Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Biologies 329 (2006) 386–397
COMPTES RENDUS BIOLOGIES
http://france.elsevier.com/direct/CRASS3/
Neurosciences

Behaviourism and the mechanization of the mind

Jean-Claude Lecas

Laboratoire de neurobiologie des processus adaptatifs, UMR 7102, CNRS, université Pierre-et-Marie-Curie (Paris-6), 9, quai Saint-Bernard, 75252 Paris cedex 05, France

Received 15 November 2005; accepted after revision 15 February 2006

Available online 27 April 2006

Abstract

The significance of Behaviourism is examined in relation to its far conceptual roots, i.e. comparative animal studies initiated by Darwin, mechanistic physiological thinking initiated by Descartes and empiricist associationism. The Behaviourist anti-mentalist position induced neuromechanistic interpretations based on Pavlovian reflexes, stimulus-response connectionism and the very first hypotheses on synaptic plasticity. As a result, the evolutionary tradition was dropped and the two other trends were combined into a new adaptive version of Cartesian automaton, with persisting influences in modern reductionist thinking, from robotics and cognitive science to the neuroscience of learning and memory. ***To cite this article: J.-C. Lecas, C. R. Biologies 329 (2006).***

Résumé

Le compartementalis et la mécanisation de l'esprit. On examine ici la signification du behaviorisme, doctrine anti-mentaliste qui domina la psychologie de l'apprentissage des années 1930–1960. Ce mouvement est issu des études comparatives de tradition darwinienne, mais aussi de l'empirisme associationniste et de la physiologie réflexiste d'origine cartésienne. Privilégiant les interprétations mécanistes, le connexionnisme stimulus–réponse et les premières hypothèses de plasticité synaptique, il abandonne la perspective évolutionniste pour une synthèse originale des deux autres courants. Ainsi apparaît une version adaptative de l'animal-machine cartésien qui annonce la pensée réductionniste moderne, la robotique cognitiviste et les neurosciences de l'apprentissage et de la mémoire. ***Pour citer cet article : J.-C. Lecas, C. R. Biologies 329 (2006).***

Keywords: Behaviourism; Cartesian automata; Evolution; Comparative psychology; Learning; Connectionism

Mots-clés : Behaviorisme ; Automate cartésien ; Évolution ; Psychologie comparative ; Apprentissage ; Connexionnisme

1. Introduction

During the decades 1930–1960, behaviourism represented the mainstream of experimental psychology. Today, it is an old anti-mentalist doctrine, quite disregarded. For example, in his book *'Descartes's error'*, Damasio invoked the rise of cognitive psychology as a salutary revolution, following the long night of behaviourism which emphasized the stimulus-response couple. However, in its time also, behaviourism presented itself as a radical opponent of the former introspectionist psychology of consciousness. Is psychology condemned to such recurrent polemics? No other science seems to give itself up to such a periodic burning of idols. Alternatively, it may be suggested that the 'cognitive revolution' was not so abrupt and that much of

E-mail address: lecas@ccr.jussieu.fr (J.-C. Lecas).

1631-0691/$ – see front matter
doi:10.1016/j.crvi.2006.03.009

the 'old' behaviouristic era survives in the 'new' cognitive psychology, now associated with the neuroscience and cognitive sciences. In other words, there is more continuity than discontinuity. Here this possibility is examined in the broadest possible perspective. The roots of behaviourism involve comparative psychology initiated by Darwin, British Empiricism emphasizing learning, and the reflex physiological model reinforced by Pavlovian conditioning. The mechanistic physiological tradition of reflexes can be traced back to Descartes and contrasted with the evolutionary perspective. This comparison suggests that behaviourism combined empiricist and mechanical trends in opposition to the evolutionary perspective. In this way, it introduced a new brand of Descartes's mechanical animal, not so far from modern robotics. In that sense, behaviourism can be seen as an important transition step towards modern reductionist thinking, and deserves more thorough study.

2. Descartes's machine animal and the emergence of physiology

2.1. The machine animal as a consequence of the cogito

Descartes's dualism, opposing soul and body, viewed animals as biological automata. Connected to this remained a famous anecdote, now frequently used by defenders of animal rights. Nicholas Malebranche, a distinguished disciple of Descartes, once kicked a dog in the belly and responded to the protests around that, since the animal had no soul, its cries were just a mechanical reaction. This caricature is unfair to Descartes whose ideas had little to do with any justification of cruelty. However, it is true that the mechanical animal is a key concept in Descartes's philosophy, especially in those aspects which favoured the development of modern science.

The appeal of the cogito approach rests on being a logical quest for undisputable truths, conducted from the point of view of the individual, and rejecting the statements of authority made by the Aristotelian scholastics. (In the context of the time, this was potentially dangerous: Galileo had been condemned in 1633, four years before the *Discourse on Method*.) What are the means, asked Descartes, for acquiring an exact knowledge of natural facts based upon unquestionable deductions? The first condition was to reject prejudices and to drop any opinions and beliefs that would not be grounded on 'clear and distinct ideas', that is on the logico-deductive method used in mathematical demonstrations. But then, everything was a matter of doubt. However, as it is widely known, this 'hyperbolic doubt' introduced the *cogito* which established the first secure ground for reasoning: to be able to doubt, I must think, and if I think then I must exist, at least as a thinking object or substance: the *'res cogitans'*. The existence of thought was primary, it satisfied the criterion of clear and distinct ideas and was probably to be distinguished in essence from all the rest (the material world, the body). From this point, the second step was a long detour to prove logically the existence and perfection of God, which grounded and strengthened the validity of the 'clear and distinct ideas' rule. In other words, God (which may be given the modern meaning of Nature) had gifted us with an innate capacity for conceiving mathematical reasonings (*'innate ideas'*) on which we could safely rely. In a third step, Descartes established the probable existence of the external material world from the fact that sensation comes to the mind independent of the will and thus should be induced by external events. According to the preceding step, this physical world should now be understood with the intellectual tool of mathematics. It was an object of science. More precisely, matter (*'res extensa'*), being characterized by extent and movement (while thought, being immaterial, was not) could be investigated with geometrical analysis and mechanics.

This conceptual leap, which ruined the basis of the Aristotelian philosophy, placed Descartes with Galileo among the founders of modern physics. In parallel with establishing these epistemological premises, Descartes illustrated the new way of thinking with his works in analytical geometry and geometrical optics (reflection and refraction laws), later used and improved by Isaac Newton. However, since animal and human bodies were also capable of movement, the same principles applied to living organisms. They could be studied with a particular mechanics, i.e. physiology, grounded on a specific geometry, i.e. anatomy. Descartes aimed at extending mechanical concepts to physiological functions. He searched for *"a means of finding a medicine which would be based on infallible demonstrations"* (letter to Father Mersenne, 1630). He believed that only humans possessed a dual soul/body nature, which permitted them both involuntary movements and voluntary movements for expressing their thought. On the contrary, animals were only capable of involuntary, machine-like movements. They were 'biological robots', i.e. machine animals. Although common sense credited animals with a sort of mind because of their sensory/affective reactions, Descartes refused to admit the existence of animal thought without language. In a letter to Henry More, he explained: *"Please note that I am speaking of thought,* *and not of life or sensation. I do not deny life to animals, since I regard it as consisting simply in the heart of the heart; and I do not even deny sensation, in so far as it depends upon a bodily organ. Thus my opinion is not so much cruel to animals as indulgent to human beings... since it absolves them from the suspicion of crime when they eat or kill animals.*"

Descartes's commitment to 'physiologizing' led him to support Harvey's theory of blood circulation (1628), which initiated modern physiology. Not up to the point, however, of accepting Harvey's conception of the heart as a pump (viewing it as a boiler). Rather, Descartes extended the new concept to the circulation of Galen's 'animal spirits' in nerves. Though recognizing the necessity of experiments, he himself preferred to deduce corporeal functions from anatomy. But when data were badly lacking, theory turned to speculation. This resulted from the fact that Cartesian dualism was not based on preconceived postulates but on a logico-deductive approach. This strength eventually turned to weakness. The *'res cogitans'* (the soul), involved reflexive thought, reasoning, conceiving, judgement and will. It was *not* the whole mind. For the sake of demonstration it was opposed to the body, *'res extensa'*, viewed as a machine. Descartes opposed the extremes. No wonder that a communication problem appeared between these two immaterial and material substances, namely with voluntary movement (soul controling body) or sensation (body informing the soul). This was the 'dualist impasse'. The pineal gland hypothesis where interaction took place was simply palliative (of which substance was it made?) which Descartes was forced to acknowledge. Logically, his opponents will attack on sensation and emotional life (supposedly due to the *"intimate union of soul and body"*) and on the mental life of animals: "*they feel like us*" said Condillac.

2.2. The rise of mechanical thought; a short history of reflexes

By the end of the XVIIth century, the audience of Cartesian dualism began to decline and Descartes's physics was replaced by that of Isaac Newton. However, Descartes's mechanical physiology took roots in the medical schools as an heuristic doctrine: desanctifying the body, it favoured the dissection of corpses, vivisection and animal experiment. Famous physicians of the time referred to Descartes' writings, for example De Boë (Sylvius, 1614–1672), Hermann Boerhaave (1668–1733) of Leiden, William Cullen (1710–1790) and Albrecht von Haller (1708–1777). The machine animal concept traversed the whole XVIIIth century, quite disconnected from Cartesian dualism, since it even coexisted with vitalism (naturalistic spiritualism) in Buffon's *Natural History*. Mechanical thought was clearly gaining audience, as witnessed by its most extreme and provocative illustration: the famous *Man-a-machine* (1748) by Julien Offray de la Mettrie (1709–1751), arguing from the animal nature of man that he also was a biological automaton and that soul should be denied.

While many physiological speculations of Descartes proved fanciful, he was actually credited with the first theory of reflex, an illustration *par excellence* of the mechanical aspects of involuntary movement (see Sherrington, *Man on his Nature* [1]). The nociceptive reaction to heat, sketched in *De Homine* (1764, posthumous but written in 1633) as a reflection of animal spirits, was discussed by Robert Boyle (1627–1691) and termed 'reflex action' by Thomas Willis (1714–1766) who viewed it as an elementary movement of nervous origin. Stephen Hales (1677–1761) then showed that such 'involuntary movements' depended on the spinal cord. Robert Whytt (1714–1766) conceived it as a sentient principle dispatched back to muscles and showed experimentally that only a small segment of the cord was necessary (1751). On the next step, Charles Bell (1774–1842) recognized the sensory dorsal and the motor ventral roots (1811) an observation completed by François Magendie (1783–1855) who established the reflex arc (1822, the 'Bell-Magendie' law). These experiments were continued by Johannes Müller (1801–1858) on frogs (1831) which proved an ideal preparation for studying the different spinal reflexes. After the synthesis of Marshall Hall (1790–1857) in 1833, Emil Du Bois Reymond (1818–1896), a pupil of Müller, started the first modern experiments with electrical stimulation. However, modern physiology of reflex was mainly due to Charles Sherrington (1857–1952) who coined the term 'synapse' and established the current classification of receptors (extero-, intero-, proprioceptors). Sherrington showed how muscular proprioception controlled the myotatic (antigravitational) reflex and demonstrated reciprocal inhibition before explaining how the integration of spinal reflexes form the basis of locomotion and posture. Sherrington credited Descartes with the first concept of reciprocal inhibition, formulated from the layout of extraocular muscles suggesting that an ocular movement resulted from both the action of one muscle and the release of its antagonist. Thus it is apparent that modern physiology of reflexes owed much to Descartes's machine animal, a conclusion best illustrated by the expression *physiological mechanisms*.

3. From the dualistic impasse to the associationist psychology of consciousness

The dualist impasse was responsible for the decline of philosophical Cartesianism and for the emergence of other systems, particularly that of Spinoza, in which soul and body were two different aspects of one single divine reality. At this time, British empiricists developed quite a different philosophy, based on Hobbes's rejection of Descartes's innate faculties. Locke (1632–1704), then Berkeley (1685–1753) viewed the organism at birth as a *tabula rasa*, which gained knowledge (ideas) through sensory experience. Though they already conceived abstract ideas as compounding simpler mental elements, raw perception and context, this associationist principle was more thoroughly worked out by David Hume (1711–1776) and David Hartley (1705–1757). The Scottish school of Thomas Reid (1710–1796) and Thomas Brown (1778–1820) then distinguished the elementary sensations (related to sense organs) from perception which involved the notion of objects. Finally, during the XIXth century, James Mill (1773–1836) and his son John Stuart Mill (1806–1873) elaborated a full range associationist doctrine. In short, the empiricist-associationist tradition afforded a memory explanation of knowledge. Present and past ideas were connected through a mental synthesis based on association mechanisms comparable to the chemical combination of simple elements into compounds. These principles heavily influenced every subsequent theory of perception, from Helmholtz up to now.

The associationist paradigm became even more prominent with the birth of scientific psychology, usually dated from Fechner's *Elemente der Psychophysik* (1860) and the foundation of Wundt's laboratory (1879) in Leipzig. The explicit goal of the new psychological school was the analysis of consciousness. Perception, viewed as an associationistic combination of elementary sensations, was the basic model of mind. Consciousness resulted from the effect of attention (*'apperception'*) constraining specific associations so as to form a definite mental content. However, the debate on the localization of cortical functions which took place during the second half of the XIXth century had come to the definition of 'channels', 'centres' and cortical projection areas for each sensory system. The task of sensory physiology was thus to investigate the unconscious mechanisms of peripheral coding, conduction and transformation of sensory information up to the cerebral cortex, *a task workable on anaesthetized animals*. Then, because sensory excitation was supposed to evoke a conscious sensation upon reaching the cortex, psychology took over from physiology. Sensation, the elementary mental phenomenon of the associationists, was studied in man with the available experimental methods: psychophysical detection, reaction time and introspection. In this division of labour, the cerebral cortex was the dividing line between physiology and psychology. The former investigated the mechanical side of the nervous system, while psychology pursued the analysis at the level of conscious phenomena. It was this psychology of consciousness that Wundt's pupils Titchener and Angell imported to America and which later came under the fire of behaviorists.

3.1. Merging reflex and Associationism: Alexander Bain

A major ciriticism addressed to the empirist-associationist philosophy was its total neglect of action and movement, leading to identical neglect of important mental categories such as will and intent. The associationists' subject lacked initiative. He was contemplatively and passively imprinted by experience. When Thomas Brown attempted to reintegrate movement within associationism, he only considered its sensory side, the 'muscular sense', or self-movement perception. The turn of associationism was due to Alexander Bain (1818–1903) who introduced sensori-motor associationism in two influential books, *The Senses and the Intellect* (1855) and *The Emotions and the Will* (1859). Bain was heavily influenced by the physiological writings of Johanes Müller, whose monumental *Handbuch der Physiologie des Menschen* had been made available to English readers. He came to the conclusion that action was important both in itself and through its sensory consequences, since it "*... enters as a component part into every one of the senses, giving them the character of compounds"* (Bain, 1868, p. 59). Consequently, the spontaneous associations between movements *"... and the pleasure and pains consequent upon them, educate the organism so that its formerly random movements ... (become) adapted to ends or purposes"* (see Wozniak [2]). This important statement paved the way for the notion of reinforcement.

4. Darwin's naturalization of psychological functions

4.1. Natural selection

Descartes's dualism had brought together the whole physical world into one single category of 'matter'. Centuries later, differences unknown to him were put forth by naturalists' studies of the animal series. Perhaps may we still accept the notion of biological automata for Amoeba and Insects, but evidently not for anthropoid apes. This is the kind of difference that came with the theory of Evolution.

Darwin did not invent the concept of evolution. A number of naturalists before him, most notably Lamarck, had proposed transformist hypotheses to explain how the present species derived from older forms. However, Darwin, together with Wallace, elaborated the first vastly documented theory of evolution based on natural selection. This theory was supported by a thirty-year accumulation of notes and documents, especially those collected during the round-the-world trip of the Beagle (1831–1836). It explained the gradual differentiation of species, their appearance and extinction. In 1859, the first edition of the *Origin of Species* was exhausted in one single day. Darwin's rationale started from a close scrutiny of natural variation across generations. Breeders long knew how to use it for the selection of domestic races, but it was similarly important in the wild. Using Malthus's work (1798) showing that unlimited reproduction led to geometric growth of populations, Darwin observed that such a rapid growth that would exhaust resources was almost never seen. It was prevented by mortality and harsh struggle for existence. Accordingly, any favourable variation in the offspring allowed the endowed individuals to thrive and reproduce better. The theory of natural selection was therefore a theory of differential fertility and mortality in relation to favourable or unfavourable variation. However, variations in weight, morphology and organs were not the whole story. Darwin also investigated instincts, so essential for feeding, reproduction and survival. He examined the case of cuckoos laying eggs in nest of other birds, the behaviour of slave-making ants and cell-building in hive-bees. He showed that instincts also vary and must therefore be subject to natural selection. In short, abilities and behaviours evolve. However, in 1859, he prudently restrained from addressing the problem of man's origins. He believed necessary first to impose the idea of natural selection against creationists and to avoid unnecessary polemics that could be fatal to it.

4.2. Continuity and discontinuity; Instincts and Man

In the *Origin of Species*, Darwin argued that species evolved by accumulating slight variations on each successive generation. Darwin's creed was *"Natura non falcit saltum"*, nature never takes a leap. But this gradual view of evolution could be expected to trigger controversy concerning the origins of man, because of the philosophical or religious 'prejudices' (sic). While his disciples urged him to take a stand, Darwin awaited for twelve years before publishing the *Descent of Man* in 1871, followed the next year by *The Expression of Emotions in Man and Animals*.

In the *Descent*, only the first part treated this subject (the rest of the book was on the phenomena of sexual selection complementing natural selection). In that first part, so important, Darwin forcefully sustained the gradualist view by seeking the premises of humane intelligent conducts in animals. He examined the whole array of cognitive functions with anecdotes and anthropomorphic interpretations. He endowed animals with capacities of emotion, curiosity, imitation, attention, memory, imagination, common sense, reason, learning (*'progressive improvement'*). He credited them with the ability of using tools and with faculties of abstraction, self-consciousness and even with some language and feeling of beauty. Yet his discourse changed when coming to the belief in God (which he admitted to be specific to man), moral sense and social reflexes of solidarity. Suddenly Darwin admitted discontinuity between man and animals. Here, he developed an idea by Wallace attributing the success of man as a species to his social instincts. Behaviours that strengthened mutual cohesion, such as cooperation, mutual aid, compassion and love of one's neighbour, conferred considerable advantage to the first families and tribes. By broadening the social group and progressively building societies, man escaped the conditions of natural selection: he protected himself from cold by clothes, cooked his food, hunted in groups and invented division of labour. Natural selection had selected behavioural traits which now prevented its full action.

Yet instincts are well present in man. In the *Expression of Emotions*, which closely followed the *Descent*, Darwin analysed facial mimics in animals and man: were they habits, reflexes or instincts? Using various documents, observations of his own son (then an infant), photographs of actors and responses to a questionnaire sent to missionaries and government officials throughout the British empire, he was able to establish that facial mimics are similar in all peoples and races, that they are soon present in the infant, including the born blind, and that the state of mind, or mood, they expressed was instantaneously recognized whatever the nation and culture. These universal and innate characteristics witness for the biological unity of mankind. Mimics constituted a primitive affective language showing the crucial importance of social communication in humans. In the *Origin of Species*, Darwin had quickly sketched what he meant by 'instinct': an innate behaviour, soon mastered by the young without learning and widespread in the species. By showing that these features were present in the expression of human emotions, but in relation to social interactions, Darwin came to a more precise conception of man's origins, involving *both* continuity and discontinuity. But he also introduced a new approach. The concept of instinct can be traced far back to ancient thought. Its meaning was that remained in common language with expressions as 'lower' or 'basic' instincts. Darwin avoided any such judgement of value by describing instinctive behaviour rather than discussing preconceived instincts. From this point, studying behaviour in an evolutionary framework was a part of Biology, and called into question any psychology defined as a philosophical exegesis of 'human nature'.

5. From Darwin to behaviourism

5.1. The new comparative psychology, Romanes on continuity

While constantly pressing the case that differences among species were *"differences of degree not of kind"*, Darwin had freely interpreted animal behaviour in relation to capacities usually attributed to the sole humans. The same approach was taken and continued by Georges Romanes (1848–1894) with the support of Darwin himself who made his notes and documents available to him. Romanes launched into a meticulous inventory of anecdotes for illustrating the intelligent side of animal behaviour. His book *Animal Intelligence* (1882) was very successful in the best Victorian society and many people bombarded the author with new anecdotes concerning his/her preferred pet. But Romanes was not an inflexible anecdotalist and he established criteria for accepting data, even verifying them through simple experiments [2]. Eventually, Romanes introduced two ideas that ultimately proved very influential: the assessment of intelligence in lower forms through adaptive behaviour and *learning* and the idea of a *continuous scale* of mental abilities. He primarily wanted to show that all living animals possess to various degrees reflex, instinct and reason. But since all three may prove adaptive, he distinguished reason by the *"knowledge of the relation between means and ends"*. This principle justified a number of anthropomorphic interpretations and his ordering of the animal series (Amoeba, Worms, Insects, Fish, Amphibia, Reptiles, Rodents, Cats, Dogs, Apes, Humans) was associated with a linear increase of reason and a continuous decrease of reflex. In his later texts, Romanes transposed Haeckel's popular view that *"ontogeny recapitulates phylogeny"*, mostly based on embryologic data, by indicating the age at which the human infant or child reaches the faculty observed in the adult of a given animal species. According to this view, the mental power of each species was characterized by the distance covered on an identical scale of mental capacity.

5.2. Spencer and Lloyd Morgan on Instinct, Morgan's 'canon'

Before the publication of the *Origin of Species*, Herbert Spencer (1820–1903) already defended a global conception of evolution, more Lamarckian than Darwinian, in which he proposed that intelligence gradually emerged from the diversification and complexification of instinct. This theory influenced Conwy Lloyd Morgan (1852–1936), a pupil of the Darwinian tradition through Thomas Huxley, and then professor at Bristol University. Morgan pursued with many criticisms the work of Romanes. Between 1887 and 1900, he studied the behaviour of beetles and more comprehensively that of recently hatched birds, ducklings, pheasants and chicken. He described the acquisition of flight and the phenomenon of imprinting to the mother, already observed by Douglas Spalding (1840–1877). However, as he was ignorant of maturation processes, he concluded along with Spencer that instincts were adapted by learning. Morgan is known for the first description of 'trial and error' learning (in contrast to learning by 'imitation'). He reported on the manner his dog Tony acquired the habit of opening the exit door of the garden, after having raised the latch by chance with its head. Morgan interpreted the progress of the animal with the principle of adjustment by 'pleasure and pains' put forth by Spencer and Bain and coined the term 'reinforcement'. Morgan was the outstanding theorist of the new discipline (*Introduction to Comparative Psychology*, 1894 [2]), already illustrated by the work of behavioural naturalists (the term 'behaviour' was from Spalding), such as Lubbock and Hobhouse. He argued for an evolutionary conception of mind based on the emergence of mental phenomena within the animal series ('emergentism'). However, Morgan doubted of Romanes's version of continuity since it would lead *"to believe ... that all forms of animal life from the amoeba upwards have all the faculties of man, only reduced in degree and range..."*. Regarding methods, Morgan did not reject anecdotes but criticized their anthropomorphic interpretation: *"In no case may we interpret an action as the outcome of a higher psychical faculty, if it can be interpreted as the outcome of one which stands lower on the psychological scale"* (1894). This famous 'Morgan's canon', was improperly considered during the behaviourist period as a principle of parsimony appealing to anti-mentalist reductionism, in spite of several clarifications from the author.

5.3. Thorndike

Edward Thorndike aged 22 had come to Harvard for studying under William James. There, in the spring of 1896, he attended a series of lectures on instincts and habits delivered by C.L. Morgan then a visiting professor. Thorndike soon decided to analyse 'trial and error' learning, by *experimentally* reproducing the observation of Morgan's dog. After training chickens in the attic of William James's house, Thorndike went to Columbia (NY) to finish his PhD (1898) under Catell who soon considered him as his star pupil. This work was indeed to gain considerable impact. Thorndike used cats and dogs which he tested with various 'puzzle boxes' constructed with wooden slats and hardware cloth. A hungry animal was put in the box and had to escape from it by manipulating some device (either simple: e.g., depressing a lever, or complex: e.g., three consecutive actions) in order to reach the food outside the box. By measuring the time taken by the animal to escape on consecutive trials, Thorndike obtained the first learning curves. These curves exhibited a *gradual decrement* but no sudden decrease that might have indicated the comprehension of appropriate action. Imitation tests were also performed with an observation box from which a cat could see the training of another animal, but contrary to current opinion, this procedure did not improve subsequent learning of the observer. From these observations, Thorndike interpreted his results as a chance selection of the appropriate movements from the initially agitated behaviour. The animal went to reproduce the useful response with no knowledge of the situation, because success (reward) *'stamped in'* simple connections between *'perceptions of the situation'* and *'motor impulse'*. Later, in his book *Animal intelligence* (1911 [3]), Thorndike elaborated the 'law of effect', in association with a 'law of repetition' and a 'law of readiness'. He stated that, when followed by '*satisfaction'* the response was "*more firmly connected with the situation"* so that it became more frequently evoked by the stimulus situation. When it was followed by *'discomfort'*, it became less frequent. Basically, Thorndike's theory was very close to Bain's sensori-motor associations and it was to become a cornerstone of behaviourism. The organism continuously made associations between perceptions and movements (stimulus-response or SR connections) some of them being selected by the reinforcement. In Chapter VI of his book, Thorndike explained the law of effect in physiological terms by the connection or disconnection of neural elements through synaptic modifications, which made the animal a true neuronal automaton. By parenthesis, it is astonishing that Hebb was later (1949) credited with the idea of synaptic plasticity subserving learning since Thorndike's formulations were so clear:

"The chief life processes of a neurone concerned in learning [i.e. due to *'satisfaction and discomfort'*] *are ... reception and conduction of the nerve impulse, and modifiability or change of connections.. The connections formed between situation and response are represented by connections between neurones and neurones ... across their synapses. The strength or weakness of a connection means the greater or less likelihood that the same current will be conducted from the former to the latter rather than to some other place."* And he thought of the synaptic mechanisms as *" ...protoplasmic union, or proximity of the neurones in space, or a greater permeability of a membrane, or a lowered electrical resistance, or a favorable chemical condition of some other sort."*

6. The behaviourist revolution

6.1. The beginnings

At the turn of the century, psychology was traversing a crisis. A number of papers pointed out the unreliability of introspection and called for objective methods yielding verifiable results. Thorndike's work served as a model in the framework of animal and comparative psychology which had already collected a wealth of data. For example, the observations by Jennings (1906) on elementary organisms, by Kline and Small devising the first laboratory mazes (1899–1901) and by Yerkes (1876–1956) on various species (jellyfish, frog, birds, mouse and rat) stood in deep contrast to instrospectionists' quarrels. Time had come for a change. In 1913, the famous lecture of John B. Watson at Columbia University, *'Psychology as the behaviorist views it'* (the 'behaviourist manifesto' [4]), launched the attack on introspection, introspectionists and consciousness, based on three points. Psychology was to become a natural science; it should be aimed at studying behaviour with the methods of animal psychology; and its tangible goals were to predict and control behaviour. Long later, the event was regarded as a foundation, but in fact, it fell flat and reactions were far from being unanimously favourable (Samelson [5]). The spread of ideas was slow and gradual, as shown by the long-lasting decrease in the number of papers discussing introspection and consciousness in *Psychological Review*: 29 during the period 1920–25, 13 in 1926–30, 10 in 1931–35, 8 in 1936–40, 2 in 1941–45 and no more afterwards. However, eventually, through this long maturation, behaviouristic ideas came to a dominant position at the beginning of the thirties. These ideas did not actually define a theoretical school, but rather a sort of consensus about psychology, its methods and purposes, around which a large variety of opinions could be observed.

After 1913, Watson developed a coherent doctrine in three books: *Behavior: A Textbook of Comparative Psychology* (1914), *Psychology from the standpoint of a behaviorist* (1919), and *Behaviorism* (1924). He adopted a radical anti-mentalist position, refusing any causal or explanatory value to the very concept of consciousness. Psychology was a natural science, it was not the science of mind but that of behaviour. As a natural science, it recognized *"no dividing line between man and brute*" and only used objective experimental methods based on measurable responses. Behaviour was defined by Watson as the whole set of organized responses leading to a process of adjustment to the environment (a position close to Spencer). As there was no behaviour without a stimulus, or a stimulus-situation, all mechanisms were viewed as stimulus-response (SR) chains. A key point of his doctrine was the classification of responses into 'hereditary' (emotional and instinctive) and 'acquired'. From this he developed an ontogenesis of complex habits through the action of trial-and-error learning and Pavlovian conditioning. Both summed up for reshaping innate responses (instinctive, emotional) already present in the newborn. Finally, among the acquired responses, Watson distinguished those which are 'explicit' (overt, observable), from inner responses which he termed 'implicit'. A well-known example of the latter was thinking, viewed as a 'subvocal' discourse, sketched as a SR reaction chain in which each response (an unpronounced word) served as a stimulus for the next step, until the final explicit vocal response (1919). This notion of 'implicit', covert, responses will be heavily used by Watson's successors for (laboriously) reintroducing mental functions (e.g., perception, attention, signification, symbols, memory, intention) within the doctrine. For example, attention, anticipation and perception became preparatory 'postures', 'attitudes' or 'set' that oriented the organism towards a certain sort of responding to the situation (Dashiell, 1928 [6]).

With the assigned goal of prediction and control, the behaviorist program was overtly positivist. Science was to contribute to general welfare by initiating useful techniques. Applied psychology was rapidly growing. Hugo Münsterberg had laid the basis of industrial psychology and Thorndike had turned to educational psychology. After World War I, the spread of mental and ability testing (used by the military) was promissory of an era of 'mental technology', according to Yerkes who had played a major role in the Alpha military program. Watson himself had analyzed infant phobia by means of Pavlovian conditioning before leaving university for advertising. Contrary to the classical psychology of sensation and consciousness, headed by an academic elite fond of philosophical thought and isolated from social reality, the new behaviourism furnished these emergent disciplines with a common psychological doctrine, quite appealing to the young generation. It involved a simple conception of both the organism (individual) and his environment (stimuli), universal mechanisms of habit formation and the urge for experimental measurements in every field. However, until the end of the twenties, when Dashiell published a general synthesis (1928), many programmatic texts were to appear, but very few data, which looked puzzling on the part of a movement calling for objective experimental measure.

6.2. Divergence from comparative psychology

Yet objective data existed in other disciplinary fields that did not get the publicity granted to behaviorist authors. In human studies for example, the recording techniques devised by Dodge in 1901 had already permitted the analysis of ocular movement during reading. With such data, cognitive hypotheses about comprehension and meaning were published as soon as 1908 by Edmund Huey (*The Psychology and Pedagogy of Reading* [2]). Animal comparative psychology was elaborating new behavioural techniques, such as multiple choice discrimination (Hamilton, 1911; Yerkes, 1916), delayed response for measuring mnemonic capacity (Hunter, 1913), 'reasoning' tests in the rat (Maier, 1929), before Harlow's 'learning sets' in 1949 (see Munn, 1971 [7]). All these behaviours did not fit so easily into the stimulus-response mould. Robert Yerkes, clearly the most representative researcher of the field, rejected behaviourism and broke with Watson with whom he had worked in 1911. Long after the war and much effort, as a Yale professor, Yerkes finally succeeded in 1929 in installing the primate laboratory he had been thinking of for many years (nowadays the Yerkes Laboratory). In all respects, primate studies brilliantly pioneered by Köhler's work at Teneriffe (*The mentality of Apes*, 1925 [8]) stood in sharp contrast to behaviourism. Moreover, other comparative studies showed that to be valid, inter-species comparisons had to take into account the behavioural repertoire of the animals (Maier and Schneirla, 1935 [9]). This was quite contrary to the simplistic view of animal continuity invoked by the behaviourists.

6.3. The core principles of behaviourism: tabula rasa and physiological reductionism

To the introspectionists, Watson opposed the methods and achievements of animal comparative psychology, which primarily aimed at studying instinctive behaviour to seek the origins of intelligence. Behaviourists felt somewhat embarrassed with the concept of instinct, as with motivation. Their positions were varied and changing. They could not deny the existence of innate behavioural patterns, but they preferred to discuss innate/acquired interactions and finally retained the only emotional or motivational aspects of instinct: the 'drive'. They largely held, along with Spencer and Morgan, that instinctive behaviours were modified, if not suppressed, by experience. When ascending the phylogenetic scale, these innate behaviours were gradually replaced by acquired habits so that, in man, all behaviour was virtually acquired. For behaviourists the dismissal of instinct was certainly more essential than the rejection of consciousness. This was the basic anti-nativist principle of the empiricist-associationist philosophy rejecting innate 'ideas' or capacities: the organism at birth was a *tabula rasa* which acquired knowledge through experience. Starting from comparative psychology describing the *specific abilities* of animals and man, behaviorists came to this paradox that they treated learning as so fundamental that it transcended species. A vast majority of their studies was on learning, almost exclusively using the rat as a 'model' species.

This explained why behaviourists became infatuated with Pavlovian conditioning. Reflex was then the dominant physiological concept and was considered the basic building block of brain architecture. It was therefore most exciting to learn from Pavlov (see Buser, this volume) that the reflex wiring might change under the action of reinforcers. Pavlov's studies were known in America as earlier as 1909, but they were popularized by Watson, then by Razran and Gantt (who started a conditioning laboratory at John Hopkins, in 1929). American researchers extended classical conditioning to vegetative responses (heart rate) and generalized its concepts in psychiatry, social psychology, etc. Conditioning complemented trial-and-error learning by explaining signal substitutions and finer discriminations for adapting responses to the environment. Because of Thorndike's SR connectionism, both types of acquisition could be theorized in terms of neural circuits. Behaviorists became overtly neuroreductionists and addicted to 'neuromechanistic' (a term of Woodworth, 1924 [10]) physiological speculations, especially concerning the 'synaptic resistance' changes that closed SR links.

6.4. The second Behaviorism: the learning theorists

For three decades, between the late twenties and the sixties, the behaviourist hegemony was manifest in learning studies. Among the great names, Edward Tolman (1886–1959) and Ivan Krechevsky (1909–1977) are generally considered as the first cognitivists. On the opposite side, Edwin Guthrie (1886–1959), Clark Hull (1884–1952) and later Burrhus Skinner (1904–1990) were deliberate anti-mentalists proposing mechanical interpretations of behaviour. Karl Lashley (1890–1958) was apart because his studies with brain lesions had put him in a position for criticizing current learning theories.

The main problem was to merge into one single theory the trial-and-error model of Thorndike and the more physiological Pavlovian conditioning. However, this integration proved difficult. To make things even more complex, as early as 1930, Tolman who was not convinced by SR connectionism analyzed 'latent learning' resulting from prior exploration of the maze with no reinforcement. When reinforcement was introduced, the rat found its way in a few trials. Along these lines, Tolman further showed that the rat even acquired a global knowledge of the environment, a 'cognitive map' permitting rapid reorganization of the route when the experimenter changed the starting position or maze configuration [11]. Tolman agreed with Krechevsky that the rat might form 'hypotheses' about the correct path from the beginning of training. He elaborated a theory of 'vicarious trial-and-error learning' to explain the behaviour at the choice point, when rat explored, hesitated and finally chose an alley. In all these experiments, the role of reinforcement which remained critical for performance was minored for actual learning, in favour of some imprinting of the stimuli from the environment.

This minimalist view of reinforcement was also found in Guthrie who proposed, along with Hull, a mechanical stimulus-response interpretation quite opposite to Tolman's. His model was a simplified version of Pavlovian conditioning which evoked Pavlov's reply in the Psychological Review (1932). Guthrie held that the mere temporal contiguity between stimuli and movement was sufficient for them to be automatically associated. On the recurrence of stimuli the response tended to be evoked, which explained one trial learning. Contrary to this, Hull brought reinforcement to the fore in the only serious attempt to integrate Thorndike's SR connections and Pavlovian conditioning into one single (but complex) model. With his associate Kenneth Spence (1907–1967), he built a learning theory as completely deterministic as possible, based on the drive reduction role of reinforcement. The response strength was given by a mathematical function (equation) the variables and parameters of which represented the drive, the incentive value of reinforcement and the level already acquired. This latter variable incorporated the past Pavlovian associations, either excitatory or inhibitory, due to reward and non-reward (errors). Hull's *Principles of Behavior* (1943 [12]) left no place for whatever mental function, either intentional or conscious. He rightly warned against any subjective interpretation of the observed behaviour as *"if I were the rat what would I do?"*, but soon jumped to the machine animal. We must view *"the behaving organism as a completely self-maintaining robot, constructed of materials as unlike ourselves may be"*. Thus generations of students went to consider the laboratory rat as an artificial animal, a 'model of the organism'. Hull's equation was the first example of an *adaptive robot*, now widespread in modern cognitive sciences.

Skinner illustrated another facet of radical anti-mentalism, which denied any explanatory value to mental phenomena (although accepting their existence). As behavioural changes were basically adaptations to the external world, it was at this level that we ought to search for the ultimate causes of our actions. Inner processes were mere reflects or mediative instances of overt behaviour about which any theory was condemned to obscure physiological speculations. Skinner urged to precisely describe the particular changes in the spontaneous self-initiated behaviour due to reinforcement. After setting up original methods (response rate and intra-subject design), Skinner studied the acquisition and extinction of the 'operant behaviour' (bar-pressing or key-pecking) with rats and pigeons in his famous 'Skinner box'. He showed how response rate was affected by different reinforcement schedules (continuous, fixed/variable ratios, fixed/variable intervals). In *Walden Two* (1948), his great socio-philosophical speculation of a behaviourist 'brave new world', Skinner expressed his thorough conviction that individual's behaviour is totally and mechanically determined by the history of his positive and negative reinforcements. Logically, he also exhibited his aversion for free will, soul and dualism.

Karl Lashley had begun his career in a time when localization of function and generalization of the reflex concept made the brain equivalent to a large telephone switchboard where precise connections were established between 'centres'. In his first experiments with brain lesions, undertaken in 1914 under Watson, he vainly attempted to disrupt the circuits of a Pavlovian conditioning, then presumed to be cortical. For many years, Lashley pursued his *"search of the engram"* with various learning tasks, various sort of lesions and various animal species, including monkey. Much to his surprise, no lesion *specifically and permanently* abolished the acquired habit. By demonstrating that the degree of habit disruption after the lesion and the capability to relearn the task were both related to the size of the lesion, Lashley came to a clear commitment against sensori-motor speculations *à la* Thorndike, based on the reflex arc model and the switchboard metaphor. As an anti-connectionist and antireductionist advocate, Lashley has deeply changed our conception of brain localization of function.

6.5. The end of behaviourism and the return of mentalism

After more than two decades of effort and controversy, behaviourist theories of learning did not converge: no common concept had been worked out. The very nature of reinforcement, its mode of action, partial reinforcement, or even stimulus generalization were left unexplained. Trial-and-error ('instrumental') learning and Pavlovian conditioning were finally regarded as two distinct mode of acquisition (Hilgard and Marquis, 1940 [13]). In fact, the only common feature of these theories was still the empiricist-associationist assumption of knowledge progressively acquired through experience. It was put in two versions: passive imprinting of the subject by the environmental stimuli (Tolman, Guthrie), or the specific action of reinforcement which 'stamped in' the correct SR association/adaptive response (Thorndike, Hull, Skinner). As seen above, this was a consequence of the behaviourist's denigrative position concerning instinct and innate capacities which moved them away from their Darwinian and comparative origins. Only Lashley had taken a different position by approving the ethologist's work on instinct.

While the beginnings of behaviourism are not so easily distinguished, its fall is more clearly dated. After Hull's death (1952), his work was forgotten incredibly fast, due to the growing skepticism aroused by his mathematical models. Though reluctant to theorize, Skinner proposed an operant learning interpretation of human language acquisition (1957 [14]). He was opposed a devastating criticism by Noam Chomsky (1959 [15]) showing that the explosion of language skills in the four-five year infant could only be interpreted as resulting from the maturation of innate cognitive structures. This was a serious blow against the basic anti-nativist position of the behaviourists subtending the primacy of learning. During the sixties, a second crucial assumption was under heavy attack: that of universal learning mechanisms, equally valid across the animal series (including man), and justifying the principle of 'model' species (rat and pigeon). After the work of Garcia's group on conditioned taste aversion (see Seligman, 1970 [16]), former comparative studies were recalled showing the importance of species behavioural repertoire so as to suggest that acquisition processes could be partly species-specific. The final blow was given by the new 'cognitive psychology' program (as termed by Neisser) in human experimental psychology, advocating for the description of mental operations in information processing terms. This movement induced the return of mentalistic vocabulary and concepts.

7. The mechanization of mind

After Behavourism the scientific landscape changed rapidly under the pervasive influence of the information processing paradigm. However, the cognitivist terms employed often proved deceiving or contradictory, much like Lashley's proposal of a 'conscious machine' in the past (1923, [17]). Modern research offered a contrasted picture mixing truly mentalist approaches and mechanical interpretations and models, somewhat reminiscent of the many differences between Tolman and Hull. Without attempting any detailed analysis for which there is neither place nor historical distance, we shall concentrate on a few trends and examples showing the continuity of mechanistic-reductionist thinking between behaviourism and contemporary era.

During the 1960–1970 decades, new experimental stimulation and recording methods in behaving animals have given the impetus for studying physiological concomitants of learning, either Pavlovian or instrumental. On these grounds, the Neurosciences of 'learning and memory' built up with the substitutive-reductionist approach taken by behaviourists. As studying learning had substituted for the interest in the diversity of animal intelligence, investigating memory mechanisms now replaced the study of different learning processes. Acquisition of a new response was used for modifying memory and studying memory storage. Model tasks, chosen to be fast and simple, were combined with the model-animal principle. For example, the one-trial passive avoidance test (a rat is shocked when entering a dark shelter and its retention is measured later by the latency of this innate response) has been widely used for studying amnestic treatments or biological correlates of memory. A machine-like information processing metaphor summarized the whole approach: acquiring a new response was assimilated to putting new information into memory, the latter being viewed as general function, as in a computer.

This approach paved the way for searching for a unique cellular mechanism that could explain memory in the whole animal series, from invertebrate to man. An explosive development of studies on synaptic plasticity was observed from the early eighties. As noted above, Thorndike's synaptic connectionism (1911) was the very origin of the synaptic-learning idea. Ironically, since chemical transmission was not firmly established before the fifties (see Dupont, this volume), it appears that *the theory of synaptic plasticity as a basis for learning and memory was far ahead of the true biological synapse*. However, as could be expected, synaptic plasticity soon became a very complex field, involving development as well as memory. Different kinds of synaptic plasticity have been described, beginning with long term potentiation (LTP) and depression (LTD). Yet it was not at all evident that synaptic modification would affect neuronal discharge which also depended upon 'intrinsic' cellular excitability. Intrinsic plasticity due to membrane property changes was also demonstrated, but astonishingly it only produced about 150 papers in more that twenty years against several thousand for LTP/LTD (Zhang and Linden, 2003 [18]). There was an evident bias in favour of connectivity against reactivity, again going back to Thorndike's 'neuromechanistic' SR connectionism.

During this contemporary period, human cognitive psychology undertook the analysis of mental processes with novel experimental designs and improved versions of very old methods: reaction time and tachystoscopic detection. But now, at the time of cybernetics and computer science, mental processes were defined as an information processing cascade. Each mental operation was viewed as a transformation of information transmitted by the preceding process or by sensory organs. This new SR reductionism could be sketched as a stimulus-response flowchart composed of a succession of 'black boxes' corresponding to the traditional mental functions: perception, memory, decision, motor response. Such a model of mind, inspired by the computer metaphor (a very simplified computer) resembled a machine blueprint. As roughly put by Varela: *"computers offer a mechanical model of thought"* (*Invitation aux Sciences Cognitives*, 1989, p. 39). On the next step, criticisms did not tackle this mechanical aspect but the sequential design of the flowchart. McClelland and Rumelhart (1986 [19]) proposed a new connectionist architecture, termed 'Parallel Distributed Processing' (PDP) and inspired from synaptic plasticity studies. According to the authors, cognitive functions emerged from the processing activity of distributed neuronal populations and processing was assimilated to learning by adaptation of the connections between constituent neurones. Commentators have underlined that these models explicitly implemented the anti-nativist postulate of behaviourists and empiricists. Starting as a *tabula rasa*, they adapted to the environmental configuration: experience was imprinted into their connectivity.

These models have given the impetus to the elaboration of improved mathematical algorithms leading to sophisticated robotics. In the general framework of cognitive sciences and artificial intelligence, such auto-adaptive automata have proved useful in a number of domains: applied linguistics, archiving, Internet searching, air traffic control, economic forecast, etc. But now, psychological or mental terms only seem to be useful for labelling black boxes. Is there anything basically cognitive in the concept of information? Rigorously speaking, it defines together with entropy the degree of organisation of a message within a given system. This permitted the development of computer science, which is an engineering science. After the Neurosciences of learning and memory, robotics and cognitive sciences have opened to us the era of the mechanization of mind.

8. In conclusion: the importance of behaviourism

To what extend have these modern trends been heralded by behaviourism? Of course, it played no role in the rise of information concepts and technologies, nor in the fascination they have exerted. However, it left the place prepared for their invasion of psychology in several ways. First and foremost, the rejection of consciousness, maybe not so important in itself, heavily favoured mechanistic interpretations. For example, it discarded mental functions amenable to experimental study, even when data were available (e.g., ocular movements and reading). Second, the giving up of the comparative framework precluded any attempt to define mental processes more precisely (which is after all basic to any psychology). Third, the ubiquitous invocation of physiological *mechanisms*, through SR connections and conditioned reflexes, at a level where no physiology was possible (the total 'organism'), turned to mythological and 'neuromechanistic' speculation. Indeed, this was in line with the original conception of Physiology as a mechanics of the body and with Descartes's machine animal. In fact, these combined features have led to a coherent attempt to theorize the psychological process in mechanical terms. Watson's machine-like model of thought and Hull's robotized rat may be taken as naïve but significant sketches of modern automata. The importance of the history of behaviourism thus appears: this movement prepared contemporary psychological thought.

Acknowledgements

The author wish to thank Professor Buser for useful discussions on earlier drafts of the manuscript.

References

[1] C.S. Sherrington, Man on his Nature, University Press, Cambridge, UK, 1951.
[2] R.H. Wozniak, Classics in Psychology, 1855–1914: Historical Essays, Routledge/Thoemmes Press, London, 1999.
[3] E.L. Thorndike, Animal Intelligence, MacMillan, New York, 1911.
[4] J.B. Watson, Psychology as the behaviorist views it, Psychol. Rev. 20 (1913) 158–177.
[5] F. Samelson, Struggle for scientific authority: The reception of Watson's behaviorism, 1913–1920, J. Hist. Behav. Sci. 17 (1981) 399–425.
[6] J.F. Dashiell, Fundamentals of Objective Psychology, Houghton Mifflin Company, Boston, 1928.
[7] N.L. Munn, The Evolution of the Human Mind, Houghton-Mifflin, New York, 1971.
[8] W. Köhler, The Mentality of Apes, Hartcourt Brace, New York, 1925.
[9] N.R.F. Maier, T.C. Schneirla, Principles of Animal Psychology, McGraw-Hill, New York, 1935.
[10] R.S. Woodworth, Four varieties of behaviorism, Psychol. Rev. 31 (1924) 257–264.
[11] E.C. Tolman, Cognitive maps in rats and men, Psychol. Rev. 55 (1948) 189–208.
[12] C.L. Hull, Principles of Behavior, Appleton-Century-Crofts, New York, 1943.
[13] E.R. Hilgard, D.G. Marquis, Conditioning and Learning, Appleton-Century-Crofts, New York, 1940.
[14] B.F. Skinner, Verbal Learning, Appleton-Century-Crofts, New York, 1957.
[15] N. Chomsky, A review of B.F. Skinner's *Verbal Behavior*, Language 35 (1959) 26–58.
[16] M.E.P. Seligman, On the generality of the laws of learning, Psychol. Rev. 77 (1970) 406–418.
[17] K.S. Lashley, The behavioristic interpretation of consciousness, Psychol. Bull. 30 (1923) 237–272, 329–353.
[18] W. Zhang, D.J. Linden, The other side of the engram: Experience-driven changes in neuronal intrinsic excitability, Nat. Rev. Neurosci. 4 (2003) 885–900.
[19] J.L. McClelland, D.E. Rumelhart (Eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, MIT Press, Cambridge, MA, 1986.