Place, U. T. (1991). Error-correction in connectionist networks: A new perspective on the law of effect. [Unpublished paper. Presented to the Annual Conference of the British Psychological Society, Bournemouth, 12th April 1991, Session on Behavioristic Perspectives on Cognitive Psychology and to the 17th Annual Convention of the Association for Behavior Analysis, Atlanta, Georgia, May 26th 1991. Conference presentation abstract appeared in the Proceedings of the British Psychological Society 1991 Abstracts, 67.]

ERROR CORRECTION IN CONNECTIONIST NETWORKS: A NEW PERSPECTIVE ON THE LAW OF EFFECT.

Ullin Place, University of Wales, Bangor.

The parallel distributed processor (PDP) is a pattern discrimination learning device in which ‘information’ is stored, not in a localized memory store, but in the form of changes in the ‘weights’ of synaptic connections between the large number of semi-conductor units or ‘nodes’ (corresponding to the neurons of the central nervous system) of which the device consists. The properties of such a device, whereby it learns to recognize complex patterns of sensory input, derive from the so-called ‘learning rules’ (McClelland and Rumelhart 1988) which determine the changes in the synaptic weights which result from each learning experience.

According to McClelland and Rumelhart, there are two such rules which have been followed in giving networks the capacity to learn,

> the so-called Hebbian or correlational learning rule .... and the error-correcting or ‘delta’ learning rule. (McClelland and Rumelhart *op.cit.* p.83)

It is not altogether surprising to find that these learning rules, adopted by connectionists in order to give a PDP its capacity to learn, correspond rather precisely both to the principles proposed by learning theorists on the basis the experimental analysis of learning in animals, as well as to those which go back to the earliest recorded attempts to formulate the principles of learning in humans. McClelland and Rumelhart trace back their correlational learning rule to Hebb (1949) and James (1890). They could equally well have cited the principles of classical conditioning as described by Pavlov (1927) and, before that, the principle of association by contiguity which can be traced back to Aristotle. Likewise the error-correcting or "delta" rule is a thinly disguised version of the Law of Effect as formulated by Thorndike (1911) which, in the form of the principle of psychological hedonism, goes back to Epicurus.

Once these identifications are made, it becomes apparent that one of the effects of adopting the PDP as an alternative to the conventional serial-digital computer as the preferred model for the way the brain functions in its control of behavior, is to bring back into the focus of theoretical concern many issues, issues concerning the principles required to explain the phenomenon of animal learning which dominated psychological theory in the 1940's and 1950's and which disappeared from the agenda of psychological theory when behaviorism was overtaken and submerged by the so-called ‘cognitive revolution.’

One issue which was extensively discussed in days gone by is the question whether we really need two learning principles and, if we do, whether these two principles alone are enough to account for all the phenomena of learning. Can we, as Pavlov (1927), Watson (1914) and Guthrie (1935) thought and as some contemporary connectionists are inclined to think, reduce everything to some version of the contiguity or correlational principle? Or should we follow Hull (1943) in trying to reduce all learning to some version of the Law of Effect? If we decide with Thorndike (1911), Skinner (1938) and McClelland and Rumelhart (1988) that two principles are needed, which two principles we are talking about, and how are we to construe the division of labor between the two?

Input Selection and Output Selection

Let us assume provisionally

(1) that McClelland and Rumelhart's "Hebbian or correlational learning rule", as used in the construction of connectionist networks, is identical with the classical principle of association by contiguity, Tolman's (1932) principle of "sign-gestalt-expectation" and the principle of "SS contiguity learning" (Spence 1951), as postulated to explain learning phenomena in the behavior of living organisms (to be referred to hereafter as the ‘correlational/contiguity learning rule’),

(2) that McClelland and Rumelhart's "error-correcting or ‘delta’ learning rule", as used in network construction, is likewise identical with the Law of Effect as proposed by Thorndike (1911) to account for the phenomenon of trial-and-error learning in the behavior of animals

(to be referred to hereafter as the ‘error-correcting/effect learning rule’), and, for the sake of argument,

(3) that these two principles are both necessary and jointly sufficient to explain the learning process in a parallel distributed processor (PDP) whether artificially constructed or naturally occurring as in the brain.

The division of labor between these two principles can then be explained by invoking a distinction between two kinds selection process within the overall process of input-to-output transformation as it manifests itself in the molar behavior of living organisms, namely:

(a) the process of ‘input selection’ or ‘selective attention’, and

(b) the process of ‘output’ or ‘behavior selection’.

Given this distinction, it is clear

(a) that the correlational/contiguity learning rule is an input-selection learning rule, and

(b) that the error-correcting/effect learning rule is an output-selection learning rule.

Input Selection and the Correlational/Contiguity Learning Rule

The process of *input selection* or *selective attention* is what I have called (Place 1985) a "semi-covert" behavioral process. It consists partly in overt movements of the sense organs, such as the eyes or the fingers, so as to maximize stimulation from a particular aspect of the organism's current stimulus environment and partly in covert adjustments in figure-ground relations within what has been called ‘the proximal stimulus’ within the brain. The function of this process is to determine what aspect of the total stimulus input at any moment of time is to control and determine the subsequent process of output selection.

An *input-selection learning rule* is a principle which governs changes in the input selection properties of a network/brain which develop over time as a consequence of previous inputs into the system. Considered as such an input-selection learning rule, we may state

The correlational/contiguity learning rule

> If two stimulus event types S_1 and S_2 impinge on an organism's receptors in close temporal proximity such that the onset of S_1 is regularly followed after a short and constant interval by S_2, the organism will learn to expect S_2 whenever S_1 is presented.

Positive evidence of the operation of this principle in the behavior of living organisms is provided by the phenomenon of the classical (respondent) conditioning. In this case, S_1 in the above formula corresponds to the *CS* (conditioned stimulus) and S_2 to the *UCS* (unconditioned stimulus). The resulting expectation manifests itself in the *CR* (the conditioned autonomic response - salivation, GSR, pupillary reflex, etc.) which develops after repeated pairings of the *CS* and *UCS*. Negative evidence of the operation of this principle is provided by the phenomenon whereby an *unexpected* stimulus event invariably attracts the organism's attention.

Principles of Input Selection

The principle whereby attention is caught by the unexpected is not so much a learning rule as an *input-selection principle*, a principle which determines what aspect of the total input is selected at a particular moment, rather than one which determines changes in the selective properties of the system over time. We thus have

> *The Principle of the Pre-Emptive Selection of the Unexpected Input*
>
> Where a stimulus event of type S_1 has repeatedly occurred in close temporal contiguity with a stimulus event of another type S_2 such that the organism has learned, in accordance with the correlational/contiguity learning rule, to expect an event of type S_2, given the occurrence of an event of type S_1, and a stimulus event of type S_1 has occurred followed immediately, not by a stimulus event of type S_2 as expected, but by a stimulus event of another type S_3, the organism's attention will focus on the *unexpected* stimulus event of type S_3.[1]

As we shall see later, this distinction between a *principle of selection* which determines what input or output is selected on a particular occasion and a *learning rule* which determines how the selective properties of the system change over time has an important role to play in the analysis of the process

---

[1] There would appear to be two other input-selection principles:

(a) the principle of salience whereby attention is attracted to stimulus features which stand out in sharp contrast to their background, and

(b) the principle of motivational significance whereby attention is attracted to stimulus features which are motivationally significant to the organism, both those that are attractive and those that are repulsive.

of output selection and the part played in that process by the error-correcting/effect learning rule. It is to this topic that we must now turn.

Output Selection and the Error-Correcting/Effect Learning Rule

The process of *output* or *behavior selection* is the process whereby an instrumental or ‘operant’ motor response is selected for emission by the organism at a particular occasion. Which output/behavior is selected is determined partly by the nature of the input which has already been selected as demanding a response and partly by the motivational attitude of the organism to the various possible outputs which suggest themselves based on the prevailing *establishing conditions*, as suggested by Michael (1982).[2] It is in this connection that we encounter the *error-correcting/effect learning rule*.

The original purpose of McClelland and Rumelhart's "error-correcting or ‘delta’ learning rule" is to give to a network the ability to learn a discrimination by a process of trial and error in which, after an initial phase of responding at random, the device gradually learns to eliminate errors and make only correct responses. Such discrimination learning requires two conditions:

(a) a sensorium which is sensitive to the differences between the positive cases (Skinner's S^D) and the negative cases (Skinner's S^Δ),

(b) a feed-back mechanism from the environment which "tells" the system when it has made a correct response and when it has made an error.

The effect of the ‘delta’ rule is that, when an error-message is received, the weights of the synaptic connections which were activated by the initiation of the immediately preceding response are reduced by a process of ‘back-propagation of error-correction.’ This process starts at the output end of the network and works backwards through it towards the input end weakening the weights of the connections involved in the erroneous response as it goes. Given the additional principle that the

---

[2] For those not familiar with this important concept, a good example of an establishing condition is the situation in which one is confronted by a plate of food without the cutlery needed to consume it. Under this condition any behavior expected to result in obtaining cutlery has a high probability of being selected.

weight of any activated connection which is not weakened in this way will be strengthened, it is easy to see how, after a number of trials, the tendency to respond correctly would be enhanced with a gradual elimination of errors.

Serious doubts have been raised concerning the principle of the back-propagation of error-correction. These doubts relate in part to the difficulty of envisaging how such a process might be realized in terms of any process which we know about in brain, and in part to concerns about the limitations to the pattern-discrimination properties of networks which appear to be due to the adoption of this principle (Chater 1989). What is not in doubt is the reality of the process of pattern-discrimination learning by random-trial-and-progressive-error-correction which is manifested by a network whose synaptic weights are changed in accordance with this principle. Nor is there any doubt that at the molar behavioral level this process of trial-and-error learning is the same process which has been familiar to students of animal behavior since it was first described by E. L. Thorndike (1898) in the light of his classical study of cats learning to escape from a puzzle box (Figure 1).

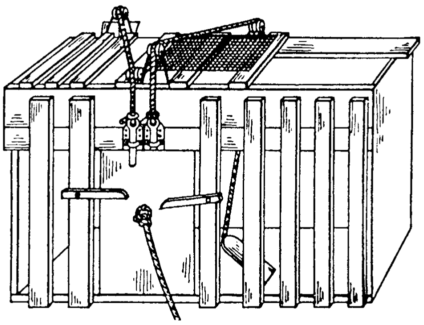

Figure 1. A Thorndike Puzzle Box

The first clear demonstration of process of trial-and-error-correction in pattern discrimination learning comparable to that found in connectionist networks was by Lashley (1930) using the so-called 'jumping stand' (Figure 2). In this situation, a rat is compelled by a blast of air from behind to jump towards one or other of two cards displaying different patterns. It receives a 'correct' message in the form of the collapse of the card allowing it to gain access to the food behind it, if it makes what is defined by the experimenter as the correct response. It receives an 'error' message in the form of bumping its nose against a card which is locked into place and falling down into the net below, when it makes what is defined as an error. It was left, however, to Skinner to provide the classical theoretical analysis of this type of instrumental or, as he calls it, "operant" discrimination learning in Chapter Five of *The Behavior of Organisms* (Skinner 1938).

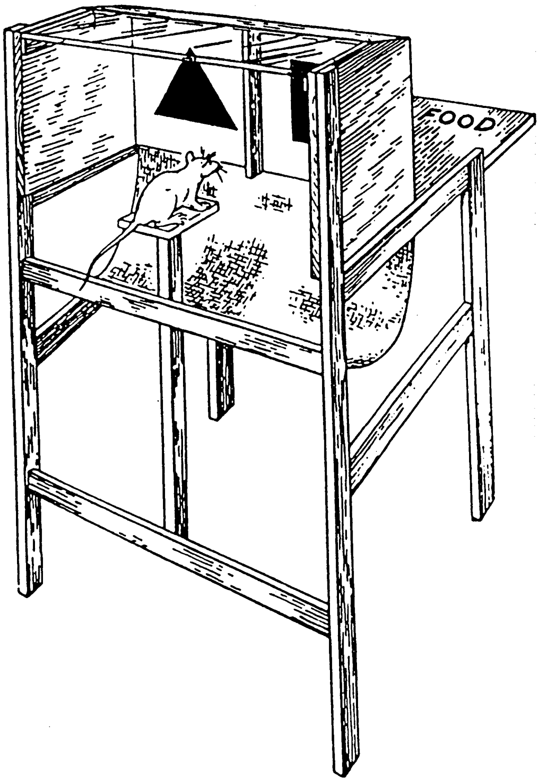

Figure 2. The Lashley Jumping Stand

What this shows us is that the process of trial-and-error-correction whereby a connectionist network learns to recognize patterns is indistinguishable from the process of pattern recognition as it is observed in experimental studies of animal behavior. There are, of course, differences. Subsequent learning to discriminate another pair of stimuli interferes with previous learning in an artificial network in a way and to an extent which it does not do in the case of a living organism. Artificial networks have to be designed differently in order to discriminate a diachronically presented pattern such as a melody, in contrast to a synchronically presented pattern such as a musical chord or a visual scene; whereas brains can handle both equally well, without any apparent change of gear. But these differences do not affect the essential identity of the learning process itself in the two cases. It is not surprising, therefore, to find that the principle of the back-propagation of error-correction, as used to generate such learning in a PDP, bears a remarkable resemblance to the "Law of Effect" as originally proposed by Thorndike (1911) in order to account for trial-an-error learning in the cat.

Picking out what is common to the two principles and presenting the result as a learning rule for the discrimination of two classes of stimuli ***S+*** and ***S-***, we can state

> *The Error-Correcting/Effect Learning Rule*
>
> If an output (response) of type ***R+*** is emitted in the presence of an input (stimulus) of type ***S+*** and that combination is regularly accompanied in close temporal proximity by a 'correct' message or reinforcement, whereas the emission of an ***R+*** in the presence of an input of a discriminably different type ***S-*** is regularly accompanied in close temporal contiguity by an 'error' message or "disinforcement" (to use the terminology proposed by Harzem and Miles, 1978[3]), and if an output (response) of type ***R-*** which is antagonistic to ***R+*** is emitted in the presence of an ***S-*** and *that* combination is regularly accompanied by a 'correct' message or reinforcement, whereas the emission of an ***R-*** in the presence of an ***S+*** is regularly accompanied by an 'error' message or disinforcement, a network whether artificial or naturally-occurring will gradually learn to emit ***R+*** and not to emit ***R-*** in the presence of ***S+***, and to emit ***R-*** and not to emit ***R+*** in the presence of ***S-***.

The Law of Effect as a Derivative Principle at the Molar Level

So stated, the error-correcting/effect learning rule is a principle of output- or behavior-selection whose operation at the molar level of analysis at which the organism-environment relations are observed cannot be denied. As the philosopher Daniel Dennett observes in his book *Brainstorms* (Dennett 1978) "the Law of Effect will not go away." Nevertheless, unlike the case of the artificial connectionist network in the back-propagation of error-correction is a fundamental learning rule governing changes in synaptic weights at the molecular level, there is reason to think that in the case of living organisms the Law of Effect is not a fundamental principle at this level of analysis. It is

---

[3] Harzem and Miles propose the term "disinforcement" primarily as a substitute for the more traditional term `punishment' on the grounds that it avoids the unfortunate emotive connotations of the latter term, while at the same time emphasizing the parallel and contrast with reinforcement. Disinforcement parallels reinforcement in that it is an immediate consequence of emitting or failing to emit an operant response which has an effect on the strength of the organism's propensity to emit that response on similar occasions in the future. It differs from reinforcement in that the effect is to weaken rather than strengthen the response tendency. More recently Professor Miles (Emeritus Professor T.R.Miles of the Department of Psychology, University of Wales Bangor) has accepted (personal communication) the present author's suggestion that the use of the term be extended so as [to] cover *any* immediate consequence of behavior (including non-reinforcement in conditions where an expectation of reinforcement has been established, as in the experimental extinction of a positively reinforced operant) which has the effect of weakening the tendency to respond in that way. However, Professor Peter Harzem of Auburn University who first proposed the term is not persuaded of the utility of this extended usage.

rather an emergent principle at the molar organism-environment level of analysis which results from the operation, at the relatively molecular level of analysis at which synaptic weights are changed, of three distinct principles:

(1) the Principle of Contingency Expectation,
(2) the Principle of Psychological Hedonism, and
(3) the Law of Exercise (Thorndike 1911).

This is the conclusion to which we are led when we consider what are the counterparts in the case of a living organism learning to distinguish stimulus patterns in the natural environment of the 'correct' and 'error' messages which need to be fed back to an artificial network, if it is to be trained to make this kind of discrimination. When this question is asked, two things become clear:

(1) in an organism operating in the natural environment, the correct/error-message is constituted - in the case of an organism which has not acquired the ability to respond to verbal stimuli and whose behavior is "contingency-shaped" in Skinner's (1966) sense[4] - by *the immediate consequences of behavior* - broadly speaking what happens or fails to happen within 15 seconds of the emission of the affected behavior on the evidence of Perin (1943), or 0.5 seconds, if the effect of expectations based on the correlational/contiguity principle are excluded as in Grice (1948);
(2) the immediate consequences of behavior are treated by the organism as a correct-message (reinforcement) or as an error-message (disinforcement) depending on the motivational attitude of the organism to those consequences.

In other words, if an animal 'likes' the consequences of behavior, it receives a 'correct' message when those consequences register; if it 'dislikes' the consequences, an error-message is received. In the real world there is no such thing as an error or correct-message which is defined as such independently of the organism's motivational state.

---

[4] One of the effects of "specifying" the consequences of behavior verbally in the form of what Skinner calls "a rule" is to bring behavior under the control [of] its more remote consequences.

The fact that it is the motivational attitude of the organism to the immediate consequences of behavior which determines whether those consequences are ‘read’ by the organism as a ‘correct’ message (reinforcement) or as an ‘error’ message (disinforcement) shows us that when a living organism learns a discrimination by means of the process of trial-and-error-correction, it does not simply learn to *emit* the correct responses and *omit* the errors. What it learns to do is to expect a three-term sequence of events of the type which Skinner (1969) calls "a contingency."

A *contingency* for Skinner is a sequence of states of affairs and events whereby, given certain *Antecedent* conditions, *Behaving* in a certain way, has certain *Consequences*[5]. Whenever, in the course of its interaction with its environment, an organism repeatedly encounters a particular contingency in this sense, the correlational/contiguity learning rule, as stated above, allows us to deduce

The Principle of Contingency Expectation

> Given that the combination of an Antecedent condition with the sensory feedback from the Behavior emitted under that condition (S_1) has been regularly and closely followed by a certain Consequence (S_2), a network, whether artificial or natural, will learn to expect the Consequence (S_2), given the Antecedent condition combined with the sensory feedback from an incipient evocation of the Behavior regularly emitted under that condition (S_1)[6].

---

[5] Skinner defines "a contingency" as a relation between a Stimulus, a Response and a Reinforcement. The more generalized conception of a contingency as a relation between Antecedent, Behavior and Consequence which is employed here derives, as far as the writer is concerned from Dr. Ogden Lindsley of the University of Kansas. It is to be preferred to Skinner's own formulation on the grounds that

(a) the concept of an Antecedent allows us to include an establishing condition (Michael 1982), such as a state of food deprivation which is not a stimulus,

(b) the concept of Behavior allows us to include non-responding alongside responding, and

(c) the concept of a Consequence allows us to include disinforcing consequences (Harzem and Miles 1978), such as non-reinforcement or punishment, in addition reinforcing ones.

[6] This formulation is to be preferred to that proposed by Dickinson (1988) in which what is learned is the *propositional attitude* or ‘belief’ that S_1 will be followed by S_2. This is because propositions, such as that expressed by the embedded sentence represented by the letter p in the sentence frame

O believes that p,

are linguistic entities constituted by the class of all possible sentences in any natural language which are semantically equivalent to p. It follows that this way of talking, when taken strictly and non-metaphorically, presupposes both linguistic competence on the part of the behaving organism and linguistic control of the behavior which is thereby selected. This presupposition is avoided if the matter is formulated by using the verb ‘expect’, provided that what is expected is an event rather than that a proposition be true.

Although it is deduced from the correlational/contiguity learning rule, this principle is not itself a learning rule. Moreover, although it combines with an output selection principle, the Principle of Psychological Hedonism, to determine output selection, it is an input selection rather than an output selection principle.

Given such an expectation, the output which is selected will be determined by

> *The Principle of Psychological Hedonism*
>
> If a particular Consequence is expected on the basis of sensory feedback from the incipient evocation of a particular pattern of Behavior (***R+***), and the prevailing Antecedent establishing conditions are such that the motivational attitude of the organism to that Consequence is favorable, the Behavior in question (***R+***) will be selected and completed; if, on the other hand, the motivational attitude to that Consequence is unfavorable, the incipient Behavior (***R+***) will be inhibited and replaced by the selection of Behavior (***R-***) which is antagonistic to ***R+***.[7]

The operation of this principle has been demonstrated by Adams and Dickinson (1981) who showed that if a rat's lever-pressing response is reinforced by the opportunity to eat a particular foodstuff and that foodstuff is then subject to a taste aversion procedure outside the lever-pressing situation, lever-pressing is suppressed by the contingent presentation of the devalued foodstuff despite the fact that the rat has never experienced the taste aversion procedure as a consequence of lever-pressing.

The Law of Exercise as the Learning Rule in Output Selection

The fact demonstrated by the Adams and Dickinson experiment, that what an animal learns in the process of trial-and-error-correction is to expect a certain consequence rather than perform a particular response, does not mean that habits of responding in a particular way are not ‘stamped in’ by this kind of procedure, while the antagonistic responses are ‘stamped out’, as originally proposed by Thorndike (1911). What it does mean, however, is that the stamping in and out is not to be attributed *directly* to the effect on behavior of its consequences in the past, as proposed by traditional formulations of the Law of Effect. It is simply a consequence of the fact that, by virtue of the operation of other principles, a particular motor output has been repeatedly selected in response to a particular input, while its antagonist has been repeatedly inhibited under those same input

---

[7] It will be observed that this is an output selection principle rather than a learning rule.

conditions. In other words, particular input-output connections are stamped in and stamped out in accordance with

> *The Law of Exercise*
>
> If an output (response) of type ***R+*** is repeatedly selected by and in the presence of an input of type ***S+***, and an antagonistic output ***R-*** is actively inhibited under the same input condition, the tendency for an ***S+*** to select an ***R+*** and inhibit an ***R-*** will be strengthened every time an ***S+*** evokes an ***R+*** and inhibits an ***R-***.

Conclusion

This way of engineering the Law of Effect may not be a feature that needs to be incorporated into a PDP in order to give it the discrimination learning abilities required for the various practical purposes for which such devices are increasingly being used. That these are the kinds of consideration that need to be born in mind in constructing models of how such matters are engineered by the brain is not in doubt.

References

Adams, C. D. and Dickinson, A. (1981) Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, 33 B: 109-112.

Chater, N. (1989) Learning to respond to structure in time. *R.S.R.E. Malvern Technical Report*, September 1989.

Dickinson, A. (1988) Intentionality in animal conditioning. In L. Weiskrantz (Ed.) *Thought without Language*. Oxford: Oxford University Press.

Dennett, D. C. (1978) *Brainstorms: Philosophical Essays on Mind and Psychology*. Montgomery, Vermont: Bradford Books.

Grice, G. R. (1948) The relation of secondary reinforcement to delayed reward in visual discrimination learning. *Journal of Experimental Psychology*, **38**, 1-16.

Guthrie, E. R. (1935) *The Psychology of Learning*. New York: Harper.

Harzem, P. and Miles, T. R. (1978) *Conceptual Issues in Operant Psychology*. Chichester: Wiley.

Hebb, D. O. (1949) *The Organization of Behavior*. New York: Wiley.

Hull, C. L. (1943) *Principles of Behavior*. New York: Appleton-Century-Crofts.

James, W. (1890) *Principles of Psychology*, Vol.1. New York: Holt.

Lashley, K. S. (1930). Basic neural mechanisms in behavior. *Psychological Review*, 37: 1-24.

McClelland, J. L. and Rumelhart, D. E. (1988) *Explorations in Parallel Distributed Processing*. Cambridge, Mass.: MIT Press.

Michael, J. (1982) Distinguishing between discriminative and motivational functions of stimuli. *Journal of the Experimental Analysis of Behavior* 37: 149-155.

Pavlov, I. P. (1927) *Conditioned Reflexes*. English Translation by G. V. Anrep. London: Oxford University Press.

Perin, C. T. (1943) A quantitative investigation of the delay-of-reinforcement gradient. *Journal of Experimental Psychology*, **32**, 37-51.

Place, U. T. (1985) Semicovert behavior and the concept of pain: a comment on H. Rachlin 'Pain and behavior'. *The Behavioral and Brain Sciences*, *8*, 70-71.

Skinner, B. F. (1938) *Behavior of Organisms*. New York: Appleton-Century-Crofts.

Skinner, B. F. (1966) An operant analysis of problem solving. In B.Kleinmuntz (Ed.) *Problem Solving: Research, Method and Theory*. New York: Wiley.

Skinner, B. F. (1969) *Contingencies of Reinforcement*. New York: Appleton-Century-Crofts.

Spence, K. W. (1951) Theoretical interpretations of learning. In S. S. Stevens (Ed.) *Handbook of Experimental Psychology*. New York: Wiley.

Thorndike, E. L. (1898) Animal intelligence: an experimental study of the associative process in animals. *Psychological Monographs*, **2**, No.8.

Thorndike, E. L. (1911) *Animal Intelligence*. New York: Macmillan.

Tolman, E. C. (1932) *Purposive Behavior in Animals and Men*. Berkeley: University of California Press.

Watson, J. B. (1914) *Behavior, an Introduction to Comparative Psychology*. New York: Holt.