Learning and Memory

C.R. Gallistel, Rutgers University

In Luria's (1973) pioneering work on neuropsychology, *The working brain*, he opened his chapter on Memory by commenting on the limited progress that had been made in establishing its material basis:

> "…the cerebral organization of complex forms of mnestic activity still remains a new and largely unopened chapter of neuropsychological science.'

> "The reasons for this situation are the oversimplified views of memory held by the great majority of physiologists, and the total failure to appreciate the complex structure of mnestic processes which psychologists have studied only in the last twenty or thirty years."—p. 280

And

> "Research into the physiology of the conditioned reflex, entirely devoted to processes concerned with the fixation of experience, established nothing more than some basic physiological factors of what was very conventionally described as 'opening the pathways' and 'the reinforcement of conditioned connections'', but in fact contributed nothing of substance to the elucidation of the nature of memory. Equally unproductive were the investigations of learning conducted by the American behaviourists, for despite the fact that there are many thousands of publications on this subject, the nature of memory still remains completely unknown. It was facts such as these that led the American psychologists Lashley, in his well-known paper 'In Search of the Engram' (1950), to the pessimistic conclusion that the material nature of memory is just as much an enigma as it was many decades ago."

As someone whose career has been devoted to trying to find a way to establish the material basis of memory, it is painful to confess that the situation has not changed in the 35 years since he wrote--and the 57 years since Lashley wrote. The question is, Why have we made no progress? The answer, I believe, is that we have misconceived the nature of learning and, hence, the nature of memory.

At this time within cognitive science and its allied disciplines, there are two quite different conceptual frameworks for thinking about learning and memory. These conceptual frameworks interact strongly with our conception of the functional architecture of the brain. Part of what motivates a commitment to one story about the functional architecture of the brain is a commitment to the corresponding story about learning and memory. Conversely, the assumption that the currently accepted story about the functional architecture of the brain is the right story biases one strongly toward one of the stories about the nature of learning and memory.

In the first story about the nature of learning, which is by far the more popular one, particularly in neurobiologically oriented circles, learning is the rewiring by experience of a plastic brain so as to make the operation of that brain better suited to the environment in which it finds itself. This is the story that adheres to the "opening the pathways" or "reinforc[ing] the conditioned connections" view of memory.

In the second story, learning is the extraction from experience of information about the world that is carried forward in memory to inform subsequent behavior.

In the first story, the brain has the functional architecture of a neural net. In the second story, it has the functional architecture of a Turing machine.

Learning As Rewiring

The first story about learning is plausibly exemplified by the well known experiments by Pavlov on conditioning the salivary reflexes of dogs. Pavlov came to his experiments after a long study of the reflex secretions of stomach juices that aid in digestion, because his original interest was in the physiology of digestion. He won the Nobel prize for his work on digestion, but today he is much better remembered for his work on the foundations of learning. The digestive reflexes he studied can be understood in terms of the reflex pathways described by Sherrington in the classic behavioral experiments that delineated the integrative properties of the synaptic connections between neurons (Figure 1). That functional architecture remains the dominant one in contemporary attempts to understand the function of the brain. It is the architecture we describe in the chapter on computation by neural nets.

The first element of a reflex arc is a sensory neuron. It's key function to confer sensitivity to environmental conditions by means of differential sensitivity to stimuli. A sensory receptor responds to some stimuli and not to others. The receptors in the retina, for example, are sensitive to light but not to the chemical composition of the fluid within the eye. The receptors in the nose and tongue, by contrast, are not sensitive to light, but they are sensitive to extremely low concentrations of certain molecules absorbed into either the nasal mucosa or the present in the fluid bathing the tongue.

The second functional elements are interneurons. The function of interneurons is to integrate signals from many different sensory neurons and to send signals reflecting the results of this integration to many different motor neurons and/or to other interneurons, including themselves. The first function, the bringing together of many different inputs, is called convergence. The second, relaying the results of convergence to many different outputs is called either divergence or fanning. An interneuron can also send a signal back to itself, either directly or by way of a loop through other neurons. A signal that circles back to its source is called a *recurrent* signal. As we will see in the next chapter, recurrent signals play a fundamental role in contemporary nueral net models.

The function of motor neurons is to translate the results of the neural integration that has occurred among the interneurons into observable outputs, that is, muscular contractions and relaxations and glandular secretions (like salivation).

In contemporary neural net modeling, these three kinds of functional elements are commonly called the input nodes (or input layer), the hidden layer, and the output layer. The layers in question are layers of schematic neurons.

Figure 1. *The functional architecture of a reflex arc. S = sensory neuron (selectively sensitive input element); I = interneuron (integrative element mediating convergence and divergence of signals); M = motor neuron (output element). When a path loops back to connect a neuron to itself, it is called a recurrent connection. The junctions between neurons are called synapses. Signal transmission within neurons is electrical, but transmission across the junctions (synaptic transmission) is chemical. The integration of signals occurs when multiple input pathways synapse on a common postsynaptic neuron. The dashed presynaptic input represents an input created (or changed in its conductance) by experience. This is also the functional architecture of a neural net, except that in a neural net each different kind of neuron (S, I and M) is replicated and arranged (conceptually) in layers, with the S neurons constituting the input layer, the I neurons the hidden layer, and the M neurons the output layer.*
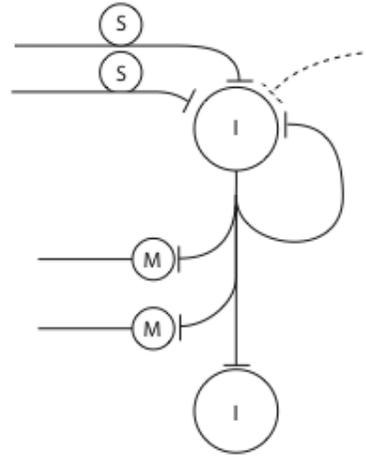
In the case of the digestive reflexes, the sensory neurons (input elements) are selectively sensitive to the chemical composition of the substances in the mouth, throat and stomach. The motor neurons (output elements) control the secretion of different digestive juices. The integration of input signals by the interneuron elements (elements of the hidden layer) determines the pattern of secretion, that is, the mixture of juices.

How strongly the incoming signal from a given sensory neuron affects the outgoing signal from an interneuron depends on the strength of the synaptic connection that the incoming neuron makes with the interneuron. If the signal arriving at a presynaptic ending (-| in the figure) releases a large amount of transmitter substance into the gap between the presynaptic ending and the postsynpatic neuron, and if there are many receptors sensitive to that substance in the membrane on the other side of the gap (the postsynaptic membrane), then that input signal has a strong effect on the postsynaptic output signal. Such a synapse is said to have a high *conductance*. A synapse at which little substance is released and/or there are few receptors in the postsynaptic membrane sensitive to the released substance is said to have a low conductance. The direction of a signal's postsynaptic effect, whether excitatory or inhibitory, depends on the transmitter substance released from the presynaptic membrane and on the receptors sensitive to that substance in the postsynaptic membrane.

The process of signal integration in the postsynaptic neuron is at least crudely analogous to arithmetic summation. Broadly speaking, the postsynaptic neuron adds up the positive and negative inputs. If the resulting sum exceeds a threshold value, the postsynaptic neuron transmits a signal to all the neurons with which it connects. The integration by the interneurons and the subsequent integration of interneuron input by output neurons determine the output pattern. In Pavlov's case, the output pattern was the relative amounts of digestive juices secreted.

In the course of his digestive studies, Pavlov noticed that in dogs that had been used repeatedly, secretions were observed to occur in response to stimuli that had nothing inherently to do with digestion. He noticed that the sight or sound of the experimental preparations for delivering a digestion-relevant stimulus (e.g., food on the tongue) triggered secretions even when the digestion-relevant stimuli were not in fact delivered. He realized that this was an example of learning. There are no genetically specified connections leading from neurons in the auditory system sensitive to sound to the motor neurons innervating the secretory glands of the digestive system, which is why sounds do not elicit secretions in the inexperienced dog. Pavlov assumed that this learning took to form of a rewiring of a system. The assumed potential for rewiring is indicated by the dashed input pathway in Figure 1. In modern thinking, this represents an input whose synaptic conductance is initially essentially 0, but which can be increased by means of a experience-sensitive conductance-altering mechanism. To this day, there are many who do not see how it is possible to draw any other conclusion from these simple behavioral observations. The stimulus in question (for example, the sound of food being prepared) does not elicit the secretory response in the experimentally naive dog, but it does do so in the dog who has repeatedly heard those sounds and then been given food. Must not the system in some sense have rewired itself? There once was no path from the ear to the stomach and now there apparently is.

Pavlov was understandably vague about exactly how this rewiring should be imagined to occur, but he set out to study the conditions under which it occurs. In doing so, he was guided by the congruence between what he seemed to have observed and one of the oldest and most popular ideas in the philosophical literature on the theory of mind, the notion of an associative connection. In the seventeenth century, the English philosopher and political theorist, John Locke, argued that our thoughts were governed by learned associations between 'ideas.' Locke understood by 'ideas' both what we might now call simple sense impressions, for example, the impression of red, and what we might now call concepts, such as the concept of motherhood. He called the first, simple ideas and the second complex ideas. Whereas rationalists believed that ideas were connected by some kind of preordained intrinsic logical system, Locke argued that the connections between our ideas were in essence accidents of experience. One idea followed the next in our mind because the stimulus (or environmental situation) that aroused the second idea had repeatedly been preceded by the stimulus (or situation) that aroused the preceding idea. The repeated occurrence of these two ideas in close temporal proximity had caused an associative connection to grow between them. The associative connection conducted excitation from the one idea to the other. When the first idea was aroused, it aroused the second by way of the associative connection that experience had forged between the two ideas. Moreover, he argued, the associative process forged

complex ideas out of simple ideas. Concepts like motherhood were clusters of simple ideas (sense impressions) that had become strongly associated with each other through repeated experience of their co-occurrence.

There is enormous intuitive appeal to this concept. It has endured for centuries and is as popular today as it was in Locke's day, probably more popular. Some years ago, I gave a talk arguing that associations did not exist. Afterwards, I had a long and passionate argument with a distinguished modern proponent of the associative theory of learning and of the corresponding story about memory and the functional architecture of the brain. Toward the conclusion of the discussion, the famous theorist asked rhetorically, "So how come when I hear 'salt', I think of pepper?" We tell this anecdote to illustrate the enduring intuitive appeal of Locke's argument and the influence it has on contemporary theories about the functional architecture of the brain and how that architecture explains behavioral phenomena. It has this appeal in part because on first hearing, it seems that it just has to be right. Our ideas do seem to become associated because they occur together.

The influence on Pavlov of this well known line of philosophical thought was straightforward. He translated the doctrine of learning by association of ideas into a physiological hypothesis. He assumed that it was the temporal proximity between the sound of food being prepared and the delivery of food that caused the rewiring of the reflex system, the formation of new connections between neurons. These new connections are the physiological embodiment of the psychological and philosophical concept of an association. Pavlov set out to vary systematically the conditions of this temporal pairing—how close in time the neutral (sound) stimulus and the innately active food stimulus had to be, how often they had to co-occur, the effects of other stimuli present, and so on. In so doing, he gave birth to the study of what is now called Pavlovian conditioning. The essence of this approach to learning is the arranging of a predictive relation between two or more stimuli and the study of the behavioral changes that follow the repeated experience of this relationship. These changes are imagined to be the consequence of some kind of rewiring within a functional structure like that portrayed in Figure 1.

The functional structure portrayed in Figure 1 is not that of a representational system. There is no symbolic memory; hence, no mechanism for carrying forward in time the information gleaned from past experience. It is extremely important to understand precisely why this is so. The only thing in the functional structure in Figure 1 that could serve as a symbolic memory is the experience-sensitive synaptic connection represented by the dashed synaptic input. A physical mechanism that is to function as an enduring symbolic memory must be capable of adopting more than one enduring physical state. By enduring, we have in mind time scales measured in days, months and years. That is the time scale appropriate to what Pavlov studied and, more generally, to learning as a behavioral phenomenon.

The only element in Figure 1 that can be enduringly placed in more than one state depending on the past experience of the system is the synapse represented by the dashed input, the synapse whose conductance can be changed by experience. However, for the changes in synaptic conductances to function as symbols, the mapping from experience

to the state of the symbol must be invertible, that is, it must be in principle possible to deduce from the state of the synapses an aspect of the experience that produced that state—the aspect represented by the symbol. An encoding procedure must be such as to make decoding possible. The complex processes by which the brain retrieves highly encoded representations of the past were the "mnestic activity" that was the focus of Luria's (1973) chapter on Memory. But the changes in synaptic conductance cannot encode anything.

What Pavlov and those who have followed his experimental lead have discovered is that many different aspects of experience combine to determine the apparent strength of the association between two stimuli. The strength of the association--that is, the magnitude of the change in synaptic conductance—appears to depend on the delay between the sound and the food, on the strength of the sound, on the tastiness of the food, on the interval between successive pairings of sound and food, on how many pairings of sound and food there have been, and on many other aspects of the animal's experience. Thus, the strength of an association depends on many different and quite unrelated aspects of the animal's experience. The values of those different experiential variables are all confounded—run together, mixed up—in the determination of the strength of an association, Thus, from the resulting strength of a single association, it is impossible to determine the value of any single one of those variables. The function relating the strength of an association to experience is a many-one function. It maps the values of many different experiential variables to a single variable, the strength of the association. Many-one functions are not invertible. You cannot get from the value of the one back to the values of the many.

Moreover, even if the mapping from experience to the strengths of malleable synapses were in principle invertible, the architecture in Figure 1 does not give the system access to the state of its synapses. Put another way, the states of the synapses are not accessible to computation.

The issue of accessability to computation is a critical issue. The processes in the postsynaptic membrane that integrate the inputs from different synapses are computational in that they combine the inputs to yield an output. Indeed, as already noted, the combinatorial operation generally assumed in current models based on the architecture in Figure 1 is arithmetic addition of the weighted input signals, followed by a decision based on the sum. However, these postsynaptic computational processes do not have access to the synaptic conductances. What they have access to is the term by term product of the synaptic conductances and the presynaptic signals. In mathematical language and notation, there are two vectors that combine to determine the signals that are postsynaptically integrated. The first vector specifies the strengths of the signals arriving at the presynaptic endings. The second vector specifies the conductances of the synapses that must be crossed to reach the postsynaptic integration processes.

The first vector may be symbolized $\langle S_1, S_2, ... S_n \rangle$ and the second $\langle w_1, w_2, ... w_n \rangle$. In this symbolization, S_1 stands for the strength of the signal arriving at the first synapse, S_2 stands for the strength of the signal arriving at the second synapse, and so on, S_n standing for the strength of the signal arriving at the nth synapse. (Signal strength may be thought

of as the number of spikes arriving per second.) Similarly, w_1 stands for the conductance of the first synapse, w_2, the conductance of the second synapse, and so on. (The conductance is the size of the postsynaptic effect of a presynaptic spike; the bigger the effect, the higher the conductance of the synapse.) If there are any enduring symbols in the architecture portrayed in Figure 1, then the vector of synaptic conductances, $\langle w_1, w_2, ... w_n \rangle$, is the equivalent of the strings of 0s and 1s in a binary symbol stored on Turing's tape. What the integrative postsynaptic process sees, however, is not this vector. What it sees is $\langle w_1 S_1, w_2 S_2, ... w_n S_n \rangle$. If it does not know the strengths of the presynaptic signals, $\langle S_1, S_2, ... S_n \rangle$, then it cannot determine $\langle w_1, w_2, ... w_n \rangle$ from $\langle w_1 S_1, w_2 S_2, ... w_n S_n \rangle$. If the postsynaptic computational processes have no way of knowing what $\langle S_1, S_2, ... S_n \rangle$ was—and in the architecture in Figure 1, they do not—then there is no way that these processes can decode $\langle w_1 S_1, w_2 S_2, ... w_n S_n \rangle$.

In short, there are fundamental reasons why the architecture in Figure 1 is not the architecture of a representational system and why the only possible enduring symbols in Figure 1, the synaptic conductances modifiable by experience, are not in symbols. This is why those who think that learning is the rewiring of a plastic brain by experience shy away from representational theories of brain function. The anti-representational strain in this line of thought has been evident in its every manifestation during the past 100 years. Early champions of this view, like Clark Hull (1929; , 1952) and B. F. Skinner (1938; , 1957) went to great lengths to argue that behaviors that looked like they were informed by a representation of the subject's environment could in fact be better understood as chains of stimulus-response events involving no underlying representations. More recent versions of this view have been similarly antirepresentational (Rumelhart & McClelland, 1986), arguing that neural processing is "subsymbolic" (cite Smolensky chapter). The intuitive essence of the anti-representational view is that animals act intelligently because of the rewiring effects of experience on their brains, but they do not know why what they do is intelligent because they have no enduring representation of the aspects of the world to which their action is adapted.

Learning As the Extraction and Preservation of Useful Information

The second story about learning is exemplified by dead reckoning, which is a key component of animals' navigation. Dead reckoning is a simple computation, for which effective (machine implementable) procedures have long been known and routinely used. It illustrates the fundamental role that a symbolic memory plays in computation. It is important to our argument because there is strong behavioral evidence that this procedure is implemented in the brains of a wide range of animals, including many insects. Indeed, the best behavioral data come from studies on ants and bees.

Figure 2. *Tracing of the track of a foraging Cataglyphis bicolor. N = nest hole. X = location of food. Solid line = outward journey; dashed line = homeward journey. Distance scale on the order of 10 m (After Harkness & Maroudas, 1985.)*

Figure 2 is a tracing of the track of a foraging ant of the species *Cataglyphis bicolor*, which inhabits the Tunisian dessert. The ant emerged from the nest, struck off to the northwest for several meters, then broke into the tortuous circling pattern characteristic of this ant when it is searching for something. At the spot marked with an X, it found the carcass of a scorpion. It bit off a chunk to carry back to the nest. The homeward journey is traced by the dashed line. Clearly, the ant knew the direction home. Experiments, in which the ant is captured just as it starts for home, carried across the dessert and released in unfamiliar territory show that it also knows the distance to the nest (Wehner & Srinivasan, 1981). The released ants run the compass course from the capture point at the food to the nest, which is, of course, not the course to the nest from the release site. This proves that they do not require familiar landmarks to set and maintain their nestward course. When they are a meter or two past where the nest should have been encountered, they stop and begin to search for it. This proves that they do not need the familiar landmarks surrounding the nest in order to estimate that they are in the vicinity of the nest, because, in the unfamiliar terrain to which they have been experimentally transported, those landmarks are absent.

The ants learn from their experience the compass direction and distance of their nest from the site at which they find food. This information cannot have been carried in their genes. The questions, then, are: What is the experience from which they learn? How do they learn from it? What does this tell us about the nature of learning? What does it tell us about the functional architecture of the brain?

The experience from which they learn is the experience of moving themselves from location to location en route to finding the food. The learning process is the continual updating of a representation of their current location by summing the successive small displacements (small changes in position) by which they have reached it. This tells us that learning is the extraction from experience of symbolic representations, which are carried forward in time by a symbolic memory mechanism, until such time as they needed in a behavior-determining computation. This in turn tells us that the functional architecture of the brain must have as one of its basic elements a symbolic read-write memory mechanism, a mechanism whose structure suits it to carry

information forward in time in a manner that makes the information accessible to computation.

Updating an estimate of one's location. The ant is like almost all other animals whose navigational behavior has been subject to experimental study in that it keeps an ongoing estimate of its current location by updating that estimate as it moves. The mathematical term for this is *path integration*, but in English texts on the principles of navigation, it is commonly called *dead reckoning*, which is thought to be a corruption from the abbreviation 'ded. reckoning,' for 'deduced reckoning.' The term reckoning already tells us that it is a computational process. It is a simple computation. It depends on updating one's displacement vector step by step, adding each small displacement to the sum of the previous displacements to obtain net displacement. The function of memory is to carry the vector sum forward in time in a manner that makes it possible to add new displacements to it, as they occur.

The learning mechanism is inherently modular, because it takes different computations to extract different representations from different data. The structure of the learning mechanisms that perform the computations must reflect analytic aspects of the computational problem. In the present example, the mechanism that extracts a representation of current position from the experience of getting to it integrates the velocity vector with respect to time. This process in the brain reflects an analytic fact about the world, namely that position as a function of time is the integral of velocity as a function of time.

Other learning mechanisms have structural features that are similarly reflective of the universal analytic aspects of the computational problem that a learning mechanism must solve. For example, the mechanism that learns the solar ephemeris—the sun's compass direction as a function of the time of day—has built into it the universal truth that, regardless of where one is in the world, the sun is in the east in the morning and in the west in the afternoon. What varies is how the sun gets from east to west: in the northern hemisphere above the tropics, it moves from the east clockwise around the horizon into the south at noon and then continues around the horizon into the west in the afternoon. In the southern hemisphere below the tropics, it moves counterclockwise around the horizon into the north at noon and thence into the west. In the tropics, it stays more or less in the east all morning, rising to almost directly overhead at noon, and then sinks into the west during the afternoon. The universal aspect is built into the mechanism by which animals learn the locally appropriate solar ephemeris (Dyer & Dickinson, 1994). What the learning mechanism extracts from experience are the values of the parameters of the universal function, the values that adapt the universal function to the local ephemeris. Again, the function of memory is to carry the values of these parameters forward in time, so that they can continue to inform navigational behavior. By reducing the learning problem to one of estimating a few parameters, the universal ephemeris built into the learning mechanism makes it possible for animals to learn the local function from very limited experience. What they learn transcends what is contained in the experience from which they learned it. Thus, for example, bees that have never seen the sun in the morning, only in the afternoon, nonetheless believe that in the morning it will be in the east, not in the west, where they have so far seen it (Dyer & Dickinson, 1994).

A similar story applies to the mechanism by which humans learn language (Baker, 2001). A universal grammar that specifies the binary dimensions along which the grammars of human languages may vary is built into the learning mechanism. What the learning mechanism extracts from the experience of the local language are the values that these parameters have in that language. Other animals close to us on the evolutionary bush lack this learning organ, which is why they cannot learn a human language. What humans extract from limited experience of the local language is like what bees extract from limited experience of the local solar ephemeris in that it transcends the evidence from which it is extracted. Chomsky (1980) coined the term "poverty of the "stimulus" to refer to the ubiquitous finding that the conclusions learners extract from data are not justified by those data. To explain the conclusions drawn, one has to take into account the innate problem-specific structure of the learning mechanism that made the induction.

It turns out that even in the learning that Pavlov studied, the learning that underlies conditioned reflexes, there is a domain-specific learning mechanism at work. Although Pavlov and his followers generally paid no attention to the formal structure of the problems that they set the subjects in their learning paradigms, the problems in fact fall within a well-specified domain, the domain of multivariate, non-stationary time series analysis (C.R. Gallistel, 1999). Traditional general-purpose associative theories of learning have never been able to account for some of the most basic aspects of the learning that mediates the appearance of conditioned reflexes. For example, they have never been able to define temporal pairing, even though the notion of temporal pairing has always been taken to be a sine qua non of association formation (C.R. Gallistel, 2007). A learning mechanism whose structure reflects the demands inherent in such an analysis (C. R. Gallistel & Gibbon, 2002) does a much better job of explaining the quantitative facts that have emerged from the tens of thousands of research papers in this area mentioned by Luria in the quote with which this essay opened.

Summary

There are two conceptions of the nature of learning intimately linked to two conceptions of the functional architecture of brains. In the first conception, learning is a rewiring process. Experience rewires the brain so as to make behavior better adapted to the experienced environment. The rewiring does not represent that environment. It does not represent anything. It simply alters the mapping from brain inputs to brain outputs. On this conception, the functional architecture of the brain is the architecture described in any textbook on neurobiology. This congruence with our current understanding of neurobiology is a great strength of this conception. Its great weakness is the difficulty it has in giving a coherent account of well-documented behavioral phenomena like dead reckoning, which seem to require for their explanation the assumption that the brain is capable of representing simple behavior-relevant aspects of the animal's situation, such as its position relative to home.

In the second conception, learning is the extraction from experience of behaviorally useful information about the environment, which information is carried forward in time in a symbolic read-write memory. The great strength of this conception is the ease with which it explains the many well-documented behavioral phenomena that

imply a symbolic representation of the world and the animal's situation in it. It is the foundation for the complex mnestic activity that Luria discusses in his many works on memory. The weakness is that this conception requires us to assume that a symbolic read-write memory is a fundamental part of the functional architecture of a brain (including an insect brain). One would look in vain in contemporary textbooks of neurobiology for a well-documented mechanism well-suited for this function. The essential question then becomes, Whose problem is this? Is it a problem for cognitive psychologists, who conceive of their task as the development of computational theories of brain function capable of explaining behavioral data? Or is it a problem for neurobiologists, who conceive of their task as the discovery of those neurobiological mechanisms that explain the fundamental behavioral capacities of brains? If among those capacities demonstrably depend on a symbolic read-write memory, then it is the neurobiologists' problem.

References

Baker, M. C. (2001). *The atoms of language*. New York: Basic Books.

Chomsky, N. A. (1980). *Rules and representations*. New York: Columbia University Press.

Dyer, F. C., & Dickinson, J. A. (1994). Development of sun compensation by honeybees: How partially experienced bees estimate the sun's course. *Proceedings of the National Academy of Sciences, USA, 91*, 4471-4474.

Gallistel, C. R. (1999). The replacement of general-purpose learning models with adaptively specialized learning modules. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences. 2nd ed.* (2 ed., pp. 1179-1191). Cambridge, MA: MIT Press.

Gallistel, C. R. (2007). Flawed foundations of associationism? Comments on Machado and Silva. *American Psychologist, 62*(7), 682-685.

Gallistel, C. R., & Gibbon, J. (2002). *The symbolic foundations of conditioned behavior*. Mahwah, NJ: Lawrence Erlbaum Associates.

Harkness, R. D., & Maroudas, N. G. (1985). Central place foraging by an ant (Cataglyphis bicolor Fab.): a model of searching. *Animal Behaviour, 33*, 916-928.

Hull, C. L. (1929). A functional interpretation of the conditioned reflex. *Psychol. Rev., 36*, 498-511.

Hull, C. L. (1952). *A behavior system*. New Haven, CT:: Yale University Press.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology* (B. Haigh, Trans.). New York: Basic Books.

Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel distributed processing*. Cambridge, MA: MIT Press.

Skinner, B. F. (1938). *The behavior of organisms*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1957). *Verbal Behavior*. New York: Appleton-Century-Crofts.

Wehner, R., & Srinivasan, M. V. (1981). Searching behavior of desert ants, genus *Cataglyphis* (*Formicidae*, Hymenoptera). *Journal of Comparative Physiology, 142*, 315-338.