JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2007, **88,** 115–130 NUMBER 1 (JULY)

AUTOSHAPING AND AUTOMAINTENANCE: A NEURAL-NETWORK APPROACH

José E. Burgos

UNIVERSITY OF GUADALAJARA - *CEIC*

This article presents an interpretation of autoshaping, and positive and negative automaintenance, based on a neural-network model. The model makes no distinction between operant and respondent learning mechanisms, and takes into account knowledge of hippocampal and dopaminergic systems. Four simulations were run, each one using an *A-B-A* design and four instances of feedfoward architectures. In *A*, networks received a positive contingency between inputs that simulated a conditioned stimulus (CS) and an input that simulated an unconditioned stimulus (US). Responding was simulated as an output activation that was neither elicited by nor required for the US. *B* was an omission-training procedure. Response directedness was defined as sensory feedback from responding, simulated as a dependence of other inputs on responding. In Simulation 1, the phenomena were simulated with a fully connected architecture and maximally intense response feedback. The other simulations used a partially connected architecture without competition between CS and response feedback. In Simulation 2, a maximally intense feedback resulted in substantial autoshaping and automaintenance. In Simulation 3, eliminating response feedback interfered substantially with autoshaping and automaintenance. In Simulation 4, intermediate autoshaping and automaintenance resulted from an intermediate response feedback. Implications for the operant–respondent distinction and the behavior–neuroscience relation are discussed.

Key words: autoshaping, automaintenance, interpretation, neural networks, directedness, response feedback, operant–respondent distinction, behavior–neuroscience relation

This article is an exercise in scientific interpretation, the use of principles derived through experimental analysis of certain phenomena to account for other, typically more complex phenomena. In this sense, it is like Skinner's (1957) *Verbal behavior*, where principles derived through experimental analyses of relatively simple behaviors (key pecking in pigeons and bar pressing in rats) were used to account for more complex ones (speaking and writing in humans). The present interpretation, however, differs in three respects.

First, the interpretation targets two behavioral phenomena that were discovered through experimental analysis of nonhuman behavior, namely autoshaping and automaintenance. They have received much attention largely because of their implications for the operant–respondent distinction, a distinction that has been central in conditioning research since it was drawn by Skinner (1935, 1937, 1938; cf. Konorski & Miller, 1937a, b; Miller & Konorski, 1928). Second, the principles to which the interpretation appeals have been derived through independent experimental analyses of the structure and functioning of nervous systems at the cellular and anatomical levels. In this sense, it is a neural interpretation. Third, the interpretation takes the form of computer simulations that are based on a mathematical formulation of such principles. In this sense, it is a formal interpretation (Donahoe & Palmer, 1994).

The primary aim of this exercise is to show that a neural-network model with the following two interrelated features can simulate autoshaping and automaintenance. 1) It makes no distinction between operant and respondent learning mechanisms, nor does it reduce either category to the other, although it makes a distinction between types of responding and types of contingencies. 2) The learning mechanism was not inferred from observed performance, but informed by independent knowledge from neuroscience. In the first section, I briefly review highlights of the literature on autoshaping and automaintenance and their relevance for the operant–respondent distinction. The second section is devoted to the basic aspects of the model used, with an

This research was partly funded by Grant 42153H from the Mexican National Council for Science and Technology (CONACYT).

I thank William Timberlake and two anonymous reviewers for useful comments to previous drafts.

Correspondence and reprint requests may be addressed to the author at Francisco de Quevedo 180, Col. Arcos de Vallarta, Guadalajara, Jalisco 42230, Mexico. (E-mail: jburgos@cucba.udg.mx; web: www.ceic.cucba.udg.mx).

doi: 10.1901/jeab.2007.75-04

elaboration of the aforementioned features. The simulations are described in the third section. I discuss the implications for the operant–respondent distinction and the behavior–neuroscience relation in the last section.

THE PHENOMENA

Autoshaping (AS) was first reported by Brown and Jenkins (1968). Food-deprived pigeons were given pairings of a keylight followed a few seconds later by access to the food magazine. After a number of pairings, pigeons pecked the illuminated key. The first key peck caused the procedure to switch to a standard operant-conditioning schedule where key pecks gave access to the food magazine. Brown and Jenkins defined AS as the occurrence of the first key peck, but other acquisition criteria can be used (e.g., three out of four consecutive trials with at least one key peck; see Gibbon & Balsam, 1981). AS has been observed in different species and with different types of responses and reinforcers (e.g., Gamzu & Schwam, 1974; Jenkins & Moore, 1973; Sidman & Fletcher, 1968; Smith & Smith, 1971; Squier, 1969; Stiers & Silberberg, 1974; Timberlake & Grant, 1975; Wasserman, 1973).

The implications of AS for the operant–respondent distinction were immediately apparent. AS seems to blur the distinction, for it exemplifies the acquisition of a prototypical operant response through a respondent-conditioning procedure (Hearst, 1975; Pear & Eldridge, 1984; Schwartz & Gamzu, 1977; Tomie, Brooks, & Zito, 1989). Key pecking is a prototypical operant response in three senses.

First, it is directed in that it involves approaching and making physical contact with an identifiable part of the environment. In the autoshaped key peck, that part is the source of the light (viz., the response key). This feature is the reason behind the alternative label ''sign-tracking'' (Hearst & Jenkins, 1974). The defining characteristic of this feature for operant responding has been challenged (e.g., Pear & Legris, 1987). However, it remains a defining feature of autoshaped responding. Second, the response is emitted in that it is not elicited in an obvious way by any stimulus, especially the reinforcer, grain (although grain appears to elicit grain pecking in birds). Third, the response is modifiable by consequences. AS shows that a response with these features can be reliably acquired in a procedure where the response is neither elicited by nor required for the reinforcer.

The question thus arises as to whether autoshaped key pecking is operant, respondent, or both. The evidence strongly suggests that AS is a respondent-conditioning phenomenon. According to this account, the keylight functions as a conditioned stimulus (CS) and the food as an unconditioned stimulus (US). AS has been shown not to occur (or to occur very weakly) in zero- or negative-contingencies between CS and US (e.g., Gamzu & Williams, 1973), and to depend on temporal variables such as the trial-duration and the intertrial interval, in a way comparable to respondent conditioning of autonomic responses (e.g., Gibbon & Balsam, 1981; Jenkins, Barnes, & Barrera, 1981; Ricci, 1973; Terrace, Gibbon, Farrell, & Baldock, 1975). Also, the topography of the autoshaped response closely resembles the topography of the response specifically elicited by the US (e.g., Jenkins & Moore, 1973).

Some authors (e.g., Gormezano & Kehoe, 1975) have argued that the unconditioned response (UR) is not directly measured in AS, which makes the labels ''US'' and ''UR'' misnomers. They thus have rejected AS as a clear case of respondent conditioning. Others (e.g., Herrnstein & Loveland, 1972; Hursh, Navarick, & Fantino, 1974) have reached a similar conclusion, but based on the directed character of autoshaped responses. According to the basic argument, AS is not a clear case of respondent conditioning because the former involves directed responses whereas the latter involves autonomic responses, which are nondirected in character. The problem with this argument is that it neglects the fact that directed responses also have been observed in classic respondent conditioning of autonomic responses. For instance, Pavlov (1941) observed that when the CS was a light and its source (a bulb) was within reach of the subjects (dogs), they licked the source (see also Pavlov, 1955; Zener, 1937).

Further doubts on the respondent character of AS have appealed to the fact that, pre-

sumably in contrast to respondent conditioning of autonomic responses, AS seems to be relatively insensitive to intermittent reinforcement (e.g., Gamzu & Williams, 1971, 1973; Gonzalez, 1973, 1974; Schwartz & Williams, 1972a). However, it is unclear that intermittent reinforcement is harmful to all respondent conditioning of autonomic responses (Gibbs, Latham, & Gormezano, 1978; Gormezano & Moore, 1969).

The uncertainty increased with the discovery of automaintenance. Two sorts of automaintenance have been identified. Positive automaintenance (PAM) is the maintenance of an autoshaped response under the same conditions in which it was autoshaped (e.g., Gonzalez, 1974). In contrast to Brown and Jenkins' (1968) procedure, the first response does not produce the reinforcer. Instead, light–food pairings keep occurring, independently of the organism's behavior. PAM suggests that stimulus–reinforcer relations also are sufficient to maintain directed responding. However, responses in PAM are accidentally followed by the reinforcer, which introduces a positive response–reinforcer contingency that may also play a role.

The role of response–reinforcer relations in autoshaped key pecking has been assessed through negative automaintenance (NAM), the reliable maintenance of an autoshaped response in an omission-training procedure (Williams & Williams, 1969). In its simplest form, the procedure gives the reinforcer at the end of each trial unless the animal responds. If the animal responds, the trial ends without reinforcement. Despite this negative response–reinforcer contingency, responding was maintained, albeit at a reduced level (cf. Hursh et al., 1974; Sanabria, Sitomer, & Killeen, 2006). Autoshaped key pecking thus seems to be sensitive to response–reinforcer relations. However, stimulus–reinforcer relations are still considered to play a major role (see also Myerson, Myerson, & Parker, 1979; Schwartz & Williams, 1972a, b; Wessells, 1974).

BASICS OF THE MODEL

The neuroscientific rationale and mathematical details of the model have been discussed in other papers (Burgos, 2001, 2003, 2005; Burgos & Donahoe, 2000; Burgos & Murillo, in press; Donahoe, 2002; Donahoe & Burgos, 1999, 2000; Donahoe, Burgos, & Palmer, 1993; Donahoe & Palmer, 1994; Donahoe, Palmer, & Burgos, 1997a, b). Accordingly, I shall focus on those aspects that are most directly relevant to the present study. In particular, I shall clarify the sense in which the model makes no distinction between operant and respondent learning mechanisms.

A neural network is a set of interconnected neural processing elements or units. A unit is a sort of abstract neuron that can be activated through a mechanism that is described by an activation function. In the present model, activations are real values between 0 and 1, and can be neurally interpreted as the probability of occurrence of an action potential. A connection is the theoretical analogue of a synapse or small group of synapses, and it consists of a presynaptic unit and a postsynaptic unit. The strength of a connection is represented numerically by a weight that represents the efficacy with which the presynaptic unit can activate the postsynaptic unit(s) to which it is connected. In the present model, a weight is a real number between 0 and 1. It can be neurally interpreted as the proportion of postsynaptic receptors that are controlled by a presynaptic process. In both functions, time is divided into discrete time steps (ts) of an undefined duration.

Weights change through a learning function that includes two free parameters that determine rate of change (one for increments, one for decrements), the activation of the presynaptic unit, the activation of the postsynaptic unit, and a signal that arises from certain specialized units in the network (*ca1* and *vta*; see below). The function also includes the amount of weight available on the postsynaptic unit. This amount is proportional to the amount of weight already gained by other presynaptic units that are connected to the same postsynaptic unit. The more weight a presynaptic unit gains, the less weight will be available to be gained by other presynaptic units that are connected to the same postsynaptic unit. The present learning function, then, is competitive in character. Presynaptic units connected to a certain postsynaptic unit compete for a limited amount of weight available on the postsynaptic unit.

Models of conditioning typically have been developed following a top-down strategy (e.g.,

Gibbon & Balsam, 1981; Miller & Matzel, 1988; Rescorla & Wagner, 1972; Wagner, 1981). This strategy begins with processes that are inferred from observed performance, independent of the biological structure and functioning of the brain. The processes are postulated to involve some sort of relationship (competition or comparison) among hypothetical entities (associative strength, expectancy) that are used to explain known phenomena and predict novel ones. The correspondence between the processes and the biological structure of the brain is not integral to the construction of the model. That is to say, the models are not informed by neuroscientific knowledge about the biological structure and functioning of brains. In some cases (e.g., Grossberg, 1968; Meeter, Myers, & Gluck, 2005; Schultz, 2002; Wagner & Donegan, 1989), a correspondence is sought after the model has been built. The logic of such models, however, remains top-down.

In contrast, the present model was developed following a bottom-up strategy in that it involved drawing on independent knowledge from neuroscience about vertebrate brains at the cellular, microcircuit, and anatomical levels. Hence, some correspondence with the biological structure and functioning of the brain was integral to the construction of the model. Conditioning phenomena are thus explained as emergent properties of artificial systems whose parts function and relate in ways that are consistent with that knowledge. The learning function, in particular, is informed by knowledge about hippocampal and dopaminergic systems, which have been shown to be involved in conditioning (e.g., Berger, Alger, & Thompson, 1976; Christian & Thompson, 2003; Pan, Schmidt, Wickens, & Hyland, 2005; Power, Thompson, Moyer, & Disterhoft, 1997; Schultz, Dayan, & Montague, 1997). This feature distinguishes the model from others that do not take such knowledge into account (e.g., Kehoe, 1988; Sutton & Barto, 1981), or that do but focus only on one system (e.g., Meeter, Myers, & Gluck, 2005; Schmajuk & DiCarlo, 1992; Schultz, 2002; Zipser, 1986).

The model makes no distinction between operant and respondent learning mechanisms in that nothing in its learning function corresponds to any of the standard ways of making the distinction. The two descriptive ways date to Skinner (1935). One way refers to the distinction between elicited and emitted responses. The other refers to the distinction between stimulus–stimulus and response–outcome contingencies. Nothing in the learning function corresponds to either distinction. As I argue below, the model makes these distinctions, but not at the mechanism level of the learning function. Rather, they are made at the system or network level.

There also is a theoretical distinction that refers to different types of associations between internal representations of events. One type is postulated to be between stimulus representations (e.g., Gibbon & Balsam, 1981; Miller & Matzel, 1988; Rescorla & Wagner, 1972; Wagner, 1981), the other between response and outcome representations (e.g., Colwill & Rescorla, 1986, 1990; Rescorla, 1991). The former are supposed to result from Pavlovian contingencies, the latter from operant contingencies. Nothing in the learning function corresponds to such a distinction either. Nor does anything in a network correspond to the distinction either. To be sure, a connection in a network is a sort of associative bond. Also, one could speak of stimulus and response representations in a network. However, no single connection in a network in this model corresponds to an elementary stimulus–stimulus or response–outcome bond. In order to appreciate this and how the model makes the descriptive distinctions, consider Figure 1, which shows a typical network in this model.

Units are represented as circles and connections as arrows. Thin arrows represent connections that are modifiable through the learning function. Thick arrows represent what are initially maximally strong, unmodifiable connections. The units are organized into layers in a feedforward manner, where the activations propagate in only one direction, from one layer to the next (left to right in the diagram). Here, the model assumes no logical equivalence between a network and its constituting parts (units and connections), just as there is no logical equivalence between a whole organism and its neurons and synapses. Hence, what applies to a whole system (whether an organism or an artificial neural network) does not

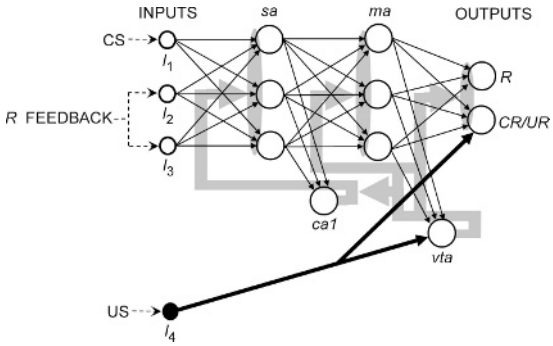

Fig. 1. A typical network architecture. Units labeled as I_1, I_2, and I_3 are input units whose activations represent the kinds of exteroceptive stimuli used in respondent conditioning as CSs (e.g., light and tones). Reinforcement (US occurrence) is represented by an activation of the input unit labeled as I_4. The dashed arrows labeled as ''CS'', ''*R* feedback'', and ''US'' represent the input activations that defined the stimuli used in Simulation 1. Thin black arrows represent variable connections whose weights changed according to a learning function. Thick black arrows represent maximally strong unmodifiable connections. Gray arrows and areas represent the signals that influence weight changes. Responding is represented by the activation of the output units, labeled as *R* and *CR/UR*. The other labels denote: *sa*: sensory-association; *ma*: motor-association; *ca1*: Cornu Ammon 1; *vta*: ventral tegmental area.

necessarily apply to its parts. The descriptive operant–respondent distinctions were originally made in reference to whole organisms, not their neurons and synapses. The model seeks to be consistent with this usage by making the distinctions at the level of the whole network, not its units and connections. In this manner, the model avoids the so-called ''mereological fallacy,'' the assumption that what applies to the whole necessarily applies to its parts (see Bechtel, 2005; Bennett & Hacker, 2003).

The input (leftmost) layer consists of units labeled as I_1 through I_4, whose activations represent environmental stimuli. Activations of I_1, I_2, and/or I_3 represent sensory–exteroceptive stimuli like lights and sounds, typically used as CSs in respondent conditioning and as discriminative stimuli in operant conditioning (see Simulations for an explanation of the label ''*R* feedback''). Activations of I_4 represent the occurrence of a US or primary reinforcer, like food, water, or shock. Strictly speaking, input activations are not computed according to the activation function. Rather, they are assigned according to a training protocol that simulates the kinds of procedures used in conditioning studies.

Units I_1, I_2, and I_3 connect to the sensory-association (*sa*) units, which simulate the kind of integration found in secondary sensory areas. These units, in turn, connect to *ca1*, a hippocampal-like change detector that is the source of a signal that influences changes in all the weights of the input–*sa* and *sa*–*ca1* connections. In this sense, the signal is diffuse. Its diffuse character is represented in the diagram by the grey arrows and areas arising from the *ca1* unit. The signal is defined as a temporal difference between the activations of *ca1* in successive time steps (ts), for which it also is a discrepancy signal. The *sa* units also are connected to the motor-association (*ma*) units, which represent the kinds of neurons found in secondary motor areas.

The *ma* units connect to *vta*, a dopaminergic-like unit that is the source of a signal that influences changes in the *sa*–*ma*, *ma*–*vta*, and *ma*–output weights. Like the signal arising from *ca1*, the *vta* signal is a diffuse discrepancy. Its diffuse character is represented in the diagram by the grey arrows and areas arising from the *vta* unit. At any ts *t*, *vta* can be activated in one of two ways. If the activation of I_4 at *t* is above zero, *vta* is activated at *t* with the same level as I_4. Otherwise, *vta* is activated by I_1, I_2, and/or I_3 via the *sa* and *ma* units, according to the activation function. Substantial activations of *vta* by the *ma* units thus require learning in that they occur only after the appropriate connection weights have increased. Typically, this increase obtains in a training protocol that simulates the kinds of procedures that have been observed to promote excitatory conditioning in natural systems.

The *ma* units also connect to the output units, labelled as *R* and *CR/UR* (rightmost layer). The activations of these units represent the system's overt responding. Like *vta*, *CR/UR* also receives an initially maximally strong and unchangeable connection from I_4. Like *vta*, *CR/UR* can be activated in one of two ways at any ts *t*. If the activation of I_4 is above zero at *t*, *CR/UR* is activated at *t* with the same level as I_4. This activation represents elicited or unconditioned responding. Otherwise, *CR/UR* is activated by I_1, I_2, and/or I_3, via the *sa* and *ma* units, according to the activation function.

This activation represents conditioned responding; hence the label ''*CR/UR*''. Like *vta*, substantial conditioned activations of *CR/UR* require learning in that they occur only after the appropriate weights have increased. Such an increase, too, results from exposure to a training protocol that simulates the kinds of procedures that lead to excitatory conditioning in natural systems.

The *R* unit, in contrast, can only be activated conditionally. Such activation represents emitted responding in that it is not elicited by the US (i.e., not activated by I_4) and requires learning to occur at substantial levels. In this model, then, the distinction between emitted and elicited responding is made in terms of the distinction between *R* activations and unconditional *CR/UR* activations, respectively. Again, this distinction is made at the network level, between types of output activations, not between learning mechanisms or aspects of the learning mechanism. Hence, the distinction is nowhere to be found in the learning function.

The model makes the distinction between respondent and operant contingencies in terms of two different conditions of activation of I_4, which define two types of training protocols. In respondent protocols, I_4 activations depend only on the activations of the other input units, regardless of the activation of *R*. In operant or response–outcome protocols, I_4 activations depend also on *R* activations. In this kind of protocol, I_4 activations may additionally depend on the activations of the other input units, in order to simulate discriminated or occasion-setting procedures. This distinction, too, is made at the network level, not at the learning-mechanism level. Therefore, the distinction is not part of the learning function either.

SIMULATIONS

The primary main aim of this exercise is to include AS, PAM, and NAM in the range of behavioral phenomena that a model with the above features can simulate. Simulation 1 is the basic demonstration of how the model simulates the phenomena. The rest of the simulations explore two more specific aspects. One aspect is the role of directedness as a defining feature of autoshaped responding. The other is how this role is affected by the network architecture. The exercise is only a starting point, so the simulations are very simple. In each simulation, four neural networks and a within-subject design that approximates the *A–B–A* research design were used. More complex simulations are possible, but they are not the best starting points and thus better left for further studies. In order to have some continuity with previous studies with this model, and rely as little as possible on parametric manipulation, all of the model's free parameters were the same as those used in previous simulations. Also, in agreement with most simulation studies with this model, initial connection weights were set to 0.01 (cf. Burgos, 2003, 2005).

Simulation 1

Four instances of the architecture depicted in Figure 1 were used. This architecture is the one most used in previous research with this model (but see the other simulations). All instances were first given 300 trials of a CS–US contingency, where the CS was defined as the activation of I_1 at a level of 1.0 for 8 ts (dashed arrow labeled as ''CS'' in Figure 1). The US was defined as the activation of the I_4 unit also at a level of 1.0 at the last CS ts (dashed arrow labeled as ''US'' in Figure 1). The CS–US interval, then, was 7 ts. As in most previous simulations with this model, for convenience the intertrial interval was not explicitly simulated (cf. Burgos, 2005). Instead, it was assumed to be long enough for all activations to decrease to a near-zero level.

The networks were assumed to dwell in an extremely simple environment, consisting of only one identifiable part with which they could make physical contact and function like an operandum. Hence, there could only be one directed response. No other identifiable part of a standard operant-conditioning chamber, like a food magazine, was simulated. Rather, the US was supposed to be directly administered to, not actively consumed by, the networks. In order to simulate active consumption of the primary reinforcer from a food magazine, a more complex simulation would be required. But as a starting point, I preferred to keep the study as simple as possible.

Response directedness was defined as follows. Whenever the activation of *R* was equal to or greater than 0.5, input units I_2 and I_3 were activated at a level of 1.0 (dashed arrows

labeled as "*R* feedback" in Figure 1). Otherwise, I_2 and I_3 were not activated. *R* activations below 0.5 thus represented undirected responding such as orientation toward the key (e.g., Wessells, 1974). This criterion implies that the same neurons are involved in directed and undirected responding, which seems implausible. However, the criterion is only a simplification, not a theoretical assertion about the functioning of natural nervous systems.

The above definition of response directedness was guided by two assumptions (e.g., Dinsmoor, 1985, 1995, 2001; Notterman & Mintz, 1965; Wyckoff, 1952; Zeigler & Wyckoff, 1961). First, approaching and making direct physical contact with some part of the environment generates sensory feedback. Second, this feedback can come to control behavior. In key pecking, for instance, the feedback would be given at least by the proximate sight of the lighted key (which makes it look larger and brighter), the sound made by the beak making contact with the key, and the proprioceptive stimulation that arises from the resistance offered by the key upon pecking it. Sensory feedback from directed responses thus seems considerably richer than the putative CS. This richness was represented by the activation of two input units (I_2 and I_3), as opposed to just the one for the CS (I_1).

Following the CS–US trials, the networks were then given 300 trials of omission training, defined as follows. If *R* was activated at a level of 0.5 or greater at ts 8, I_4 was not activated. Otherwise, I_4 was activated at a level of 1.0. After this training, the networks reverted to the first condition. In the three phases, *R* activations above 0.5 that occurred before ts 8 did not deactivate I_1. This procedure was intended to simulate one in which key pecking does not turn off the keylight (see Schwartz & Williams, 1972a).

Figure 2 shows the results for all networks (rows labeled as N1 through N4) and phases (columns labeled as AS/PAM, NAM, and PAM). The dashed lines mark the *R* activation criterion for response feedback. *R* activations were lower during omission training (NAM) than during the other two phases. This effect is explained as follows. Early in training, during AS, reinforcements (I_4 activations) were correlated only with I_1 activations, since *R* activations were below 0.5. This arrangement caused

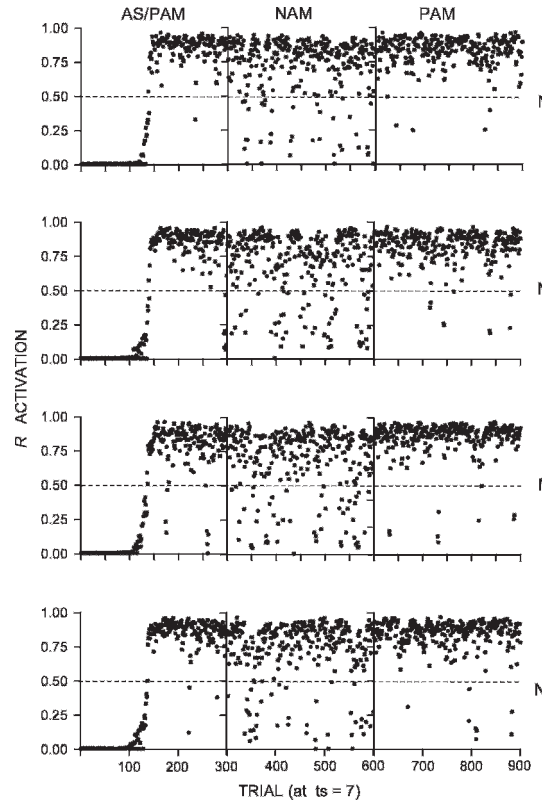

Fig. 2. Results of Simulation 1. The rows represent the individual networks (labeled as N1, N2, N3, and N4). Each network was given a sequence of three phases, represented by the columns labeled as AS/PAM (autoshaping/positive automaintenance), NAM (negative automaintenance), and PAM. The dots represent *R* activations at ts = 7 (the moment before reinforcement). The dashed lines mark the *R* activation criterion for sensory feedback from responding.

all connections to gain weight, except for the I_2–*sa* and I_3–*sa* connections. These connections remained unchanged at their initial values because of the inactivation of I_2 and I_3. After AS, during the first phase of PAM, weights were sufficiently high to allow I_1 to activate *R* above 0.5, which caused I_2 and I_3 to be activated and correlated with I_4 activations, which resulted in substantial maintenance. During NAM, *R* activations equal to or greater than 0.5 at ts 8 prevented the activation of I_4, inducing a decrease in all connection weights and, consequently, the activation of *R* by I_1, I_2, and I_3. Whenever the activation of *R* was below 0.5, I_4 was reactivated and all connections regained weights, which increased the activation of *R*, and so on. The net effect was the

maintenance of R's activation at substantial levels.

These results show that the same network architecture used in previous simulation research with this model can simulate the phenomena of interest, at least in their simplest form. The results are consistent with the evidence that directed responding that is not elicited by the reinforcer can be acquired and maintained under a respondent contingency, without an explicit positive operant contingency. This is in agreement with the dominant account, according to which AS is a purely respondent-conditioning phenomenon. Additionally, the reduced activations in NAM were due at least partly to the negative stimulus–reinforcer relation, which induced a reduction in the activation of R by I_1. This result agrees with the account that stimulus–reinforcer relations play a role in NAM.

The evidence also suggests that response–reinforcer relations may play a role in PAM and NAM. The present results are consistent with this evidence as well. Not only were the networks sensitive to the explicit negative response–reinforcer contingency in NAM. Also, in PAM high R activations at ts 7 were adventitiously followed by I_4 activations at ts 8, which may have contributed to the substantial PAM. Neither the present nor the subsequent simulations, however, allow for an unequivocal determination of this contribution. Such a determination would require a more complicated simulation that is better left for future study.

The connectivity of the architecture used here was complete in that all units in a layer were connected to all units in the next layer. However, aside from its neuroanatomical implausibility (there is no such complete connectivity in vertebrate brains), another simulation showed that directedness made no difference in this architecture. Four other networks whose R activations had no sensory feedback (i.e., did not activate I_2 and I_3) behaved nearly identically to those shown in Figure 2, under the same conditions. Activations of I_1 (CS) correlated with activations of I_4 (US) were sufficient for simulating acquisition and maintenance of undirected responding that was not elicited by the reinforcer. The model's relevance to AS, PAM, and NAM is thus diminished. Fortunately, the model allows for a way to make directedness more relevant. In order to understand how the model can do this, it will be convenient to see exactly why directedness is unnecessary in the architecture shown in Figure 1.

The problem with the architecture in question is its full connectivity. Such connectivity favored a strong competition between the CS (activation of I_1) and response feedback (I_2 and I_3 activations correlated with R activations larger than 0.5). Early in training, R was activated below 0.5, for which only the CS occurred, with no response feedback. The CS thus had a competitive advantage. As a result, I_1 gained weight before I_2 and I_3 could gain any weight. By the time response feedback started to occur, I_1 had gained much of the weight available on the *sa* units. As a result, I_2 and I_3 gained little weight, so their activations did not contribute substantially to the activation of R. The implication is that directedness could play a more substantial role in networks with a reduced competition between the CS and response feedback. The next two simulations were devised to explore this implication.

Simulation 2

Infinitely many different network architectures are possible. However, a neuroscientifically meaningful choice is underdetermined by current experimental evidence. In particular, the *sa–ca1*, *ma–vta*, and *ma*–primary motor cortex connectivity is not known with the specificity that is required for handcrafting an artificial neural architecture. Consequently, at this point in time, the choice has to be guided more by logical than evidential considerations. The present choice was constrained by two logical considerations. First, the architecture must allow for a determination of whether sensory feedback from directed responses plays any role when competition between the CS and the feedback is reduced. Second, a sensible first step for this determination is to explore the extreme case where there is no competition.

Figure 3 shows an architecture that satisfies both considerations (for convenience, the US and *CR/UR* units, as well as the diffuse signals, were omitted from the diagram). It is a simple architecture with minimal synaptic competition between different groups of input units. It has two distinguishable portions joined by a single input unit (I_3). The output unit (R) is part of the upper portion, which allows the unit to be

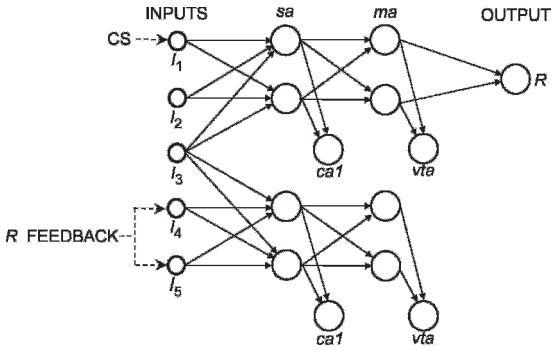

Fig. 3. Network architecture used in Simulations 2, 3, and 4. Labels I_1 through I_5 represent the exteroceptive sensory input units. The dashed arrows labeled as "CS" and "*R* feedback" represent the activations that defined the stimuli used in the simulations. The US and *CR/UR* units, as well as diffuse signals (shown in Figure 1), were omitted for simplicity. Signals influencing weight changes were averaged across *ca1* or *vta* units.

activated only by I_1 and/or I_2. Similarly, the lower portion could be activated only by I_4 and/or I_5.

Each portion has its own *ca1* and *vta* units, because having only one *ca1* and one *vta* for both portions would have increased competition. This feature raises the technical issue of how to compute the diffuse signals. The model is neutral in this respect and many solutions are possible, but, once again, the choice is underdetermined by the experimental evidence. Logically, a sensible strategy would be to preclude any learning bias towards either portion of the architecture. The simplest way to achieve this is to assume that a signal of either type (*ca*1 or *vta*) is distributed evenly between the two units of that type. The *ca1* signal thus was defined as the sum of the diffuse signals from both *ca1* units, divided by the two *ca1* units. The same was done for the *vta* signal and units.

In order to prevent competition between CS and feedback, the CS was defined as the activation of I_1 at a level of 1.0 for 8 ts (dashed arrow labeled as "CS" in Figure 3). The sensory feedback was defined as the activation of I_4 and I_5 at a level of 1.0 whenever *R* was activated at 0.5 or more (dashed arrows labeled as "*R* feedback" in Figure 3). I_2 and I_3 thus remained inactive throughout the entire simulation. Unlike the sensory feedback in Simulation 1 and the CS in both simulations, sensory feedback in this simulation did not activate *R*. This feature allowed for the determination of whether such activation was necessary for sensory feedback

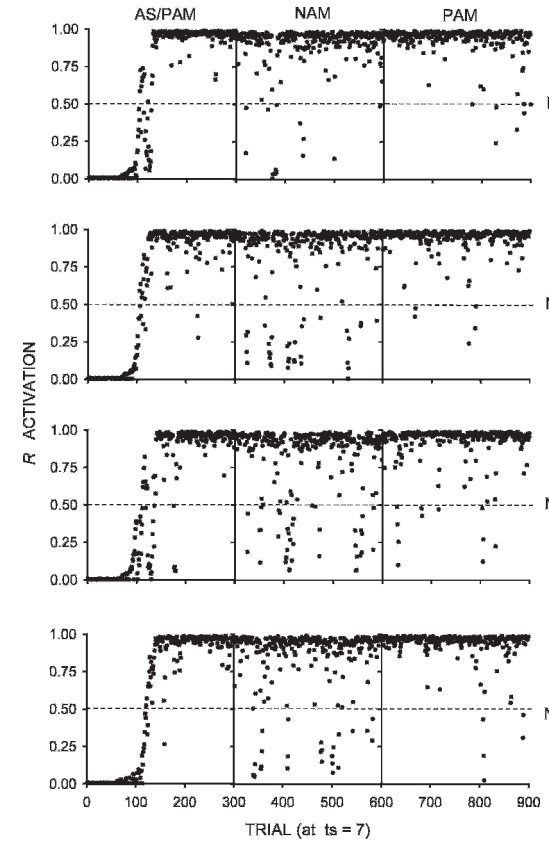

Fig. 4. Results of Simulation 2. The rows represent the individual networks (labeled as N5, N6, N7, and N8). Each network was given a sequence of three phases, represented by the columns labeled as AS/PAM (autoshaping/positive automaintenance), NAM (negative automaintenance), and PAM. The dots represent *R* activations at ts = 7 (the moment before reinforcement). The dashed lines mark the *R* activation criterion for sensory feedback from responding.

from *R* to play a role (see Simulation 3). The activations still qualify as feedback in that they can activate the *ca1* and *vta* units of the lower portion of the network, which influences changes in the weights of the *ma–R* connections and, to this extent, the *R* activations themselves. The US was defined as before. For simplicity, only one *R* unit was used in order to avoid the issue of whether one or two types of responses were involved.

The results are shown in Figure 4. *R* activations were higher and less variable than those observed in Simulation 1. The reason was that I_4 and I_5 could acquire substantial control over the lower portion of the architecture. The *ca1*

and *vta* units in this portion were thus substantially activated, which increased the average *ca1* and *vta* diffuse signals and, hence, the amount of weight increment. As before, omission training was detrimental, although less so than in Simulation 1. This result confirms the ones obtained in the previous simulations, with the concomitant reduction in variability due to an interaction between the architecture and the procedure used.

Simulation 3

To assess the role of sensory feedback in the new architecture, four other instances were given the same procedure as in the previous simulations, but without any sensory feedback correlated with *R* activations. Units I_2 through I_5 of Figure 3 thus remained inactive throughout the entire simulation. Hence, the dashed arrows labeled as ''*R* feedback'' in Figure 3 do not apply here. As Figure 5 shows, the absence of sensory feedback from responding was noticeably detrimental to *R* activations in all three phases, compared to Simulation 2. *R* activations were considerably lower and more variable, and omission training was far more effective. The results can be explained as follows. The *ca1* and *vta* units in the lower portion of the networks remained inactive because of the inactivation of I_3, I_4, and I_5. This situation caused a substantial reduction in the average *ca1* and *vta* diffuse signals and, hence, in the amount of weight increment.

Directedness thus played a more substantial role in a partially connected architecture under a procedure where there was no competition between CS and response feedback. The main implication of this for natural neural systems is that the importance of sensory feedback from responding and, to this extent, directedness, is inversely proportional to the amount of competition between CS and response feedback, as determined by an interaction between the architecture and the environmental conditions. Neural connectivity can vary across different species, individuals of the same species, and sensory and response systems of the same individual. Therefore, differences in the importance of directedness could be partly due to differences in species, individuals, and/or sensory and response systems, environmental conditions being equal.

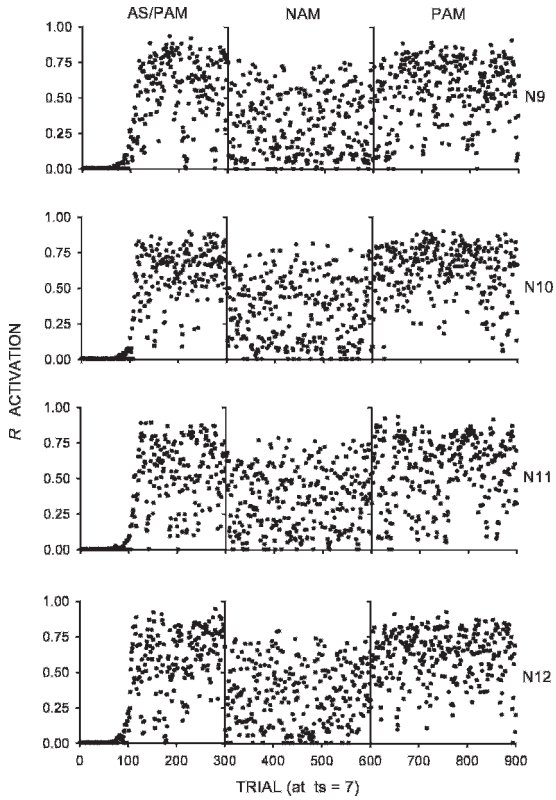

Fig. 5. Results of Simulation 3. The rows represent the individual networks (labeled as N9, N10, N11, and N12). Each network was given a sequence of three phases, represented by the columns labeled as AS/PAM (autoshaping/positive automaintenance), NAM (negative automaintenance), and PAM. The dots represent *R* activations at ts = 7 (the moment before reinforcement). In this simulation, *R* activations had no sensory feedback. Hence, dashed lines are not shown, and the dashed arrows labeled as ''*R* feedback'' in Figure 3 do not apply here.

Sensory feedback played a role despite the fact that it did not activate *R*. The implication of this for natural systems is that response feedback can control the response that produces it without activating the corresponding response system. This possibility allows for a distinction between control and activation by some stimulation. In the present model, control may (as in Simulation 1) but need not (as in Simulations 2 and 3) involve activation by the controlling stimulus.

Simulations 2 and 3 also suggest that AS, PAM, and NAM are proportional to the intensity of sensory feedback. In these networks, this intensity was given by the level of activation of the feedback input units (I_4 and

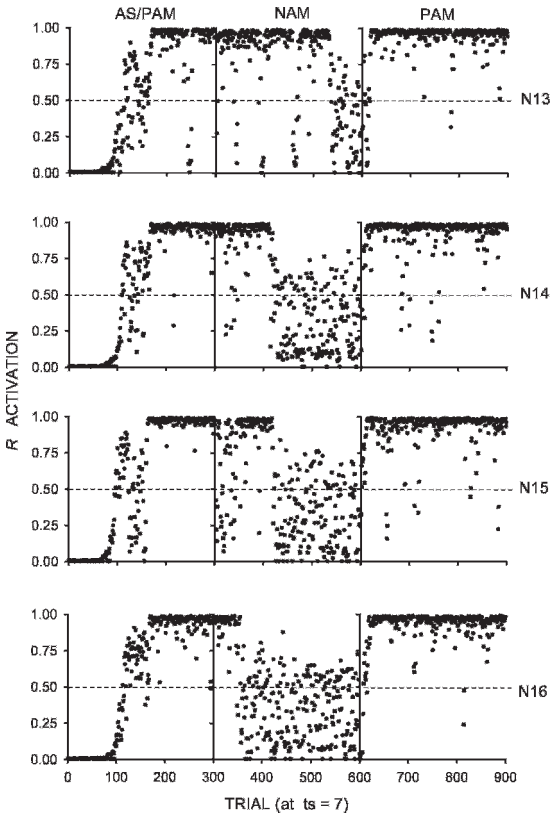

Fig. 6. Results of Simulation 4. The rows represent the individual networks (labeled as N13, N14, N15, and N16). Each network was given a sequence of three phases, represented by the columns labeled as AS/PAM (autoshaping/positive automaintenance), NAM (negative automaintenance), and PAM. The dots represent *R* activations at ts = 7 (the moment before reinforcement). The dashed lines mark the *R* activation criterion for sensory feedback from responding.

I_5 in Figure 3). Intermediate feedback activations should thus produce less substantial AS, PAM, and NAM than that observed in Simulation 2, but more than the present simulation. The following simulation was devised to explore this implication.

Simulation 4

This simulation was identical to the previous two except that the sensory feedback units (I_4 and I_5 in Figure 3) were now activated at a level of 0.6. As Figure 6 shows, there was less substantial AS, PAM, and NAM than Simulation 2, but more than Simulation 3. The explanation is the same as before. This time, the *ca1* and *vta* units in the lower portion of the networks were less substantially activated by I_4 and I_5 than those in Simulation 2, but more so than Simulation 3. This situation caused roughly intermediate average *ca1* and *vta* diffuse signals and, hence, weight increments.

These results allow for a plausible resolution of the following dilemma. On the one hand, Simulation 2 involved substantial sensory feedback from responding. However, it also showed substantial NAM, which is inconsistent with the results reported by Sanabria et al. (2006). On the other hand, Simulation 3 showed less NAM, which is consistent with those results. However, it involved no sensory feedback from responding. The implication is that directedness, conceived as sensory feedback from responding, plays little or no role. In the present simulation, however, an intermediate level of sensory feedback from responding resulted in substantial AS and PAM, but an extinction-like effect in NAM. This result represents a better approximation to those reported by Sanabria et al.

Directedness can thus be conceived as sensory feedback from responding in a way that retains its relevance for AS, PAM, and NAM, and is reasonably consistent with the evidence reported by Sanabria et al. (2006). The implication, again, is that the magnitude with which these phenomena occur is proportional to the intensity of the feedback. Of course, this implication was explored here in a very simple way. Consequently, it is not directly applicable to the much more complex experiment by Sanabria at al. One important simplification in the simulation was that the intensity of response feedback was constant across trials, which is clearly implausible. Notterman (1959) showed that extended exposure to an explicit response–reinforcer contingency (continuous reinforcement) tended to decrease response force. On this basis, it seems reasonable to believe that the intensity of response feedback also decreases with the number of sessions. The most effective procedure used by Sanabria et al. for reducing NAM precisely involved extended PAM followed by NAM followed by instrumental learning. Perhaps the larger number of sessions in this procedure, compared to the other procedures used by Sanabria et al., contributed to a reduction in key pecking force and duration and, to this extent, in sensory feedback from key pecking. This situation, in

turn, may have contributed to the more successful NAM reported by the authors.

GENERAL DISCUSSION

The aim of this study was to show that AS, PAM, and NAM can be simulated with a neural-network model that has two features. First, its learning function makes no distinction between operant and respondent mechanisms, although it makes the descriptive distinctions between response and contingency types at the network level. Second, the model was built in a bottom-up fashion, taking into account independent evidence from neuroscience. The learning function, in particular, takes into account the structure and functioning of hippocampal and dopaminergic systems in conditioning.

In contrast to other theoretical accounts of the distinction (e.g., Bindra, 1972; Mowrer, 1947; Rescorla & Solomon, 1967; Trapold & Overmier, 1972), the present one does not invoke hypothetical inferences from behavior in a top-down fashion. As such, the model represents an advance towards a truly unified theoretical account of conditioning that really takes into account knowledge from neuroscience as integral to model building. The model still involves many simplifying assumptions, but any model does. Besides, the assumptions were dictated primarily by practical considerations of empirical underdetermination by the relevant neuroscientific evidence, not by a top-down hypothesizing strategy.

Hippocampal and dopaminergic systems are found in only some animal species, so the proposed mechanism has a limited scope of application. However, the mechanism is not intended to be universally applicable to all animal species. That would be premature and overly sweeping at this point in time. The mechanism represents only a small step toward a comprehensive account of the neural substrates of conditioning. The relation between the behavior of organisms and the structure and functioning of their nervous systems is exceedingly complex and thus requires progressive, piecewise study.

The preliminary character of the present effort is most obviously seen in the extreme simplicity of the architectures and procedures used. The networks were assumed to be in an environment that had only one identifiable part with which they could make direct physical contact and produce sensory feedback. Further research would be needed to simulate more complex scenarios where the environment has several parts, and the networks have the minimal output capacity to make direct physical contact with every part and the minimal input capacity to detect the sensory feedback that is produced by this contact.

In such scenarios, the present account assumes that different parts produce different sensory feedbacks that differentially control their respective producing responses. For instance, the sensory feedback that is produced by pecking an intermittently illuminated key must be somewhat different from that produced by pecking grain in an intermittently illuminated food magazine. The two feedbacks differentially control different aspects of the pigeon's behavior. In the model, this difference could be simulated by having activations of different *R* units produce activations of different input units. In this manner, the networks could discriminate between responding directed towards one part and responding directed towards the other part. All of this can be accomplished without having the feedback inputs activate the *R* units that produce the feedback.

There also is the issue of the role of response–reinforcer relations in these networks. Response–reinforcer relations (implicitly positive in the case of PAM, explicitly negative in the case of NAM) did occur in the present simulations, so they may have played some role. However, the results were explained purely in terms of the dynamics of activations and weights, as determined by stimulus–reinforcer relations, without appealing to response–reinforcer relations. The latter thus add nothing to an understanding of such dynamics, above and beyond stimulus–reinforcer relations. It remains to be seen whether this obtains in future, more complicated simulations.

Finally, there is the relationship between behavior and neuroscience, which was explored in a recent volume of this journal (Green, 2005). Behavior analysts tend to be suspicious of neural interpretations of behavioral phenomena, under the assumption that they are inherently reductionistic. This assumption, however, is unwarranted. All scien-

tific accounts are neutral with respect to the reductionism–antireductionism debate, simply because this debate is philosophical. As such, the debate involves additional concepts, assumptions, and methodologies that transcend any scientific proposal. The very same scientific account, then, can be philosophically treated in a reductionistic or a nonreductionistic manner. To wit, the model could be taken either as a causal explanation of AS, PAM, and NAM (reductionistic treatment), or as a description of the basic happenings in an organism's nervous system when these phenomena occur (nonreductionistic treatment). There is nothing inherent to the model that favors one or the other treatment.

Another eyebrow raiser for behavior analysts is the potential character of artificial neural networks as ''conceptual'' nervous systems (CNS). Skinner's (1938, Ch. 12; 1950, p. 193; 1974, pp. 217–218) misgivings in this regard are well known, and make some sense. His point that reference to the nervous system is not necessary for the study of behavior in its own right to be scientifically respectable is most reasonable. Equally sound is his point that a scientific study of behavior in its own right is essential for a proper scientific understanding of its neural substrates. The present study is entirely consistent with these two points. Nothing in it questions the scientific respectability of the experimental analysis of behavior in its own right, or its relevance for a proper scientific understanding of the neural substrates of behavior. On the contrary, the study was motivated by phenomena that were discovered through such analysis.

Less reasonable, however, is Skinner's suggestion that a theory is scientifically respectable only if it is supported by direct experimental data. If he indeed suggested this, the present proposal departs from it. Under Skinner's criterion, the theories of Sherrington, Newton, Darwin, and Mendel, for instance, would not have been scientifically respectable at the moment they were proposed since they were not supported by any direct experimental evidence at that moment. By Skinner's standards, then, there was no scientific reason to pursue those theories at that moment, even if later on (as he acknowledged) they received direct experimental support. And yet, scientists did pursue them at that moment, and to good effect. The key is maintaining a balance between the openness of concepts and their potential causal concreteness.

For reasons that remain unclear, scientists often bet on a new theory, as it were, and wait to see if it receives direct experimental support. So, either scientists should not do this or Skinner's criterion is too restrictive. The present proposal leans toward the latter option. To be sure, the wait is not always worth it (although it is always possible that we did not wait long enough). One might agree with Skinner that betting on a new theory carries considerable risk. However, the history of science shows that scientists seem to be willing to take the risk, at least occasionally. Betting on new theories may not be common in science, but it has contributed to scientific progress.

REFERENCES

Bechtel, W. (2005). The challenge of characterizing operations in the mechanisms underlying behavior. *Journal of the Experimental Analysis of Behavior, 84*, 313–325.

Bennett, M. R., & Hacker, P. M. S. (2003). *Philosophical foundations of neuroscience.* Malden, MA: Blackwell.

Berger, T. W., Alger, B., & Thompson, R. F. (1976, April 30). Neuronal substrate of classical conditioning in the hippocampus. *Science, 192*, 483–485.

Bindra, D. (1972). A unified account of classical and operant training. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 453–481). New York: Appleton-Century-Crofts.

Brown, P. L., & Jenkins, H. M. (1968). Auto-shaping of the pigeon's keypeck. *Journal of the Experimental Analysis of Behavior, 11*, 1–8.

Burgos, J. E. (2001). A neural-network interpretation of selection in learning and behavior. *Behavioral and Brain Sciences, 24*, 531–532.

Burgos, J. E. (2003). Theoretical note: Simulating latent inhibition with selection neural networks. *Behavioural Processes, 62*, 183–192.

Burgos, J. E. (2005). Theoretical note: The *C/T* ratio in artificial neural networks. *Behavioural Processes, 69*, 249–256.

Burgos, J. E., & Donahoe, J. W. (2000). Structure and function in selectionism: Implications for complex behavior. In J. Leslie, & D. Blackman (Eds.), *Issues in experimental and applied analyses of human behavior* (pp. 39–57). Reno, NV: Context Press.

Burgos, J. E., & Murillo, E. (in press). Neural network simulations of two context-dependence phenomena. *Behavioural Processes.*

Colwill, R. M., & Rescorla, R. A. (1986). Associative structures in instrumental learning. In G. H. Bower (Ed.), *The psychology of learning and motivation: Vol. 20* (pp. 55–104). New York: Academic Press.

Colwill, R. M., & Rescorla, R. A. (1990). Evidence for the hierarchical structure of instrumental learning. *Animal Learning & Behavior, 18*, 71–82.

Christian, K. M., & Thompson, R. F. (2003). Neural substrates of eyeblink conditioning: Acquisition and retention. *Learning and Motivation, 10*, 427–455.

Dinsmoor, J. A. (1985). The role of observing and attention in establishing stimulus control. *Journal of the Experimental Analysis of Behavior, 43*, 365–381.

Dinsmoor, J. A. (1995). Stimulus control: I. *The Behavior Analyst, 18*, 51–68.

Dinsmoor, J. A. (2001). Stimuli inevitably generated by behavior that avoids electric shock are inherently reinforcing. *Journal of the Experimental Analysis of Behavior, 75*, 311–333.

Donahoe, J. W. (2002). Behavior analysis and neuroscience. *Behavioural Processes, 57*, 241–259.

Donahoe, J. W., & Burgos, J. E. (1999). Timing without a timer. *Journal of the Experimental Analysis of Behavior, 71*, 257–263.

Donahoe, J. W., & Burgos, J. E. (2000). Behavior analysis and revaluation. *Journal of the Experimental Analysis of Behavior, 74*, 331–346.

Donahoe, J. W., Burgos, J. E., & Palmer, D. C. (1993). A selectionist approach to reinforcement. *Journal of the Experimental Analysis of Behavior, 60*, 17–40.

Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior.* Boston: Allyn & Bacon.

Donahoe, J. W., Palmer, D. C., & Burgos, J. E. (1997a). The S-R issue: Its status in behavior analysis and in Donahoe and Palmer's *Learning and Complex Behavior. Journal of the Experimental Analysis of Behavior, 67*, 193–211.

Donahoe, J. W., Palmer, D. C., & Burgos, J. E. (1997b). The unit of selection: What do reinforcers reinforce? *Journal of the Experimental Analysis of Behavior, 67*, 259–273.

Gamzu, E., & Schwam, E. (1974). Autoshaping and automaintenance of a key-press response in squirrel monkeys. *Journal of the Experimental Analysis of Behavior, 21*, 361–371.

Gamzu, E., & Williams, D. R. (1971, March 5). Classical conditioning of a complex skeletal act. *Science, 171*, 923–925.

Gamzu, E., & Williams, D. R. (1973). Associative factors underlying the pigeon's key pecking in autoshaping procedures. *Journal of the Experimental Analysis of Behavior, 19*, 225–232.

Gibbon, J., & Balsam, P. (1981). Spreading association in time. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 219–253). New York: Academic Press.

Gibbs, C. M., Latham, S. B., & Gormezano, I. (1978). Classical schedule and resistance to extinction. *Animal Learning & Behavior, 6*, 209–215.

Gonzalez, F. A. (1973). Effects of partial reinforcement (25%) in an autoshaping procedure. *Bulletin of the Psychonomic Society, 2*, 299–301.

Gonzalez, F. A. (1974). Effects of varying the percentage of key illuminations paired with food in a positive automaintenance procedure. *Journal of the Experimental Analysis of Behavior, 22*, 483–490.

Gormezano, I., & Kehoe, E. J. (1975). Classical conditioning: Some methodological-conceptual issues. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 2. Conditioning and behavior theory* (pp. 143–179). Hillsdale, NJ: Erlbaum.

Gormezano, I., & Moore, J. W. (1969). Classical conditioning. In M. H. Marx (Ed.), *Learning: Processes* (pp. 121–203). London: Macmillan.

Green, L. (Ed.). (2005), The relation of behavior and neuroscience [Special issue]. *Journal of the Experimental Analysis of Behavior, 84* (3).

Grossberg, S. (1968). Some physiological and biochemical consequences of psychological postulates. *Proceedings of the National Academy of Sciences, 60*, 758–765.

Hearst, E. (1975). The classical-instrumental distinction: Reflexes, voluntary behavior, and categories of associative learning. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes: Vol. 2. Conditioning and behavior theory* (pp. 181–223). Hillsdale, NJ: Erlbaum.

Hearst, E., & Jenkins, H. M. (1974). Sign-tracking: The stimulus-reinforcer relation and directed action. *Monograph of the Psychonomic Society.* Austin, TX,

Herrnstein, R. J., & Loveland, D. H. (1972). Food avoidance in hungry pigeons and other perplexities. *Journal of the Experimental Analysis of Behavior, 18*, 369–383.

Hursh, S. R., Navarick, D. J., & Fantino, E. (1974). ''Automaintenance'': The role of reinforcement. *Journal of the Experimental Analysis of Behavior, 21*, 112–124.

Jenkins, H. M., Barnes, R. A., & Barrera, F. J. (1981). Why autoshaping depends on trial spacing. In C. M. Locurto, H. S. Terrace, & J. Gibbon (Eds.), *Autoshaping and conditioning theory* (pp. 255–284). New York: Academic Press.

Jenkins, H. M., & Moore, B. R. (1973). The form of the autoshaped response with food or water reinforcers. *Journal of the Experimental Analysis of Behavior, 20*, 163–181.

Kehoe, E. J. (1988). A layered network model of associative learning: Learning to learn and configuration. *Psychological Review, 95*, 411–433.

Konorski, J., & Miller, S. (1937a). On two types of conditioned reflex. *Journal of General Psychology, 16*, 264–272.

Konorski, J., & Miller, S. (1937b). Further remarks on two types of conditioned reflex. *Journal of General Psychology, 17*, 405–407.

Meeter, M., Myers, C. E., & Gluck, M. A. (2005). Integrating incremental learning and episodic memory models of the hippocampal region. *Psychological Review, 112*, 560–585.

Miller, S., & Konorski, J. (1928). On a particular form of conditioned reflex (B. F. Skinner, Trans.). *Journal of the Experimental Analysis of Behavior, 12*, 187–189, (Reprinted from *Comptes Rendus Hebdomadaires des Séances et Mémoires de la Société de Biologie, 99*, 1155–1157).

Miller, R. R., & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation: Vol. 22* (pp. 51–92). San Diego, CA: Academic Press.

Mowrer, O. H. (1947). On the dual nature of learning: A reinterpretation of ''conditioning'' and ''problem-solving.'' *Harvard Educational Review, 17*, 102–148.

Myerson, J., Myerson, W. A., & Parker, B. K. (1979). Automaintenance without stimulus-change reinforcement: Temporal control of key pecks. *Journal of the Experimental Analysis of Behavior, 31*, 395–403.

Notterman, J. M. (1959). Force emission during bar pressing. *Journal of Experimental Psychology, 58*, 341–347.

Notterman, J. M., & Mintz, D. E. (1965). *Dynamics of response.* New York: Wiley.

Pan, W.-X., Schmidt, R., Wickens, J. R., & Hyland, B. I. (2005). Dopamine cells respond to predicted events during classical conditioning: Evidence for eligibility traces in the reward-learning network. *Journal of Neuroscience, 25*, 6235–6242.

Pavlov, I. P. (1941). *Lectures on conditioned reflexes, Vol. 2: Conditioned reflexes and psychiatry* (W. H. Gantt, Trans.). New York: International.

Pavlov, I. P. (1955). *Selected works* (S. Belsky, Trans.). Moscow: Foreign Languages Printing House.

Pear, J. J., & Eldridge, G. D. (1984). The operant-respondent distinction: Future directions. *Journal of the Experimental Analysis of Behavior, 42*, 453–467.

Pear, J. J., & Legris, J. A. (1987). Shaping by automated tracking of an arbitrary operant response. *Journal of the Experimental Analysis of Behavior, 47*, 241–247.

Power, J. M., Thompson, L. T., Moyer, J. R., & Disterhoft, J. F. (1997). Enhanced synaptic transmission in CA1 hippocampus after eyeblink conditioning. *Journal of Neurophysiology, 78*, 1184–1187.

Rescorla, R. A. (1991). Associative relations in instrumental learning: The 18th Bartlett Memorial Lecture. *Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology, 43*, 1–23.

Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: Relationships between Pavlovian and instrumental learning. *Psychological Review, 74*, 151–182.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.

Ricci, J. A. (1973). Keypecking under response-independent food presentation after long simple and compound stimuli. *Journal of the Experimental Analysis of Behavior, 19*, 509–516.

Sanabria, F., Sitomer, M. T., & Killeen, P. R. (2006). Negative automaintenance omission training is effective. *Journal of the Experimental Analysis of Behavior, 86*, 1–10.

Schmajuk, N. A., & DiCarlo, J. (1992). Stimulus configuration, classical conditioning, and hippocampal function. *Psychological Review, 99*, 268–305.

Schultz, W. (2002, October 10). Getting formal with dopamine and reward. *Neuron, 36*, 241–263.

Schultz, W., Dayan, P., & Montague, P. R. (1997, March 14). A neural substrate of prediction and reward. *Science, 275*, 1593–1599.

Schwartz, B., & Gamzu, E. (1977). Pavlovian control of operant behavior: An analysis of autoshaping and its implication for operant conditioning. In W. K. Honig, & J. E. R. Staddon (Eds.), *Handbook of operant behavior* (pp. 53–97). Englewood Cliffs, NJ: Prentice-Hall.

Schwartz, B., & Williams, D. R. (1972a). The role of the response-reinforcer contingency in negative automaintenance. *Journal of the Experimental Analysis of Behavior, 17*, 351–357.

Schwartz, B., & Williams, D. R. (1972b). Two different kinds of key peck in the pigeon: Some properties of responses maintained by negative and positive response-reinforcer contingencies. *Journal of the Experimental Analysis of Behavior, 18*, 201–216.

Sidman, M., & Fletcher, F. G. (1968). A demonstration of autoshaping in monkeys. *Journal of the Experimental Analysis of Behavior, 11*, 307–309.

Skinner, B. F. (1935). Two types of conditioned reflex and a pseudo type. *Journal of General Psychology, 12*, 66–77.

Skinner, B. F. (1937). Two types of conditioned reflex: A reply to Konorski and Miller. *Journal of General Psychology, 16*, 272–279.

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

Skinner, B. F. (1950). Are theories of learning necessary? *Psychological Review, 57*, 193–216.

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

Skinner, B. F. (1974). *About behaviorism.* New York: Alfred A, Knopf.

Smith, S. G., & Smith, W. M., Jr. (1971). A demonstration of autoshaping with dogs. *Psychological Record, 21*, 377–379.

Squier, L. H. (1969). Auto-shaping key responses in fish. *Psychonomic Science, 17*, 177–178.

Stiers, M., & Silberberg, A. (1974). Lever-contact responses in rats: Automaintenance with and without a negative response-reinforcer dependency. *Journal of the Experimental Analysis of Behavior, 22*, 497–506.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review, 88*, 135–170.

Terrace, H. S., Gibbon, J., Farrell, L., & Baldock, M. D. (1975). Temporal factors influencing the acquisition of an autoshaped key peck. *Animal Learning & Behavior, 3*, 53–62.

Timberlake, W., & Grant, D. L. (1975, November 14). Auto-shaping rats to the presentation of another rat predicting food. *Science, 190*, 690–692.

Tomie, A., Brooks, W., & Zito, B. (1989). Sign-tracking: The search for reward. In S. B. Klein, & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 191–223). Hillsdale, NJ: Erlbaum.

Trapold, M. A., & Overmier, J. B. (1972). The second learning process in instrumental learning. In A. H. Black, & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 427–452). New York: Appleton-Century-Crofts.

Wagner, A. (1981). SOP: A model of automatic memory processing in animal behavior. In N. Spear, & R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5–47). Hillsdale, NJ: Erlbaum.

Wagner, A. R., & Donegan, N. H. (1989). Some relationships between a computational model (SOP) and a neural circuit for Pavlovian (rabbit eyeblink) conditioning. In R. D. Hawkins, & G. H. Bower (Eds.), *The psychology of learning and motivation, Vol. 22: Computational models of learning in simple neural systems* (pp. 157–203). New York: Academic Press.

Wasserman, E. A. (1973, September 7). Pavlovian conditioning with heat reinforcement produces stimulus-directed pecking in chicks. *Science, 181*, 875–877.

Wessells, M. G. (1974). The effects of reinforcement upon the prepecking behaviors of pigeons in the autoshaping experiment. *Journal of the Experimental Analysis of Behavior, 21*, 125–144.

Williams, D. R., & Williams, H. (1969). Auto-maintenance in the pigeon: Sustained pecking despite contingent non-reinforcement. *Journal of the Experimental Analysis of Behavior, 12*, 511–520.

Wyckoff, L. B., & Jr. (1952). The role of observing responses in discrimination learning. Part I. *Psychological Review, 59*, 431–442.

Zeigler, H. P., & Wyckoff, L. B., Jr. (1961). Observing responses and discrimination learning. *Quarterly Journal of Experimental Psychology, 13*, 129–140.

Zener, K. (1937). The significance of behavior accompanying conditioned salivary secretion for theories of conditioned response. *American Journal of Psychology, 50*, 384–403.

Zipser, D. (1986). A model of hippocampal learning during classical conditioning. *Behavioral Neuroscience, 100*, 764–776.

Received: August 13, 2004
Final acceptance: April 10, 2007