtz, H.L. (1867/1910). *Handbuch der physiologischen Optik*. Leipzig: L. Voss. Reprinted in A. Gullstrand, J. von Kries & W. Nagel (Eds.) *Handbuch der physiologischen Optik* (3rd edn.). Hamburg and Leipzig: L. Voss.

[9] Levelt, W. J. (1993). *Speaking: From intention to articulation* (Vol. 1). MIT press.

[10] Bruce, V. & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327.

[11] Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, *9*(6), 296-305.

[12] Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing* (No. 37). Oxford University Press.