

A Theory of Conditioning: Inductive Learning Within Rule-Based Default Hierarchies

Keith J. Holyoak
University of California, Los Angeles

Kyunghee Koh and Richard E. Nisbett
University of Michigan

We present a theory of classical conditioning based on a parallel, rule-based performance system integrated with mechanisms for inductive learning. Inferential heuristics are used to add new rules to the system in response to the relation between the system's predictions and environmental input. A major heuristic is based on "unusualness": Novel cues are favored as candidates to predict events that are important or unexpected. Rules have strength values that are revised on the basis of feedback. The performance system allows rules to operate in parallel, competing to control behavior and obtain reward for successful prediction of important events. Sets of rules can form default hierarchies: Exception rules censor useful but imperfect default rules, protecting them from loss of strength. The theory is implemented as a computer simulation, which is used to model a broad range of conditioning phenomena, including blocking and overshadowing, the impact of statistical predictability on conditioning, and conditioned inhibition. The theory accounts for a variety of phenomena that previous theories have not dealt with successfully.

Intelligence manifests itself in the adaptation of goal-directed systems to complex and potentially dangerous environments. In order to acquire nourishment and avoid injury, an animal must process a large and ever-changing array of perceptual information. From this array it must identify cues that signal important events and occasions for goal-appropriate actions. It must learn significant regularities even though valid cues are only imperfectly predictive, often conflict with one another, and are embedded among a plethora of irrelevancies.

The kinds of learning that underlie such adaptation fall under the rubric of *induction*, broadly defined as those inferential processes that expand knowledge in the face of uncertainty (J. H. Holland, Holyoak, Nisbett, & Thagard, 1986). J. H. Holland et al. presented a framework for induction that encompasses phenomena ranging from animal learning to human categorization, analogical reasoning, and scientific discovery. The basic tenet of the framework is that inductive change in a cognitive system is directed toward the development of *mental models* of the environment that guide inferences in the course of problem solving and undergo revision on the basis of feedback. Mental models can be constructed from sets of *condition-action rules*

(also called *production rules*), which have the general form, "IF Condition 1, Condition 2, . . . Condition *n*, THEN Action 1, Action 2, . . . Action *n*" (Newell, 1972). The set of rules in the system's repertoire is not fixed; rather, it is continually revised by inductive mechanisms of two basic types: (a) mechanisms for revising the assessment of the utility of existing rules, that is, rewarding or punishing them in accord with feedback regarding the consequences of their actions, and (b) mechanisms for generating plausibly useful new rules. Various inferential heuristics guide the process of rule generation.

The framework proposed by J. H. Holland et al. (1986) stresses the importance of parallelism and of the idea that rules interact among themselves. Rules enter into a variety of relations that can be complementary or competitive in nature. Multiple rules can complement each other when they provide consistent information that can be integrated to select an appropriate action. They can also compete to control behavior and obtain reward. Interaction among rules generates *default hierarchies*, in which relatively general rules govern the expectations of the system unless they are overridden by more specific rules that incorporate more detailed aspects of the situation. Default hierarchies can be found in human representations of natural categories (e.g., "If it's a bird, then it flies"; "If it's a penguin, then it doesn't"), social stereotypes, and naive physical knowledge (J. H. Holland et al., 1986, chap. 7). In the present article we present a theory, derived from the general framework, that applies to some of the simplest forms of inductive learning observed in studies of Pavlovian (classical) conditioning in animals.

In a typical experiment of the sort with which we will be concerned, a rat is first trained to press a lever to get food. After many sessions of such lever pressing, a distinctive tone (the conditioned stimulus, or CS) is presented for several seconds. Just as the tone goes off, a shock (the unconditioned stimulus, or US; also termed the *reinforcer*) is delivered to the rat's feet. As this

Preparation of this article was supported by National Science Foundation Grants BNS-8615316 to Keith J. Holyoak, SES-8507342 and BNS-8709892 to Richard E. Nisbett, and BNS-9409198 to Keith J. Holyoak and Richard E. Nisbett. The research reported here was described in a paper presented at the Ninth Annual Conference of the Cognitive Science Society, Seattle, Washington, July 1987.

We are indebted to John H. Holland for advice and assistance. We thank Tom Minor, Christopher Peterson, and three anonymous reviewers, for their helpful comments on earlier drafts. A fourth anonymous reviewer expended extraordinary effort in sparing the article from errors and helping us to refine the theory.

Correspondence concerning this article should be addressed to Keith J. Holyoak, Department of Psychology, University of California, Los Angeles, Los Angeles, California 90024.

sequence of events is repeated, the rat soon begins to show signs of fear when the tone is heard. While the tone is on, the animal suppresses its routine lever pressing and eating, and displays a collection of behaviors, such as crouching, that constitute a conditioned emotional response (CER). The tone now signals shock and the rat exhibits fear in response to it.

Studies using variants of this deceptively simple paradigm, as well as other Pavlovian procedures, have amassed a range of phenomena that have eluded comprehensive treatment within a single model, despite numerous theoretical efforts. From our perspective, classical conditioning provides a microcosm of some of the most fundamental aspects of induction. In particular, the rat in its conditioning chamber must implicitly solve the problem of detecting pragmatically important covariations among cues provided by a changing environment. This basic inductive task must be accomplished by any realistic adaptive system, whether it be a lower animal, a human, or an intelligent computer program. We will argue that conditioning can be well understood in terms of mechanisms for generating and revising systems of rules that constitute default hierarchies. Given the wealth of systematic data available from conditioning studies, the topic offers a challenging initial proving ground for theories of induction.

We will begin by briefly reviewing some of the key phenomena observed in conditioning experiments, major theoretical proposals, and the problems that have beset them. Next, we will sketch the general framework for induction presented by J. H. Holland et al. (1986) and describe the specific theory of conditioning derived from the framework. The theory, embodied in a computer simulation, takes the form of a parallel, rule-based performance system integrated with inductive-learning mechanisms. We will describe a series of simulations of various conditioning phenomena, including several that previous theories have not dealt with adequately. We will argue that our rule-based system is able to account for findings that cannot be explained by traditional associationist models of conditioning. Finally, we will describe some of the directions in which the present theory could be profitably extended.

Classical Conditioning: Phenomena and Theoretical Accounts

Association Versus Prediction

Early theories of conditioning generally assumed that temporal contiguity of the CS and the US was necessary and sufficient to establish a conditioned response (CR). The central contribution of modern work on the topic, achieved in the late 1960s and early 1970s, has been to establish that temporally contiguous pairing is in fact neither necessary nor sufficient for learning to take place.

Temporal contiguity is not necessary because conditioning can sometimes occur when there is an interval of minutes or even of hours between the to-be-associated stimuli. If rats eat or drink a novel substance and are made ill some minutes or hours later, they will appear to attribute their illness to that food or drink and, after a single experience, form a strong and lasting aversion to it (e.g., Garcia, McGowan, Ervin, & Koelling, 1968). The phenomenon of taste aversion, which is presumably

based on specific biological predispositions toward learning certain types of associations, is only the extreme of a continuum of successful conditioning that occurs despite lags between CS and US presentation. There are many demonstrations of classical and instrumental conditioning in which the delay between events is on the order of many seconds or minutes (e.g., Kamin, 1965; for a recent review, see Mackintosh, 1983). Thus, although it may well be necessary that representations of the CS and US be active at the same time, it is not essential for the stimuli to be presented in a contiguous fashion.

In addition, temporal contiguity is not sufficient for conditioning because an event may be paired with another event and still not result in conditioning. This fact was established by the *blocking* experiments of Kamin (1968). In these studies, a normally effective CS, occurring in close temporal relation to a reinforcer, showed little ability to produce conditioning if it was paired with another CS that had already been established as a signal for that reinforcer. For example, Kamin taught rats that a noise signaled electric shock. That the rats had learned this association was established by the fact that when they were placed in a situation in which they could press a lever to obtain food, they pressed the lever less frequently during periods when the noise was on than during periods when it was off. The rats were then exposed to two stimuli simultaneously, namely, the original noise and a light, and this compound was then reinforced by the shock. Rats subsequently exposed to the light alone showed little suppression of lever pressing. The ability of the noise to predict the shock blocked learning of the fact that light also predicted the shock. The phenomenon of *overshadowing* (Pavlov, 1927), in which conditioning to a target cue is diminished by the presence of another cue in compound with the target, reveals similar effects to blocking. In a variant of overshadowing (e.g., Wagner, 1969), rats are exposed on alternate trials to a compound stimulus (consisting of a light and tone, for instance) that predicts shock, and a single stimulus (for instance, the light) that also predicts shock. Such a procedure establishes the light as a more valid predictor of shock, greatly reducing conditioning to the less valid tone predictor.

It is not necessary to have a competing stimulus to show that pairing is not sufficient for conditioning. Rescorla (1968) demonstrated that if the probability of the reinforcer is the same in the absence of the CS as in its presence, little or no conditioning results. In Rescorla's "uncorrelated" procedure, the shock US and the CS are repeatedly presented, but the shock is no more likely to occur when the CS has recently occurred than when it has not recently occurred. This procedure does not establish the CS as a predictor of the shock and, accordingly, the rat does not behave as if it were especially fearful when the CS is presented. It is only when pairing serves to increase predictability in a statistical sense that conditioning takes place.

Further difficulties for the notion that pairing with the US is sufficient to generate excitatory conditioning are posed by studies of *conditioned inhibition*, a phenomenon first described by Pavlov (1927) and discussed extensively by Konorski (1948). In one experiment, Kamin (1968) trained rats to associate noise with shock. He then paired the noise with light. This compound was presented for several trials but was never paired with shock. If only second-order conditioning was operating in this procedure, the pairings of the neutral light with the excitatory tone

would cause the light to also come to evoke fear. In fact, however, the effect of the light was to inhibit fear, which it did from the very first trial. An experiment by Rescorla (1972b) showed that a stimulus can be excitatory when first presented but then can become inhibitory. Rats were first presented with flashing lights that signaled shock. Then a series of trials was begun in which the light continued to signal shock, but these trials were interspersed with trials on which the light was preceded by a tone. On the light-plus-tone trials, the rats were never shocked. Initially, the tone caused the animals to show a fear reaction. Over time, however, as the animals learned that the light-tone combination would never be reinforced, the tone ceased to elicit much fear.

The phenomena associated with conditioned inhibition indicate that whether conditioning is excitatory or inhibitory depends on the information that the cue provides about the occurrence of the US, and cannot be predicted by temporal contiguity alone. Although contiguity of the CS and US facilitates conditioning in many circumstances, other factors must also be considered. (See Mackintosh, 1983, and Rescorla & Holland, 1982, for recent reviews of these and other conditioning phenomena.)

The Rescorla-Wagner Theory

Intuitively, many of the results just described can be interpreted as evidence that animals attempt to learn to predict important events, such as the occurrence of shock. A CS can be blocked by another stimulus that already predicts the shock or be overshadowed by another stimulus that is a better predictor. Occasional pairing does not result in conditioning unless the reinforcer is better predicted by the presence of the CS than by its absence. Rescorla and Wagner (1972) captured this intuition in their important revision of previous associationist accounts of conditioning. Their theory provides an elegant account of a wide range of conditioning phenomena. Although widely criticized for various empirical shortcomings, the Rescorla-Wagner theory remains extremely influential and will serve as a standard with which our own account can be compared.

To deal with the results we have just reviewed, Rescorla and Wagner provided a simple linear model, describing changes in the strength (V_{CS}) of the association between a particular CS and the US with which it is paired. Two of the basic principles underlying the model were adapted from Hull's (1943) theory and from the Bush and Mosteller (1955) linear model of conditioning. The first of these was that the amount of conditioning that can occur on any given trial is limited by the amount of conditioning that is still possible. Early on, when a great deal of conditioning is still possible, the incremental effect of a reinforced trial will be large, but later on, when conditioning is near its asymptote, the possible increase is small. The second principle was that the amount of conditioning that can occur on any given trial, as well as the total that can ultimately occur, is limited by the magnitude of the US. A weak US can only sustain small increments on each trial and can never produce very great associative strength with any given CS.

The Rescorla and Wagner contribution was to add a third principle, namely, that the associative strength of a given stimulus is limited not only by its own present strength but also by the

associative strengths that any concurrent stimuli already have. Thus, the amount of conditioning that can accrue to a given stimulus A is limited by the amount of conditioning that has already accrued to some other stimulus X that is also present. These three principles are embodied in the following linear model describing strength revision:

$$\Delta V_{CS} = \alpha_{CS}(\lambda_{US} - \sum_{j=1}^n V_j), \quad (1)$$

where α_{CS} is a constant that determines how fast conditioning can occur for a given CS, λ_{US} denotes the asymptotic limit of conditioning that can be supported by the US, and $\sum V_j$ represents the sum of the current strengths of associations to the US from the stimuli present (the particular CS plus all other concurrent cues).¹

Equation 1 of the Rescorla-Wagner model provides a qualitative account of many of the results we have just discussed. If the animal is first conditioned to cue A, and then exposed to an AX compound that is followed by reinforcement, conditioning to X will be blocked because $\sum V_A$ is already equal to λ_{US} . Similarly, degrees of overshadowing occur whenever other cues are established as equally good or better predictors of the US than is some given CS that occurs in a compound. The same mechanism accounts for why random pairings of a CS and US produce little or no conditioning at asymptote. In this case the relatively constant background cues (e.g., the odor of the cage), which are paired with the US both in the presence and absence of the CS, will acquire greater associative strength than will the CS.

The primary reason that the Rescorla-Wagner theory has been influential is that it provides an elegant unified account of a large number of phenomena observed in classical conditioning, and can be generalized to account for parallel phenomena that arise in studies of instrumental conditioning (Rescorla & Holland, 1982). The model captures the notion that learning depends on the comparison of internally generated predictions (based on associative strengths of cues) to environmental inputs (the magnitude of the presented US). The model also embodies the essential insight that cues compete with one another to predict the US and that the most valid and most general predictors are favored in the competition.

The Rescorla-Wagner model remains influential, also, because of its close relation to a class of adaptive algorithms that have been proposed independently in contexts far removed from animal conditioning. Equation 1, the Rescorla-Wagner rule for strength revision, is essentially equivalent to the Widrow-Hoff rule familiar in adaptive-systems theory (Widrow & Hoff, 1960) and is a generalization of Rosenblatt's (1962) perceptron rule (see Sutton & Barto, 1981). The Rescorla-Wagner model is thus closely related to a major class of strength-revision procedures used to model associative learning within adaptive networks of the sort currently being explored by connectionist

¹ In the full Rescorla-Wagner equation for strength revision, the rate of learning is the product of α_{CS} and a second parameter, β_{US} , which reflects properties of the unconditioned stimulus, such as its intensity. This additional parameter does not alter any of the implications of the model that we will discuss, and therefore is omitted here for simplicity.

theorists (Rumelhart, Hinton, & Williams, 1986). Sutton and Barto (1981) provided a neural-network instantiation of an extended version of the Rescorla-Wagner theory. Within the connectionist framework, all learning is viewed as the product of incremental changes in connection strengths with experience.

Problems With the Rescorla-Wagner and Other Associative Models

Despite its notable empirical successes, the Rescorla-Wagner model does not account well for a wide variety of other phenomena, some of which antedate their treatment and some of which are more recent. Some of these phenomena simply lie outside the scope of the model, whereas others constitute clear failures of predictions. These limitations (many of which were pointed out by Rescorla and Wagner themselves) have spurred development of numerous other models that provide refinements and alternatives (Frey & Sears, 1978; Mackintosh, 1975b; Pearce, 1987; Pearce & Hall, 1980; Wagner, 1978, 1981). We will not describe these other proposals in any detail here (see Mackintosh, 1983; Rescorla & Holland, 1982) but will mention several of their features that relate to our own model. All of the models subsequent to that of Rescorla and Wagner (1972) adopt variations of the standard associationist framework in which conditioning is treated solely in terms of strength revision. In our view, none solves all of the empirical difficulties that beset the Rescorla-Wagner theory while preserving its successes. Here we describe some of these difficulties, focusing on those that have had the greatest influence on development of our theory.

Effects of CS preexposure. A major limitation of the Rescorla-Wagner formulation is that it does not account for changes in the processing of CSs with experience. Equation 1 predicts that a stimulus that is uncorrelated with a reinforcer will begin and end with zero associative strength. No distinction is drawn between stimuli that the animal has encountered before and those that it has not. But, in fact, conditioning is retarded if the to-be-conditioned stimulus has been presented previously, a phenomenon termed *latent inhibition*. Retardation is especially severe if the cue had been presented, in an uncorrelated fashion, with the reinforcer ("learned irrelevance"), but is obtained even if the CS but not the US is preexposed (Baker, 1976; Baker & Mackintosh, 1977; Gamzu & Williams, 1973; Kremer, 1971; Siegel & Domjan, 1971). Note that the preexposed cue is retarded in both excitatory and inhibitory learning (Baker & Mackintosh, 1977; Rescorla, 1971a). The fact that inhibitory as well as excitatory learning is retarded indicates that the term *latent inhibition* is actually a misnomer, as CS preexposure does not produce conditioned inhibition.

Such phenomena have led to a variety of proposals regarding how learning can affect CS processing, typically formalized in terms of changes with experience in the α_{CS} parameter of Equation 1. Two major suggestions are that the associability of a CS is reduced as it becomes more predictable, and hence is less surprising because of repetitions in a given context (Wagner, 1978), and that associability is reduced if the animal learns that the cue is a relatively poor predictor of a US (Mackintosh, 1975b). Within the account we will propose, latent inhibition and related phenomena are interpreted as consequences of gen-

eral inferential heuristics for the initial selection of potentially predictive cues.

Conditioned inhibition and nonindependent cues. The Rescorla-Wagner model can account for the basic phenomenon of conditioned inhibition, as established in an A+, AX- paradigm (that is, stimulus A is always followed by the US, whereas the compound stimulus AX never is). The model assumes that strength values can be negative as well as positive. Cue A will reach an asymptotic strength equal to λ_{US} , whereas X will reach an asymptote equal to $-\lambda_{US}$; thus, the net strength of the AX compound will be 0.

However, this account incorrectly predicts that inhibitory conditioning effects, once established, should be extinguishable by presenting the inhibitory CS alone, in the absence of either the US or of excitatory CSs. As Zimmer-Hart and Rescorla (1974) put it,

Assuming that nonreinforcement supports a zero asymptote . . . a simple nonreinforcement of a previously established inhibitor should produce a change. If V_X is negative, then the quantity $(0 - V_X)$ is positive and consequently V_X should be incremented toward zero when it is separately nonreinforced. That is, the theory predicts that repeated nonreinforced presentation of an inhibitor should attenuate that inhibition. (pp. 837-838)

A subsequent variant of the Rescorla-Wagner model proposed by Frey and Sears (1978) weakens but does not eliminate this prediction.

Although it is required by the Rescorla-Wagner formulation, this prediction is not intuitively satisfying. When a stimulus has predicted the nonoccurrence of a US, then repeated presentations of the stimulus in the absence of the US will provide additional confirmations of the expectation of nonoccurrence, and hence might be expected to enhance its inhibitory properties. And indeed (subject to qualifications to be discussed later), this is the result obtained by Zimmer-Hart and Rescorla (1974). Although some models predict that reinforcement will not extinguish conditioned inhibition (Sutton & Barto, 1987; Wagner, 1981), none can account for actual enhancement. Later studies have also demonstrated failures to extinguish conditioned inhibition (Baker & Baker, 1985; DeVito & Fowler, 1986; P. C. Holland & Gory, 1986; Pearce, Nicholas, & Dickinson, 1982), and under some conditions enhancement has been obtained (Miller & Schachtman, 1985). (See Miller & Schachtman for arguments against the possibility that apparent enhancement results entirely from extinction of second-order excitatory associations to the inhibitor.)

A corollary prediction of the Rescorla-Wagner model is that pairing a neutral cue (a CS that is either novel or previously extinguished) with an inhibitory CS, without reinforcement, will result in excitatory conditioning of the neutral cue. For if the neutral cue has a V of 0 and the inhibitory CS has a V that is negative, then the net strength of the compound will be negative, and the discrepancy from the 0 asymptote will also be negative. Both strength values will therefore be incremented, resulting in a positive V value for the previously neutral cue. Although an initial test of this prediction yielded positive results (Rescorla, 1971b), attempts at replication have failed (Baker, 1974; Soltysik, 1985).

In our view, these failures of the Rescorla-Wagner model highlight two major deficiencies. First, the model fails to pro-

vide explicit representations of the nonoccurrence of expected events. It appears necessary to assume not simply that a cue acquire negative strength with respect to its association with the US, but rather that it acquire positive strength with respect to its association with \overline{US} . Accordingly, the animal must be able to explicitly represent \overline{US} , the failure of the US to occur, as a type of event to be predicted (Konorski, 1967; Pearce & Hall, 1980).

The second deficiency, which the Rescorla-Wagner model shares with other associationist models of learning (with the important exception of multilayer connectionist networks; see Rumelhart et al., 1986), is that learning is limited so as to accrue independently to individual cues. The response elicited by a compound AX cue is assumed to be an additive function of the responses elicited by the A cue and by the X cue. However, an environment that gives rise to conditioned inhibition is one in which multiple CSs are nonindependent. In particular, the A cue generally predicts shock, but not when X also occurs. In terms we will use in our model, the association between A and occurrence of the US is the default, whereas presentation of A in conjunction with X gives rise to an exception. We assume the animal will learn explicitly about the interaction between A and X, rather than solely about the effects of each cue considered individually.

Limitations of the independent-cue approach are apparent in studies of transfer of conditioned inhibition. The Rescorla-Wagner formulation predicts that an inhibitory cue will produce complete transfer when it is paired with another CS of comparable excitatory strength to that of the CS with which it was initially paired. This prediction follows from the fact that inhibition is modeled simply as a subtraction from the positive excitation produced by any accompanying excitatory cue. In fact, however, transfer of conditioned inhibition is only partial (e.g., Rescorla, 1969).² This finding supports our contention that the animal explicitly learns that the compound cue AX, as well as its constituent cue X, is not reinforced. The inhibition elicited by a representation of the integrated cue, unlike that elicited by a representation of X, will not extend fully to a different compound that includes X, resulting in partial transfer. The need to represent interactions between cues is emphasized by other "configural" effects on conditioning. For example, with sufficient training, animals can learn to respond appropriately in an A+, X+, AX- environment (Rescorla, 1972a; Woodbury, 1943). Clearly, such adaptive behavior cannot be accounted for by any model that represents the association to a compound cue as the simple sum of associations to its individual constituent cues. (Rudy & Wagner, 1975, provide an excellent discussion of independent-cue and configural approaches to conditioning.)

Configural effects and partial transfer of inhibition have led to suggestions that conditioning accrues not only to independent cues, but also to configural cues that are unique to a compound (Rescorla, 1973; Pearce, 1987). Thus, if an excitatory tone is nonreinforced in the presence of a light, both the light cue and a tone-light configural cue will become inhibitory. Because the tone-light cue is unique to that particular compound, transfer will be incomplete when the light is subsequently paired with a different excitatory CS. Although postulating configural cues is a useful step, the problem of specifying when such

cues become available then arises. If configural cues were direct products of perception, as is commonly assumed, then they presumably would more likely arise when two cues are presented simultaneously, rather than serially. If so, transfer of conditioned inhibition should be less when the original inhibitory learning accrues to a simultaneously presented compound rather than a serial one. In fact, however, simultaneously presented compounds yield greater transfer of inhibition than do serial compounds (P. C. Holland, 1985). In our theory, configural cues are not identified with perceptual units; rather, they emerge during learning as multiple-element conditions of rules.

More generally, recent investigations of conditioned inhibition suggest that the phenomenon is more complex than envisaged by the Rescorla-Wagner model, in ways that fit nicely within a default-hierarchy view. In the Rescorla-Wagner model, inhibition is simply the opposite of excitation—a negative rather than positive strength value. As we have suggested, however, the failures of attempts to show extinction of inhibition by nonreinforcement imply that an inhibitory CS explicitly predicts absence of the US. In addition, there is evidence that inhibitors may also operate by "canceling" the usual relation between an excitatory CS and the US (P. C. Holland, 1985; Jenkins, 1985; Rescorla, 1985). As Rescorla put it,

The stimulus A is readily described as predicting the US when it is presented alone but not when it is accompanied by B; . . . it is natural to emphasize the hierarchical relation that one stimulus has, such that it modulates the strength of the association between two other stimuli. (1985, p. 300)

In one demonstration of modulatory inhibition, Jenkins (1985) placed pigeons in a situation in which a colored dot on a pecking key signaled availability of food *unless* it was accompanied by a noise. Even when the noise alone also signaled food, presentation of the noise inhibited the rate of pecking in the presence of the colored dot. Rescorla (1985) demonstrated a facilitatory version of modulation in an A-, AB+ paradigm. Pigeons learned that pecking a key (A) produced food only when a tone sounded (B), and responded accordingly.

In terms of a default hierarchy, the modulating cue is represented by an exception rule that "censors" a default rule that would otherwise have controlled behavior. A central implication of our view of inhibition is that one of its primary functions is to protect excitatory cues from extinction. That is, an exception rule based on the inhibitor prevents a default rule based on the excitor from operating when the latter rule would have failed. The notion that inhibition involves protection of the excitatory CS from extinction was proposed by Konorski (1948); Soltysik (1985) provides supporting data.

One-trial overshadowing. The Rescorla-Wagner model predicts that blocking or overshadowing will be observed whenever a CS is presented along with another cue that already predicts the US. In fact, however, multiple cues compete for conditioning even on the first trial, when no cue is yet associated with the US. That is, if cue A is paired with reinforcement either alone or in

² The failure to obtain complete transfer of conditioned inhibition can be explained in part, but not entirely, by the offsetting transfer of an excitatory second-order association between the inhibitor and the original excitatory conditioned stimulus (Rescorla, 1984).

conjunction with cue X for one trial, subsequent testing of A reveals less conditioning in the latter case (James & Wagner, 1980; Mackintosh, 1971; Mackintosh & Reese, 1979; Revusky, 1971). Prior conditioning experience is thus not necessary to obtain overshadowing, contrary to the implication of Equation 1. These findings suggest that alternative predictors inevitably compete for reinforcement, even if none of the alternatives has an advantage due to prior association with the US.

In view of the apparent empirical limitations of associationist theories of conditioning, in which learning is treated solely in terms of incremental revision of association strengths, it seems worthwhile to investigate the potential of an alternative approach.

Adaptation Within a Parallel Rule-Based Default Hierarchy

The theory of conditioning that we will present is derived from the framework for induction proposed by J. H. Holland et al. (1986). We will first sketch the general framework and then outline the specific theory of conditioning and describe its embodiment in a computer simulation.

Rule-Based Mental Models

J. H. Holland et al. (1986) proposed that representations of the environment take the form of sets of rules that compose mental models. A mental model is an internal representation that encodes the world into categories and uses these categories to define an internal transition function that mimics the state changes that unfold in the world. A valid mental model has the formal structure of a homomorphism (see J. H. Holland et al., 1986, chap. 2). In the relatively simple world of a rat's conditioning chamber, for example, the animal may learn that an occurrence of an instance of the category "loud tone" signals a transition to the environmental state "painful shock." This piece of knowledge would constitute a small but important portion of the rat's mental model of its environment.

The rat's knowledge about the relation between tones and shocks might be informally represented by a rule such as "If a tone sounds in the chamber, then a shock will occur, so stop other activities and crouch." Rules drive the system's behavior by means of a recognize-act cycle. On each cycle the conditions of rules are matched against representations of active declarative information, which we will term *messages*; rules with conditions that are satisfied by current messages become candidates for execution. For example, if a message representing the recent occurrence of a tone is active, the condition of the above rule will be matched and the actions it specifies may be taken.

Those rules that have their conditions satisfied by active messages will compete to represent the current environment and determine the system's behavior. Rules at different levels of generality form default hierarchies. For example, if tones are typically followed by shock, unless paired with a light, the rule "If tone, then expect shock" can serve as a useful but fallible default, to be overridden by the exception rule "If light and tone, then do not expect shock" when both rules are matched. A set of rules organized as a default hierarchy satisfies a weaker extension of the concept of a homomorphism (termed a *quasi-homo-*

morphism) and, under a broad range of circumstances, provides a more parsimonious representation of a transition function than would an equivalent set of exceptionless rules (J. H. Holland et al., 1986, chap. 2). In the present theory, conditioned inhibition will be analyzed in terms of rules constituting a default hierarchy.

Rules enter into relations that can be complementary as well as competitive. Most proposed rule-based systems have been serial in nature, with only one rule at a time being selected to determine the system's behavior (Anderson, 1983; Newell, 1972). In contrast, the system to be described here involves parallel rule activity (J. H. Holland, 1986). Multiple rules will often act simultaneously to complement and support each other. Through summation of converging evidence from several matched rules, the system can use multiple sources of weak support to arrive at a confident decision. The modular character of rules, coupled with a performance system based on competitive parallelism, provides the kind of flexibility required to represent mental models.

Inductive learning is based on the generation and testing of rules. In the system we will describe, rules have an attached strength parameter that reflects their past usefulness and that is revised on the basis of feedback from the environment. Thus, the strength of the rule "If tone occurs, then shock will occur" will generally be increased if a tone is indeed followed by a shock, and decreased if the anticipated shock is withheld. Rule strength in our model is thus analogous to association strength in the conditioning models reviewed earlier.

Rule-based systems differ from associationist schemes, however, in that learning is based on the generation of new rules, as well as on strength revision. Most associationist models, from early behaviorist theories to current connectionist approaches, tacitly assume that all of the perceived features of the environment are immediately available to be operated on by the conditioning mechanism. In contrast, a rule-based model postulates an initial selection stage that picks out features and feature combinations to form the conditions of rules. (This selection stage has much in common with associationist models that posit an initial attentional process, such as the model of Mackintosh, 1975b; also see Rudy & Wagner, 1975.) Unless a feature is included in a candidate rule, nothing can be learned about its relation to other features or to appropriate behavior. Moreover, a complex environment may contain many features, few of which are likely to be cues that would help form useful rules. For example, a rat may receive a shock while listening to an unfamiliar tone, scratching itself, looking left, and smelling food pellets. Intuitively, we might expect that the rule "if tone, then expect shock" will more likely be generated in this situation than the rule "If looking left, scratching, and smelling pellets, then expect shock." What constraints might guide rule generation of this selective kind?

J. H. Holland et al. (1986) proposed that constraints are provided by inferential heuristics. These include the representative heuristic, which links stimuli that are similar in salient respects (see Rescorla & Furrow, 1977, and Seligman, 1970, for examples of the role of similarity in governing the likelihood of learning in animals); the heuristic that causes precede effects, which will favor conditioning when the CS precedes rather than follows the US; and the heuristic of temporal contiguity, which

preferentially links stimuli that occur relatively close together in time. As these examples suggest, many of the most general heuristics are those used to identify plausible causes of pragmatically important events. (See Einhorn & Hogarth, 1986, for a discussion of human judgments of plausible causes.) The present theory is neutral with respect to the issue of whether particular heuristics are themselves rulelike and subject to learning or whether they are relatively immutable consequences of fundamental operating principles of the system. (But see Nisbett, Fong, Lehman, & Cheng, 1987, for a discussion of the modifiability of inferential heuristics in humans.)

The present proposal incorporates the heuristics based on forward causation and temporal contiguity. Particular emphasis is given to the *unusualness heuristic*, which selects novel stimuli as predictors of later occurrences that require explanation. This heuristic, related to Wagner's (1978) suggestion that greater conditioning accrues to unexpected CSs, encourages the tacit assumption that an unusual event signals further unusual events. In the preceding example, the unusualness heuristic will bias the rat to generate a rule that uses the unexpected tone, rather than more familiar features of the situation, to predict future occurrences of shock.

Given the parallel nature of the system, it will often be the case that several rules will be formed as candidates for predicting important events. A number of rules that make the same prediction may then be matched at once. Such sets of rules will "bid" as a block to attempt to control behavior, thus complementing each other. Only rules that actually succeed in determining behavior will be subject to strength revision. If reward accrues to the winning set of rules, the individual rules will compete among themselves to receive portions of it. This type of competition among alternative hypotheses is analogous to Kelley's (1973) "discounting principle," whereby the existence of alternative possible causes reduces belief in any one of them (see also Einhorn & Hogarth, 1986). Both rule generation and strength revision thus will be governed by general principles that implicitly guide the system in a search for rules that capture causal regularities in its environment.

The J. H. Holland et al. (1986) framework provides a number of specific principles that can be applied to animal conditioning. These include the following:

1. The probability that a set of rules will control behavior increases monotonically with the strengths of the rules in the set relative to the strengths of competing rules that are matched. The function relating strength and response probability must be probabilistic in order to allow opportunities for weaker rules to be tested and to gain strength if they prove more useful than their competitors.

2. Rules form default hierarchies in which useful but imperfect default rules are "protected" from strength reduction by more specific exception rules that can override the defaults in particular circumstances.

3. Only rules that succeed in controlling responses are subject to strength revision. In terms of an economic analogy, rules that control behavior "pay" for the privilege by a reduction in their strength, and must "earn" at least as much reward in the form of a subsequent strength increase in order to make the transaction worthwhile. Rules that do not gain control over re-

sponses produce no consequences for the system and, therefore, neither gain nor lose strength.

4. When multiple rules operate as a set to control behavior, they divide any attendant reward. This competition for reward implies that rules accrue greater reward when they uniquely make a correct prediction than when other rules make the same prediction (the discounting principle). Reward competition provides an inductive pressure that tends to favor general rules over redundant rules that are more specific and impairs learning of new rules that serve the same function as existing strong rules. This principle will prove directly relevant to providing accounts of blocking, overshadowing, and sensitivity to the relative predictive value of cues.

5. New rules are generated in response to particular states of the system that suggest when a new rule might be useful. In the current implementation of our conditioning theory, three triggering situations are identified: (a) the occurrence of an unexpected and important event, (b) the failure of a prediction based on a rule that had previously been highly successful, and (c) the occurrence of an unusual feature in temporal contiguity with a known predictor of an important event.

6. Inferential heuristics favor certain features over others as building blocks for new rules. In particular, unusual features of the environment are favored as candidates to build the conditions of new rules.

A Rule-Based Theory of Conditioning

The framework sketched above requires a much more detailed instantiation in order to constitute a process model of conditioning. Only then will it be possible to determine if the general principles of the framework—*heuristic rule generation, parallel rule competition within a default hierarchy, and competitive strength revision*—can in fact operate as an organized system that produces behavior consistent with results obtained in empirical studies of conditioning.

We have constructed such a process model, embodied in a running computer simulation described in the Appendix. In the subsequent section of the article we will describe results obtained in a series of applications of the simulation to CER phenomena. As is inevitably the case in simulation development, the simulation itself is only an approximate embodiment of the theory on which it is based. Various simplifying assumptions are required; we will attempt to make clear which aspects of the simulation are central and which are more a matter of expedience. For example, the simulation represents both cues, such as tone, and effector actions, such as crouch, as discrete elements, and time is treated simply as a series of cycles of the processing system. These simplifications do not permit us to predict the fine-grained temporal structure of conditioning, or response magnitudes, which are a focus of some conditioning models (Donegan & Wagner, 1987; Sutton & Barto, 1981, 1987; Wagner, 1981). On the other hand, our model is sufficiently detailed with respect to learning and decision-making mechanisms as to allow us to treat in considerable depth a wide range of phenomena involving statistical predictability of the US, blocking and overshadowing, and conditioned inhibition, in-

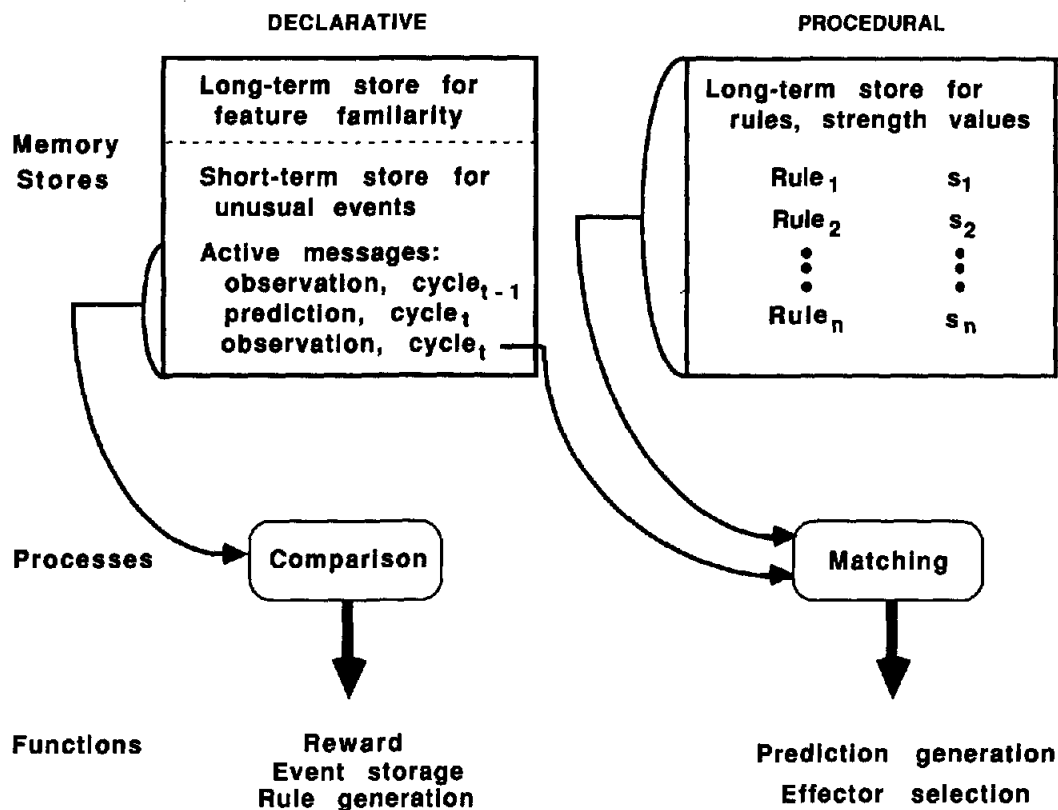


Figure 1. Memory stores and major processes involved in the process model of conditioning. (The comparison process, based on active messages, governs the reward assigned to rules, storage of unusual events in the short-term store, and rule generation. The matching process, based on the message observed on the current cycle and the conditions of rules, determines the generation of a predicted message for the next cycle and the selection of an effector action for the current cycle.)

cluding the phenomena reviewed earlier that are problematic for previous models.³

Figure 1 depicts the basic components of the processing system. In general terms, the system matches the conditions of rules against a message representing the current state of the environment (simulating perceptual input) and uses the matched rules to select an effector action and to generate a message describing the predicted next state of the environment. The predicted message is compared with that observed on the next time step, and the result governs the reward given to rules that generated the prediction, storage of unusual events in a short-term buffer, and triggering of the generation of new rules.

Knowledge in memory stores. As Figure 1 indicates, the model includes both declarative and procedural memory (Anderson, 1983; Squire, 1987). The declarative memory includes a long-term store of all the cues that have appeared in the environment, tagged with a measure of degree of familiarity, and a short-term store for recent unusual events and currently active messages. The procedural memory contains the rules in the system. Each of these types of information is dynamically updated as the model operates in a simulated environment.

The rules in the system are $S \Rightarrow R$ rules in the sense that all rules specify an effector action—either the CER or some alternative behavior that would be emitted in the absence of an ex-

pected aversive reinforcer. The simulation models a prototypical conditioning paradigm experiment in which a rat is pressing a lever for food under a partial reinforcement schedule and occasionally interrupts lever pressing when a US or CS elicits fear. For concreteness, we will refer to the CER as *crouch* and the incompatible background response as *press*. The theory is readily extendable to other conditioning paradigms; the general assumption is that the tendency to emit a response appropriate to the US will be in competition with incompatible alternative effector actions appropriate when the US is absent.

To be more precise, the rules of the model have a form that is not simply $S \Rightarrow R$, but rather $S \Rightarrow S, R$. That is, each rule specifies two action components: a description of the expected next state of the environment and an effector action appropriate to that expected state. Thus, if the expected next state includes the US, *shock*, then the specified effector will be *crouch*; if the expected next state does not include the US, the effector will be *press*. Because the expected state of the environment is directly tied to an appropriate effector action, we often will simply refer to the right-hand side of rules as shock (S) or no shock (\bar{S}), indi-

³ The model presented here differs in a number of significant details from an earlier version sketched in J. H. Holland, Holyoak, Nisbett, and Thagard, 1986, chapter 5.

cating that the rule does or does not yield an expectation of the US. This aspect of the model is presumably oversimplified in that it does not allow rules to represent expected states separately from effector actions, which could be done by means of rule sequences of the form $S_i \Rightarrow S_{i+1}; S_{i+1} \Rightarrow R_{i+1}$. We will discuss extension of the model to rule sequences later in the article; however, rules that jointly specify expectations and effector actions will suffice for our present purposes.

In the computer program, rules are represented in *classifier* notation (J. H. Holland, 1986; J. H. Holland & Reitman, 1978; see Appendix). This notation is useful for implementing the model but is not theoretically important. For purposes of exposition, we will use a more mnemonic notation throughout the article. Thus, the rule "If tone occurs, then expect shock and crouch" will be represented simply as $T \Rightarrow S$. We will use the classifier symbol # to represent a maximally general condition. Thus, the rule "If in the conditioning chamber, expect shock and crouch" will be represented as $\# \Rightarrow S$. A rule with the condition # will be matched on every processing cycle.

The declarative memory system is extremely simple. At any time, the message buffer contains three active messages: one describing the current environment; one describing the environment of the preceding cycle; and one, created on the preceding cycle, predicting the current environment. In addition, the short-term event buffer holds recent "unusual" events that occur in the environment. In the simulation, an unusual event is defined as an observed message that includes the onset of an unfamiliar feature, in which familiarity of a feature is a function of its number of occurrences. This familiarity count for each feature is maintained in the long-term feature store. When an unusual event occurs, the message representing it, with the unfamiliar features tagged, is placed in the short-term event store and held for a few cycles, during which time it may be used by the rule-generation heuristics, described later, to form conditions of new rules.

Message comparison. Each cycle of the processing system begins when a message describing the current state of the environment is received. The feature portion is compared with the corresponding portion of the prediction message posted on the previous cycle. The result determines how much reward, R (if any), is given to the rules that acted on the previous cycle. Specifically, reward is given in three circumstances: (a) if shock was predicted and occurred (a large positive reward), (b) if absence of shock was predicted and shock did not occur (a lesser positive reward), and (c) if absence of shock was predicted but shock occurred (a negative reward, i.e., a punishment). Otherwise, the reward is zero. The reward, if any, is added to the strength values of the relevant rules. The comparison is also used to trigger the generation of new rules (see ahead).

The message representing the current environment is compared with the message representing the previous environmental state to determine whether an unusual event has occurred. If a feature-onset occurs, a check of the long-term feature list is made to determine if the feature is unfamiliar, in which case the current event is defined as unusual and entered in the short-term event store.

Rule matching and response selection. Conditions of all rules are then matched against the message describing the current

environment. Each matched rule posts a bid, which is a proportion of its strength. That is, Bid b made by Rule i will be

$$b_i = k * s_i, \quad (2)$$

where k is a constant between 0 and 1. Rules that make the same prediction sum their bids and act together as a set.

The rules that will govern the system's response are then selected in accord with Principles 1 and 2, discussed earlier. Principle 1, the assumption of a probabilistic relation between relative strengths of competing rules and response selection, is realized by applying a simple version of the Luce (1963) choice model to the strengths of matched rules. Principle 2, the assumption that default rules are protected by exception rules, is realized by allowing exception rules to censor their corresponding default rules. When exception rules censor a default, the exception rules substitute for the default on that cycle. A detailed description of the response selection process is provided in the Appendix. The rules that are selected to determine the system's response on a cycle will be termed the *winning set*, W . The action called for by the winning set is performed, creating a new predicted message, and the indicated effector action is taken. The next cycle begins.

Strength revision. Strength revision takes place in two steps. In accord with Principle 3, discussed earlier, only the rules in W have their strengths changed. The rules in W effectively compete for reward, as called for by Principle 4, in accord with the following scheme. First, when the winning set is selected, each rule in the set has its strength reduced by an equal portion of the summed bid made by the rules in W , $\sum_{j=1}^n b_j$, where n is the number of rules in W . Second, when the reward, R , is assigned to the winning set on the subsequent cycle, an equal portion of R (i.e., R/n) is added to the strength of each rule in W . The net change in the strength of Rule i , then, is

$$\Delta s_i = (R - \sum_{j=1}^n b_j)/n \text{ for } i \in W, 0 \text{ otherwise.} \quad (3)$$

If a rule has its strength reduced to zero, it is eliminated from further competition and is never revived.

Properties of bidding and strength revision. The mechanisms of bidding and strength revision have several properties that deserve emphasis. A rule becomes *at risk* whenever it is a member of the winning set, W , and determines the system's behavior. The immediate consequence is a loss of strength proportional to the sum of bids. Intuitively, a winning rule is *betting* that it will garner sufficient reward by controlling the system to offset the cost it incurs to earn that opportunity. If the reward it receives is in fact larger than its bid payment, the rule will gain strength; if the reward is less, the rule will lose strength. As is apparent from Equation 3, the asymptotic strength of a rule will be the value at which the cost associated with the bid is equal to the expected reward level. The value of R together with the bid constant, k , implicitly determine the asymptotic strength of rules that make the relevant prediction.

A rule that is not in W will neither gain nor lose strength. One consequence is that an erroneous rule will tend to lose strength until it is significantly weaker than its more valid competitors; as it weakens, it will win less often and hence lose

strength more slowly. In general, the probability that a rule will govern behavior is not a function simply of its absolute strength value, but of its strength relative to rules with similar conditions that specify contrary actions. Also note that a default rule that is matched but then censored by an appropriate exception rule will not lose strength on an occasion when it might otherwise have acted and failed. This property of default hierarchies provides a natural account of the role of inhibitory cues in protecting paired excitatory cues from extinction.

In the simple case in which only one rule is in the winning set, the rule will pay out its own bid, a proportion k of its strength, and receive all the reward for a successful prediction. If multiple rules make the same prediction, they will compete for the fixed reward. Because R is divided equally, each rule in W receives a lesser strength increment than if it had acted alone. However, because each rule also pays an equal portion of the sum of bids (i.e., an amount equal to the mean bid), weaker rules will be at a relative disadvantage. A rule that bid less than the mean bid of rules in W will pay more than it would have had it acted alone; a rule that bid more than the mean will pay less. The net effect is that although all rules gain less strength if other rules are in W than if not, the weaker rules in W suffer more from the competition.

Three distinct competitive relations between rules can arise in our model. First, if two or more rules are matched and specify the same action, they will operate together to control behavior (by summing their bids), but then compete for reward if they become the winning set. Second, if two or more rules are matched, specify contrary actions, and are not in a default/exception relation, then they compete to control behavior; the outcome is determined by the relative strengths of the opposing rules, and only those that become the winning set undergo strength revision. Third, if one rule is an exception to another, they will compete to control behavior. Because of the censoring mechanism, however, the exception can substitute for the default. The exception rule thus has a special advantage over the default in the competition. If the exception rule is appropriate, its effect is actually to protect the default from strength reduction.

Rule strength in our model is analogous to associative strength, and the bidding and reward system produce the three most basic qualitative aspects of Equation 1 of the Rescorla-Wagner model. First, the asymptotic strength of a rule is limited by the magnitude of the reinforcer, which determines R . Second, the expected increase in strength due to receipt of reward will decline as a rule's strength approaches asymptote because the greater the strength of a rule the larger will be its bid, and hence the expected cost associated with winning the bidding competition. Third, the expected increase in strength for a given rule will be reduced if other rules operate concurrently. As we will see later, however, our model has a number of important empirical consequences that differ from those of the Rescorla-Wagner formulation. The relation of Equation 3 to the corresponding Equation 1 of the Rescorla-Wagner model is described in the Appendix.

Rule generation. The computer program contains three inferential heuristics for generating new rules, triggered by particular states of the system as specified by Principle 5. All three heuristics are specific instantiations of the unusualness heuris-

tic (Principle 6). As implemented, the heuristics are tailored for the prediction of shock or its unexpected nonoccurrence; however, the general theory applies to any US.

1. *Covariation detection.* If a shock occurs unexpectedly and is preceded by or concurrent with an unusual event stored in the short-term buffer, then a new rule will be constructed. The new rule will include the unfamiliar feature of the unusual event in its condition and will have an action specifying expectation of a shock and crouching. For example, if an unfamiliar tone begins prior to an unexpected shock, the rule $T \Rightarrow S$ will be generated. If no unusual event is stored and no other heuristic applies, then with some probability less than 1, a general rule is constructed with a maximally general condition and the same action as discussed earlier. Thus, if an unexpected shock occurs and no unusual event is stored, the rule $\# \Rightarrow S$ may be generated.

2. *Exception formation.* If a strong rule makes an erroneous prediction about the presence or absence of a shock, and an unusual event occurred prior to or concurrent with the prior cycle (when the failed rule was matched), then exception rules are formed by (a) adding the unusual feature to the condition of the failed rule and substituting the appropriate action and (b) using the unusual features alone to form the condition. The failed rule is preserved as a default, tagged with the newly created exception rules. For example, suppose a tone occurs paired with an unfamiliar light and the strong rule $T \Rightarrow S$ is a member of the winning set and creates an expectation of shock, which fails to occur. The rules $L + T \Rightarrow \bar{S}$ and $L \Rightarrow \bar{S}$ will then be generated. The former exception rule corresponds to the hypothesis that the unusual cue (L) signals nonoccurrence of shock only in the presence of the known predictor (T); the latter exception rule captures the more general possibility that the unusual cue might signal absence of shock regardless of whether the known predictor occurs.

3. *Chaining.* If an unusual event is stored in the buffer and a strong rule is included in the current winning set, then a new rule may be formed that uses the unfamiliar features of the unusual event in the condition and that has the same action as the parent rule. The initial strength is set to a proportion of the strength of the strong rule from which the new rule was constructed. Chaining tacitly seeks earlier predictors of the US; hence, the strength of the new rule is set higher if the onset of the unusual event preceded the old CS (that is, preceded the cycle on which the parent rule was matched) than if it was concurrent with it. As an example, suppose a tone occurs concurrently with an unfamiliar light and the strong rule $T \Rightarrow S$ is a member of the winning set. Then, the rule $L \Rightarrow S$ may be generated (at the lesser initial strength value).

In some situations, multiple heuristics will apply, in which case several new rules, sometimes contradictory to each other, may be formed. For example, suppose the strong rule $T \Rightarrow S$ is in the system and an unfamiliar light is now presented in a compound with the tone. When the light is paired with the tone and the existing rule $T \Rightarrow S$ is selected, the rule $L \Rightarrow S$ will be generated by chaining. On the next cycle, if the expected shock does not occur, exception formation will generate the rules $L + T \Rightarrow \bar{S}$ and $L \Rightarrow \bar{S}$. As we will see, the resulting rules will enter into complex interactions as the bidding and strength-revision mechanisms operate over subsequent trials to adjust the strengths of all of the rules in the system. Also note that the

introduction of new rules does not automatically result in the displacement of old ones. New rules may drive out old ones (or vice versa), strengthen an old rule (when a new exception rule prevents further losses that would otherwise accrue to the default), or reach a stable state of coexistence with the old.

The three heuristics serve related but distinct functions. Covariation detection provides initial rules to explain unexpected occurrences, whereas exception formation and chaining build on existing partial knowledge. Exception formation creates exception rules that may censor (and hence protect from further strength reduction) strong default rules that err under identifiable circumstances. Exception formation is thus a reaction to an erroneous prediction. In contrast, chaining represents an opportunistic attempt to identify an earlier predictor of a US that is already predicted by a CS. Chaining provides a simple form of second-order conditioning (i.e., a neutral cue paired with a conditioned CS comes to elicit the US).

Relations to other computational models of learning. Our theory of conditioning has some similarities to other computational models of learning that have been proposed, although the differences are extensive. The two most closely related systems are J. H. Holland's (1986) classifier systems and Anderson's (1983) ACT*, both of which are rule-based systems with strength revision. Classifier systems provided the notation for rules and messages in the present model as well as the general notions of parallel rule activity and bidding. However, the present model introduces a summation mechanism in bidding, uses an entirely different set of heuristics for rule generation, and has a somewhat different strength-revision mechanism. Our model uses neither genetic operators nor the bucket-brigade algorithm, which are central to classifier systems (J. H. Holland, 1975, 1986). Later in the article we will discuss an extension of our model that would include the latter learning mechanism.

Exception formation is similar to the process of action discrimination postulated in Anderson's (1983) ACT* theory, and chaining very loosely resembles the process of composition in ACT*. Our system differs from ACT* in that our model uses a vastly simpler representational scheme, eliminates partial matching of rules, applies rules in parallel via summation of bids, introduces the unusualness heuristic, and has a very different strength-revision mechanism.

Both ACT* and classifier systems provide more specific rules with a competitive advantage over more general rules when both are matched; however, neither includes the censoring mechanism proposed here. Our default/exception representation provides the logical equivalent of *if-then-unless* rules, and can be viewed as a probabilistic version of the censor rules described by Winston (1986), from whom we adopt the term *censor*.

Finally, our use of parallel rule activity and of strength revision give our model some commonality with connectionist networks (Rumelhart, McClelland, & the PDP Research Group, 1986). The summation mechanism used in our bidding process is related to neural-like mechanisms in which multiple inputs to units are summed to determine their outputs. Unlike connectionist models, however, our theory introduces explicit default hierarchies and heuristics for the quantum formation of new rules.

Applications of the Theory to Conditioning Phenomena

We will now apply the theory to a series of results from studies of classical conditioning, emphasizing results obtained using the CER paradigm. Our intent is to demonstrate that the model can account for a wide range of phenomena that include simple excitatory conditioning, blocking and overshadowing, sensitivity to statistical predictability of the US, and conditioned inhibition.

To increase the rigor of these tests of the model, we placed several demands on the simulations. First, unlike most previous formal models of conditioning, our simulation generates predictions of the probability of a CER, rather than solely predicting hypothetical strength values. The simulations thus test a learning model, based on rule generation and strength revision, in the context of a decision model (the bidding system) that predicts overt responses. Second, the number of free parameters was reduced by defining some parameters in terms of others, and then allowing only two parameters to vary. These two parameters corresponded to the reward for correctly predicting shock, and the proportion of strength used as a bid (which determines the rate of conditioning). Given that the data to be simulated came from studies from different laboratories and involved numerous procedural variations, small variations in these two basic parameters seem justified. (In fact, extremely similar results were obtained in a set of runs in which only one parameter was varied.) The other parameters were held constant across the entire set of simulations. For any formal theory that pretends to some generality, it is necessary to assess whether it can account for different but related phenomena with a consistent set of parameters. This constraint is particularly important for models like ours that are composed of multiple interacting processes. The parameter values used for each application of the simulation are provided in the Appendix. In addition to describing specific simulation results, we will provide qualitative accounts of a number of additional conditioning phenomena.

Let us first consider the simplest conditioning paradigm, in which a one-element CS is paired with a US. We assume that the animal, prior to receiving an aversive US (assumed to be shock in all our examples), will respond in accord with some background rule that makes no prediction about the US, and specifies some effector action appropriate in its absence. As a simple concrete case, we will refer to this initial rule, which makes no prediction about shock, as $\# \Rightarrow \text{Press}$.

Now, suppose an unfamiliar tone sounds shortly before a shock is administered. The occurrence of the tone, as an unusual event, will lead to generation of the rule $T \Rightarrow S$. Let us suppose that the tone continues to be paired invariably with the shock, which never occurs in the absence of the tone. Over trials, the new CS rule, $T \Rightarrow S$, will compete with the preexisting rule $\# \Rightarrow \text{Press}$.

In this simple case of competition between two rules that call for contrary actions, the winning set will always contain just one rule, the censoring mechanism will not apply, and reward will not have to be divided among multiple rules. Both rules will be matched whenever a tone occurs; if the former wins control of effector selection, it will be rewarded and hence gain strength; if the latter wins control, it will fail to receive reward

and hence lose strength. Over trials, the CS rule will increasingly dominate in the competition, leading to an increased probability of a CER in the presence of a tone. The preexisting rule will continue to control behavior in the absence of a tone, when it will be the only rule to be matched and will receive reward.

Conditioning functions typically reveal an initial phase of positive acceleration in responding, followed by a subsequent phase of negative acceleration to the asymptote (Mackintosh, 1974). Our model leads to the expectation that this should be the common pattern. Equation 3 predicts that when a useful new rule is formed, the growth of its strength will in fact often show an initial period of positive acceleration. The expected change in strength dictated by Equation 3 is based on a mixture of trials in which the rule gains control of response—and hence has its strength adjusted—and trials in which it fails to gain control—and hence does not have its strength adjusted. If a new rule is formed at a strength substantially lower than that of competing rules that are matched at the same time, it will at first often lose the competition and hence fail to gain strength. As it is tested and proves useful, the resulting strength increments will cause the new rule to win more often and therefore more often receive a further increment, until it reaches its asymptotic strength. Note that the magnitudes of successive strength increments that accrue *when the rule is tested* form a strictly negatively accelerated function, as does the function based on the Luce (1963) rule, which relates increases in strength to response probability (see Appendix). As the strength of the new rule increases, these factors will soon outweigh the positive acceleration contributed by the increasing frequency with which occasions for strength adjustment arise. Accordingly, the initial period of positive acceleration will be typically relatively brief and will be followed by a longer phase of negative acceleration to the asymptote.

The simulation incorporates a number of mechanisms that influence the conditionability of the CS. In accord with the forward-causation heuristic, the heuristics for rule generation favor conditioning to a CS with an onset prior to, rather than concurrent with, the US. Temporal contiguity also has an impact, as the short-term buffer for unusual events imposes an upper bound on the temporal gap between the CS and US that will allow learning. If the interval is too long, the representation of the CS will be lost before the onset of the US, and it will not be possible to form a rule linking them.

The implementation of the variants of the unusualness heuristic provides an approximate account of the effect of CS pre-exposure. If a feature has been observed a sufficient number of times, its onset will no longer trigger storage of the event in the short-term buffer, and the representation of the CS will not be available for use by heuristics in generating rules to predict the US (cf. Wagner, 1978).

Blocking and Overshadowing With Multielement CSs

Blocking. The simulation was applied to the design of the original study by Kamin (1968) that demonstrated blocking, a phenomenon providing clear evidence that conditioning does not depend simply on contiguous pairings of the CS and US. The exact design is outlined in Table 1, along with the results

obtained in Kamin's study and in our simulation. As is customary in studies using this type of CER paradigm, the dependent measure is a *suppression ratio*, calculated in this study by dividing the bar pressing during a 3-min period of CS presentation by the sum of the bar pressing during the CS period plus the immediately preceding 3-min period when the CS was absent. With this measure, complete suppression would be indicated by a ratio of 0, whereas no suppression would be indicated by a ratio of .5. Kamin ran two versions of the study that counterbalanced the use of two cues, a noise and a light, as the blocked and the blocking stimulus. We will describe the version in which the noise was the blocked stimulus.

Animals in four groups received a series of acquisition trials, involving varying numbers of pairings of shock with a light (L) and a simultaneous light-noise (LN) compound. In all groups, four trials occurred each day. On the final day, all of the groups received four nonreinforced test trials with a noise (N). The major comparison involves Groups 2 and 3. Group 2 received 16 reinforced L trials and then 8 reinforced LN trials. The subsequent test on nonreinforced N trials revealed no suppression—conditioning to the noise was blocked. In contrast, Group 3, which received reinforced LN trials without prior L trials, showed substantial suppression to the noise during the test.

The explanation of this blocking effect in terms of our theory is as follows. In Group 2, during the initial reinforced light trials, the rule $L \Rightarrow S$ will be formed and strengthened to near asymptote. When the compound LN is presented, the strong light rule will successfully predict shock, so covariation detection may never be triggered. On the other hand, the chaining heuristic will likely form the rule $N \Rightarrow S$ because the noise is paired with the light, which matches a strong rule. The new noise rule will receive relatively low strength, however, because the noise is concurrent with, rather than prior to, the light, and the noise rule will be unable to gain further strength in competition with the near-asymptotic light rule with which it will prove redundant; indeed, the noise rule may lose its initial strength. The result will be blocking of conditioning to the light.

For Group 3, with no prior experience with either cue, pairing of the LN compound with shock will trigger generation of three rules by covariation detection: $L \Rightarrow S$, $N \Rightarrow S$, and $L + N \Rightarrow S$. These three rules will divide the initial allotted starting strength. Because they always will be matched together, no rule will be selectively favored, and all three will be maintained. Thus, the noise cue will yield some suppression on the later test.

The fact that Group 1 also shows suppression to the noise indicates that it is crucial that the separate presentations of one of the cues precede rather than follow the compound presentations in order to obtain blocking. This, of course, is what our theory predicts. Group 4 is simply a control, demonstrating that the noise cue does not yield suppression if the LN compound has never been presented.

Kamin's (1968) results differ from those of the simulation in that he obtained less suppression in Group 1 than in Group 3. It seems likely that animals in Group 1 showed some forgetting of the blocked rule because 4 days had elapsed between presentation of the LN compound and the test. (Kamin reported that other evidence ruled out a retroactive-interference interpretation of this difference.) Our simulation does not include a pro-

Table 1
Simulation of the Blocking Experiment by Kamin (1968)

Group	Experimental design		Simulation	Kamin data	
				X = noise	X = light
1	AX + (8)	A + (16) test X-	.22	.36	.25
1	A + (16)	AX + (8) test X-	.50	.50	.45
3	—	AX + (8) test X-	.21	.25	.05
4	—	A + (24) test X-	.50	.49	.44

Note. The simulation data plotted are median suppression ratios to four nonreinforced presentations of the test stimulus, based on 50 simulated subjects in each condition. The two sets of data from Kamin (1968) differ as to which of two cues, a noise or a light, served as the X cue.

cess of passive forgetting. Kamin's results also indicate that the light was a more potent cue than was the noise when the roles of the two cues were reversed. The simulation does not model differences in salience of stimuli.

Our theory predicts that a rule relating the blocked cue to the US is very likely to be generated either by covariation detection (if the prior rule is only moderately strong) or by chaining (if the prior rule is strong enough to produce second-order conditioning). However, the rules based on blocked cues will acquire little if any additional strength because of their redundancy with the stronger initial rule. This predicted pattern is consistent with empirical studies that have observed conditioning to the blocked cue on the first presentation of the compound (when new rules are formed), but little if any additional conditioning on later trials (e.g., Kamin, 1968; Mackintosh, 1975a).

In general, although a blocked cue is disadvantaged in the competition for a limited niche as predictor of the US, it is not simply ignored. In our theory, the blocked cue will give rise to weak rules and also will have its familiarity incremented because of its occurrences within the compound (James & Wagner, 1980). As a result, a blocked cue will show loss of conditionability if it is later presented alone in conjunction with a US, as Pearce and Hall (1979) have demonstrated. The cue will have become familiar and hence fail to satisfy the unusualness criterion for use in forming new rules. The Rescorla-Wagner model is unable to account for the decrease in future conditionability of a blocked cue.

Overshadowing and partial transfer. Our theory, given the centrality it ascribes to the discounting principle that operates whenever multiple rules act together, can also account for one-trial overshadowing of the sort described earlier, in which a cue produces less conditioning when presented once in a compound than when presented once alone. When a compound cue with multiple unusual components is used to form rules predicting the US, the initial strength is divided among the resulting multiple rules. Thus, each cue will receive less initial strength as a predictor of the US than it would have had it been the only one available for use in rule formation.

For similar reasons, the model also correctly predicts that if an animal is initially conditioned to a compound AB cue, it will subsequently exhibit a reduced but significant degree of conditioning to either of the constituent cues, A or B, presented separately. The compound cue will give rise to three rules, based on cues A, B, and A + B, respectively, that will share the strength

supported by the reward provided on trials with the compound. If the rat is now tested with one of the cues alone, only one of the three learned rules will be matched, yielding a lesser level of suppression than observed in the presence of the compound.

Our theory also deals in a natural way with the variant of overshadowing in which a stimulus that predicts the occurrence of the US to a moderate degree shows little conditioning when it is paired with a stimulus that, by itself, predicts the US quite well. An illustrative study is one by Wagner, Logan, Haberlandt, and Price (1968). These investigators presented cue X simultaneously with cue Y one half of the time and with cue Z one half of the time. For Group 1, the compound XY was always reinforced with shock, and the compound XZ was never reinforced; for Group 2, the compound XY was reinforced 50% of the time, and so was compound XZ. Thus, in both groups, cue X was paired with shock 50% of the time. However, for Group 1, cue X was unhelpful relative to cue Y as a predictor of shock. Although it was present on one half of the occasions of reinforcement, another cue, Y, was present on all of the occasions and thus was a much better predictor. For Group 2, on the other hand, Y was no better a predictor than X (or Z). Thus, we would expect more conditioning to X in Group 2 than in Group 1, which was in fact the result obtained.

In terms of our theory, covariation detection will generate several rules in both Group 1 and Group 2 that will proceed to compete with one another for reward as shock predictors. Animals in Group 1 would be likely to form the rules $Y \Rightarrow S$ and $X \Rightarrow S$; those in Group 2 would also form the rule $Z \Rightarrow S$. (For simplicity we will ignore the possible rules with conjunctive conditions.) In Group 1, the Y rule was always confirmed, whereas the X rule was confirmed half as often. Moreover, the occasions of confirmation were arranged so that when it was successful, the X rule was always redundant with the Y rule. Whenever the X rule alone was matched (that is, when an XZ compound was presented), it failed and was driven toward a zero asymptote; accordingly, it suffered relative to the Y rule when the two rules jointly formed the winning set. The strength-revision procedures would thus strongly favor the Y rule over the X rule. In Group 2, by contrast, all three rules were confirmed half the time, so that no one rule would have been favored to drive out any of the others. Consequently, there was more conditioning to X in Group 2 than in Group 1.

Statistical Predictability and Conditioning

We now present a simulation of a study of excitatory conditioning that provided clear evidence that conditioning depends

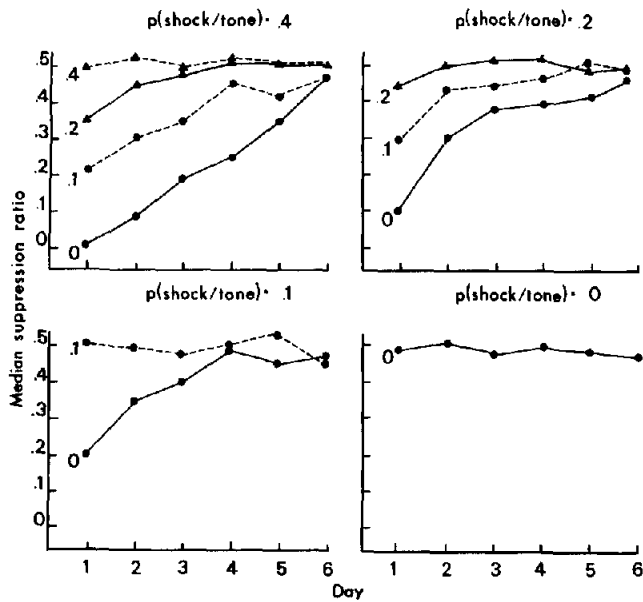


Figure 2. Median suppression ratios for each group over six extinction test sessions. (Within each panel, all groups had the same probability of the shock [unconditioned stimulus, US] during the tone [conditioned stimulus, CS]; the parameter in each panel is the probability of the US in the absence of the CS. Data are from Rescorla, 1968.)

not on simple pairing of the CS with the US, but rather on the extent to which the CS is a statistically useful predictor of the US. Rescorla (1968) placed hungry rats in a box with a bar that could be pressed to obtain food. At various times, a brief electric shock was administered, resulting in an immediate suppression of any bar pressing. Rescorla divided his animals into several groups. For some groups, shock was always presented—and only presented—during 2-min periods in which a tone sounded. For the other groups, shocks also occurred during the intervals in which the tone was absent. After this initial training period, all of the groups underwent a set of extinction trials in which shock was no longer presented at all. Figure 2 shows changes in bar pressing, calculated as suppression ratios, during the extinction sessions.

The detailed results of the Rescorla (1968) experiment displayed in Figure 2 are as follows. Each panel of Figure 2 shows suppression ratios under a different contingency between the tone and the shock during learning trials. The top-left panel shows results when there was a .4 probability of the shock in the presence of the tone (i.e., $p(\text{shock}/\text{tone}) = .4$), the top-right panel gives results for a .2 probability, the bottom-left corresponds to a .1 probability, and the bottom-right corresponds to a 0 probability. The parameter associated with each curve in a given panel (0, .1, .2, or .4) is the probability of the shock in the absence of the tone (i.e., $p(\text{shock}/\text{no tone})$).

As the data in the upper-left panel of Figure 2 indicate, suppression was complete at the beginning of extinction trials when the tone predicted shock with high probability and the absence of the tone predicted absence of shock with certainty (.4-0). Suppression was nil when the presence of the tone predicted shock with high probability, but so did its absence (.4-.4). Sup-

pression was intermediate when the absence of the tone was sometimes associated with shock (.4-.2 and .4-.1). Comparable results, but with less initial suppression, were obtained when the tone predicted the shock with probability .2 or .1. In particular, note that the top line in each panel is approximately flat at a suppression ratio of .5, indicating that suppression was always nil when the probability of shock did not differ during learning as a function of whether or not the tone was presented. Thus, sheer number of pairings was irrelevant to the degree of conditioning obtained. The key variable is clearly degree of predictability in a statistical sense.

Figure 3 presents the results of our simulation of Rescorla's (1968) experiment. (For reasons of programming convenience, slightly different conditional probabilities were used in the simulation.) Details regarding parameter values for this and other simulation results are provided in the Appendix; here, we will describe in qualitative terms the mechanisms responsible for the fit.

The simulation begins with only the initial background rule, # \Rightarrow Press (Rule 1); other rules enter the system as they are generated by covariation detection in response to simulated events. When the tone occurs together with shock, covariation detection will apply to generate the rule T \Rightarrow S (Rule 2). In addition, animals in groups that receive shocks in the absence of the tone will eventually, for lack of any unusual event that might account for these occurrences, form the rule # \Rightarrow S (Rule 3) by covariation detection.

The asymptotic suppression levels attained during training result from the competitive interactions among Rules 1, 2, and 3. The more valid the tone is a predictor of shock, the stronger Rule 2 will become relative to Rule 3, and hence the greater will be the asymptotic suppression at the end of training. Note that whenever the probability of shock does not vary as a function of the presence of the CS, Rule 2 will fail completely in competi-

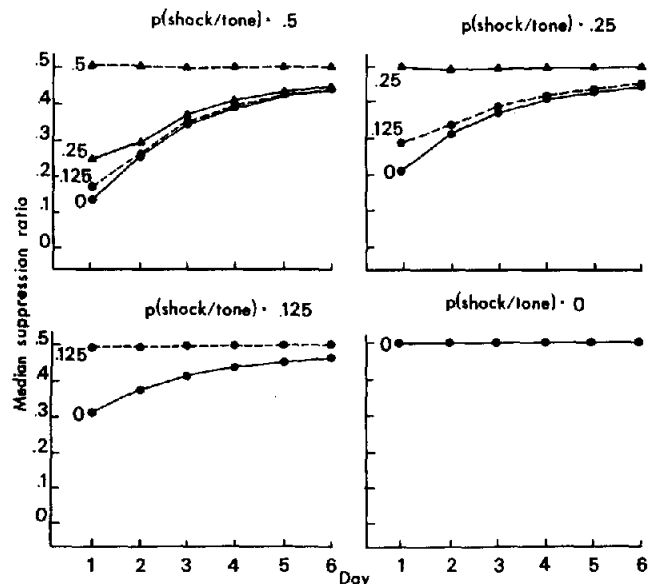


Figure 3. Simulation of the results of Rescorla (1968), depicted in Figure 2. (Data are based on 12 simulated subjects in each condition.)

tion with Rule 3. In general, the theory predicts that the asymptotic strength of a rule that is redundant with a more general rule that sometimes applies alone is zero. Accordingly, the top line in each panel of Figures 2 and 3 begins at about .5 (zero suppression).

In all conditions, as extinction proceeds, neither Rule 2 nor Rule 3 is ever correct (because there are no shocks). Each steadily loses whatever strength it acquired during training, with the consequence that Rule 1 wins ever more frequently. Ultimately, Rule 1 wins almost all of the time, regardless of presence or absence of the CS; accordingly, bar pressing approaches equal frequency in both intervals and the suppression ratio goes to .5. This general pattern is apparent in Figures 2 and 3.

Although the simulation results agree with the data obtained by Rescorla (1968) in general qualitative aspects, the fit is only approximate. In particular, the curves generated by the simulation are less clearly separated than are the actual data curves. The fit could be improved by applying a nonlinear transformation to strength values; however, we did not feel that more precise curve fitting would be useful at this point.

An important implication of our account is that, during training, animals should show an initial degree of suppression in response to the tone during the period in which Rule 2 is generated and remains for a time in competition with Rule 3, even if the tone does not actually predict shock. That is, a rule that uses an unusual event to predict the US should be generated, even though it later proves to be redundant with a more general rule and is therefore eliminated in the competition for reward.

This prediction is supported by data reported by Rescorla (1972b). He examined the suppression ratio during initial training on uncorrelated conditions (e.g., .4 p (shock/tone), .4 p (shock/no tone)). The results indicated that the rats did indeed initially entertain Rule 2. Animals that had received no shocks showed no suppression in response to the tone. However, after the first session, which included 12 random tone occurrences and several uncorrelated shock occurrences, there was substantial suppression to the tone. At this point, covariation detection would have generated the rule $T \Rightarrow S$. After several sessions, however, Rule 2 was apparently vanquished; there was no suppression to the tone at all. Thus, a variant of the unusualness heuristic, "unexpected events signal other unexpected events," generated a rule that was then tested and ultimately found to be redundant. Such results are consistent with the most basic distinction that our framework makes between mechanisms for induction, namely, rule generation and rule assessment.

Conditioned Inhibition

We will now consider some of the most complex phenomena that have been observed in CER studies, namely, those associated with conditioned inhibition. In experiments on conditioned inhibition, a cue A is first conditioned as a predictor of the US. In a subsequent phase, cue A is paired with cue B and presented without the US. Animals will learn that cue B is a safety signal and will inhibit the usual fear response to cue A when B is present. The conditioned inhibition paradigm, like the blocking paradigm, involves presentations of multi-element cues. Unlike the blocking case, in which the cues are redundant

with each other, each cue in a conditioned-inhibition paradigm conveys important information that the animal must learn in order to behave in an adaptive manner.

Several experiments establish that when a new cue is associated with an established excitatory CS, this does not result automatically in the new cue taking on the fear-evoking properties of the CS. Indeed, other experiments go farther and show that, under some conditions, the new cue can take on properties that are opposite to those of the established CS. In general, a new stimulus can have excitatory effects (that is, effects similar to the US), no effects, or inhibitory effects (that is, effects that counter those of the US), depending on the expectations generated by the pattern of associations in a given time period.

Initial excitation followed by inhibition. An experiment by Rescorla (1972b) establishes the point that a given cue may serve initially as an excitatory agent and then as an inhibitor. On the first two days of conditioning, rats were presented with four 10-s flashing lights, each of which ended with a $\frac{1}{2}$ -s shock. At this point, the rule $L \Rightarrow S$ should be established. For the next 15 days, the rats received one trial with the 10-s flashing light ending in shock and three trials with a 40-s tone, the last 10 s of which was accompanied by the light. The second kind of trial was never accompanied by shock.

This second phase of the procedure, according to our model, will trigger generation of three new rules. Chaining will generate the rule $T \Rightarrow S$, which conveys the hypothesis that occurrence of a tone may be an earlier predictor of the shock known to be signaled by the light. In addition, because the expectation of a shock generated by the strong light rule fails after the unusual occurrence of a tone, the rules $T + L \Rightarrow \bar{S}$ and $T \Rightarrow \bar{S}$ will be generated by exception formation. The first of these new rules creates an expectation of shock when the tone occurs and, hence, has an excitatory influence; the latter two rules suggest that tone signals absence of shock and, hence, have an inhibitory influence.

The observed performance pattern over trials will reflect the competition among the aforementioned rules. During the periods when the tone is being presented alone prior to the light, the competition will center on the contradictory rules $T \Rightarrow S$ and $T \Rightarrow \bar{S}$. In the model, the former rule will receive a substantially greater initial strength, because when chaining generates a new rule on the basis of an unusual event occurring prior to an established CS, the new rule begins with a strength equal to a high proportion of that of its parent. Accordingly, the simulation predicts that the tone will initially act as an excitatory stimulus. However, the resulting expectation of shock will never be confirmed, and hence the rule $T \Rightarrow S$ will eventually fail in competition with the valid rule $T \Rightarrow \bar{S}$, so that the tone will become an inhibitor.

Figure 4 presents the data obtained by Rescorla (1972b), and Figure 5 presents comparable data produced by the simulation. As may be seen in Figure 4, initial suppression in the presence of the tone, indicating an initial excitatory effect, was indeed found, followed by a period of increasing inhibition. Figure 4 also shows the remarkable speed of the initial suppression effect. The average suppression ratio on the very first three trials (on the first day) was .23. By the second day, the rule $T \Rightarrow S$ is apparently well established as a hypothesis: The suppression ratio on the second day is lower than on the first. Eventually, how-

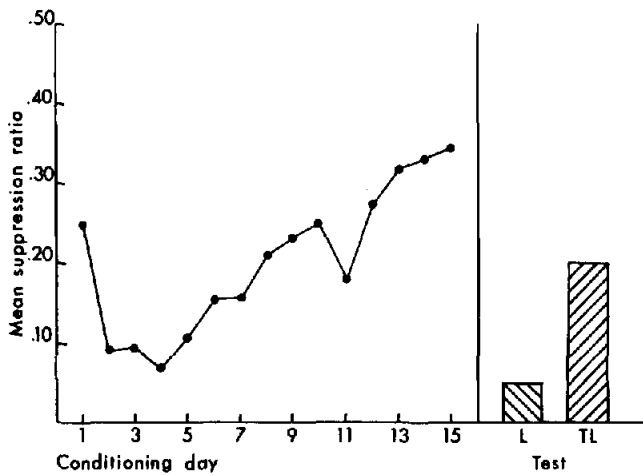


Figure 4. Median suppression ratios for each group over a series of extinction test sessions. (The panel to the left shows suppression to the tone, T, over the course of a TL-, L+ procedure, in which the tone proceeds the light, L, on the compound trials. The right-hand panel shows suppression to the light when presented alone and when preceded by the tone on a test session administered after Day 15. Data are from Rescorla, 1972b.)

ever, the erroneous rule loses strength, and the tone becomes an inhibitory cue.

The data presented in the far-right portion of Figure 4 indicate what is learned during the periods when the light occurs with the tone, without shock following. The exception rule $L + T \Rightarrow \bar{S}$, formed by exception formation, will gain strength and censor its parent rule, $L \Rightarrow S$. At the end of the 15 conditioning trials depicted on the left of Figure 4, the rats were given a series of three test sessions during which the light was presented alone on two trials and with the tone on four other trials. It is clear that there is much greater suppression to the light alone than to the light-plus-tone compound. Indeed, the suppression to the light alone is nearly asymptotic. This is as it should be, given that the light in the absence of the tone continued to signal shock over the entire training period, sustaining the original rule, $L \Rightarrow S$. Also, when the light was presented with the tone and without shock, it was at least partially protected from strength reduction by the censoring mechanism because of the exception rules. The dominance of the exception rules is evidenced by the substantially reduced suppression, relative to the light alone, that was observed when the compound stimulus was presented.

Rapid inhibitory conditioning. Not all studies of conditioned inhibition find that the cue paired with an excitatory CS is initially excitatory. In some circumstances a novel stimulus superimposed on a CS that produces a fear response will immediately inhibit the CER. The Rescorla-Wagner model cannot account for this type of finding, as Rescorla and Wagner (1972, p. 79) acknowledged, except by assuming that stimuli are normally inhibitory. In fact, however, stimuli are normally initially excitatory.

An experiment by Kamin (1968) provides a dramatic demonstration of rapid inhibitory conditioning. Kamin trained rats for 16 trials to associate white noise with shock. He then created

two different groups. Group LN received eight trials of a compound, simultaneous light-plus-noise stimulus that was never reinforced by shock, followed by four trials of the original noise stimulus that was again nonreinforced. Group N simply received 12 standard extinction trials, during which the noise was presented but never with shock. All animals received four trials per day.

Let us analyze the predictions our model makes for this study, on the basis of rule generation and subsequent competition. Group N is, of course, expected to show just the customary gradual extinction as the rule $N \Rightarrow S$ dies a slow death because of nonreinforcement. The situation is much more complex for Group LN. Given that a trial in Kamin's (1968) experiment spanned a fairly long interval (several minutes), this time period would correspond to several cycles of matching and firing rules. Because no shock was presented on the first trial, in which the light occurred along with the noise, the strong rule $N \Rightarrow S$ would repeatedly fail. Given the availability of an unusual event—the occurrence of the light—heuristics for rule generation will be triggered. In particular, exception formation should on the initial extinction trial generate new exception rules, $L + N \Rightarrow \bar{S}$ and $L \Rightarrow \bar{S}$. Both new rules will have an immediate inhibitory influence.

When the light is first paired with the excitatory noise, chaining may create the excitatory rule $L \Rightarrow S$, which will compete with the other new rules. However, the new cue we are considering here differs from the comparable cue in the Rescorla (1972b) study discussed earlier; in that in Kamin's (1968) experiment, the new cue does not occur prior to the original CS; accordingly, the initial strength of the rule generated by chaining will be low. Consequently, the influence of the new inhibitory rules will outweigh the influence of the excitatory ones, so that Group LN might be expected to show some inhibition of suppression even on the very first trial. Furthermore, the new inhibitory rules will of course be confirmed, and so the animal should show continued development of inhibition over trials.

What should happen when, after the first eight trials, the noise alone is presented? For Group N, nothing interesting. This is merely a continuation of the slow competition between

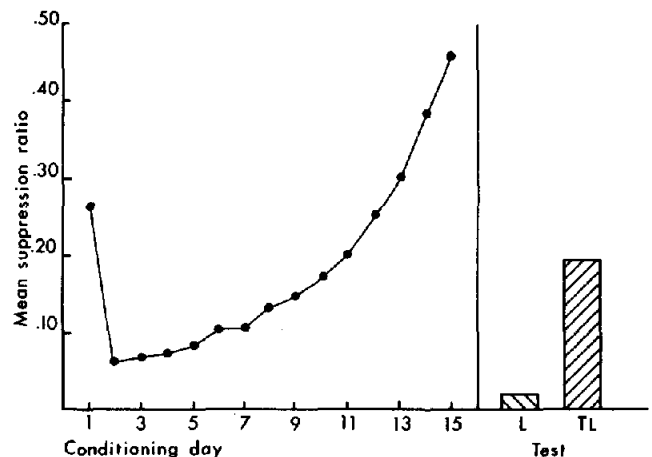


Figure 5. Simulation of the results of Rescorla (1972b), depicted in Figure 4. (Data are based on 50 simulated subjects in each condition.)

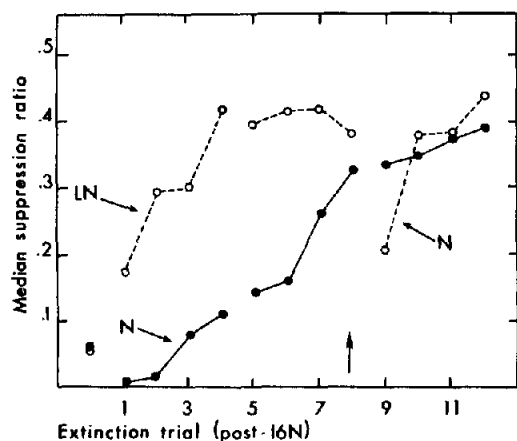


Figure 6. Extinction of suppression, by trial, after 16 sessions of conditioning to noise (N). (The left-most points represent suppression on the last day of training. The groups were extinguished either to noise alone or to a light-plus-noise, LN, compound. The arrow in the abscissa indicates the point at which group extinguished to compound is switched to noise alone. Data are from Kamin, 1968.)

the rule $N \Rightarrow S$ and its original competitor $\# \Rightarrow \text{Press}$. For Group LN, however, we expect a reversion to substantial suppression effects, because for rats in this condition the rule $N \Rightarrow S$ will have been protected to some extent because of the censoring effect of the successful exception rules.

Figures 6 and 7 present the results from the actual experiment and the simulation, respectively. The results for Group N, presented with the noise alone, may be seen at the bottom of Figures 6 and 7. These animals showed the customary slow-extinction process. The results for group LN are utterly different. The very first trial shows a substantially reduced suppression effect. The next trial, the first that confirms the new inhibitory rules, showed a further reduced-suppression effect. By the fourth experience of the nonreinforced compound, the suppression ratio has become asymptotic.

Then, four trials after that, the single stimulus N is introduced. For Group N, this is by now simply the standard occurrence, but for Group LN it is an event not encountered since the original conditioning trials, during which N alone was always accompanied by shock. Because the rule $N \Rightarrow S$ has been partially protected from strength reduction by the exception rule $L + N \Rightarrow \bar{S}$, Group LN rats show considerable suppression on the very first presentation of N alone. (The rate at which this initial suppression extinguishes appears to be more rapid in the data presented in Figure 6 than in the simulation results presented in Figure 7. This discrepancy, if it is reliable, is not accounted for by our model.)

One might question whether the decreased suppression attributable to presentation of the light with the noise, especially on the first trial, is really attributable to conditioned inhibition. One alternative possibility is that the initial decrease in suppression is due to Pavlovian external inhibition—a distraction effect on performance, rather than learning. As Kamin (1968) pointed out, however, the effect certainly cannot be attributed to elicitation of an investigatory response incompatible with bar pressing (the usual interpretation of external inhibition) be-

cause the effect of the light is actually to increase bar pressing. The conditioned-inhibition interpretation might also be questioned because Kamin did not perform either a transfer summation test (test of the light as an inhibitor with a different excitor) or a retardation test (test for reduced speed of excitatory conditioning to the light), which are the traditional operational tests for conditioned inhibition (Hearst, 1972; Rescorla, 1969). However, the use of these tests as necessary criteria for conditioned inhibition has been questioned on theoretical grounds (Jenkins, 1985). In any case, Kamin's study clearly shows that the light served to protect the excitatory noise from extinction, which is also a theoretically central property of conditioned inhibition (Konorski, 1948; Soltysik, 1985). On balance, then, an interpretation of Kamin's results in terms of conditioned inhibition seems justified.

Our theory thus explains how the introduction of a new stimulus can sometimes be immediately inhibitory, as long as its occurrence triggers formation of an exception rule predicting absence of an otherwise-expected CS, and the conditions for second-order conditioning of excitation to the new stimulus are unfavorable. Because it allows quantum introduction of new rules, the present theory is also consistent with the apparent one-trial learning effect found by Kamin.

Increased inhibition due to an extinction procedure. We discussed earlier a phenomenon that is especially problematic for the Rescorla-Wagner formulation—failure to demonstrate extinction of the inhibitory power of an inhibitory cue that is presented in the absence of either the excitatory cue or reinforcement (Zimmer-Hart & Rescorla, 1974). Recall that in the Rescorla-Wagner model, an inhibitory cue acquires negative associative strength; accordingly, when it is presented alone it will predict less than zero shock, and hence Equation 1 will adjust its strength upward toward zero. As a result, the cue should produce less inhibition when again paired with the excitatory cue. This prediction of the Rescorla-Wagner model was repeatedly disconfirmed in the Zimmer-Hart and Rescorla study.

As we argued earlier, this inadequacy of the Rescorla-Wagner model highlights the need to be able to predict explicitly the nonoccurrence of the US. In our theory, the failure of a strong

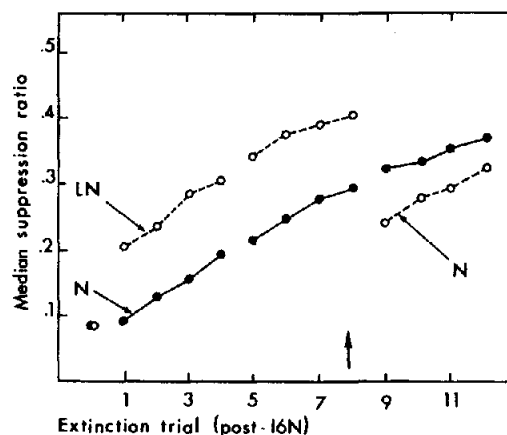


Figure 7. Simulation of the results of Kamin (1968), depicted in Figure 6. (Data are based on 50 simulated subjects in each condition.)

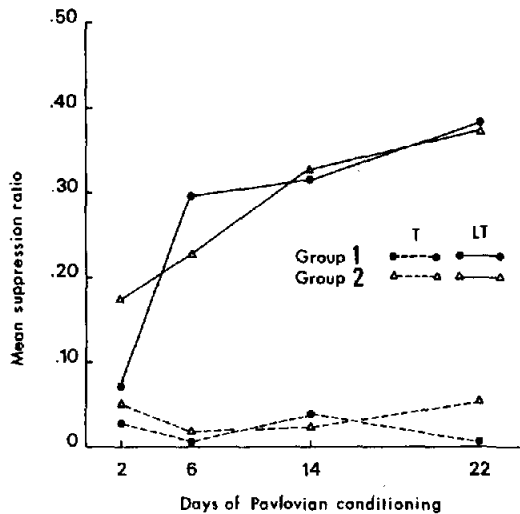


Figure 8. Mean suppression ratios during tone-alone (T) and light-tone (LT) trials in single test sessions following various amounts of training. (Both groups received T+, LT- presentations; Group 2 also received intermixed L- presentations. Data are from Zimmer-Hart and Rescorla, 1974.)

expectation of shock will trigger generation of new rules that predict that negative outcome. To illustrate the behavior of our model in the situation investigated by Zimmer-Hart and Rescorla (1974), we will present the results of our simulation of one of their experiments.

In most of Zimmer-Hart and Rescorla's experiments, the attempt to extinguish the inhibitory cue was made after the compound had reached an asymptotic level of inhibition; accordingly, any increase in the inhibitory power of the inhibitory cue (which our theory predicts if the inhibitor is presented without the US) would not be observable. In one experiment, however, the extinction attempt was made prior to asymptotic conditioning of the compound. All rats were first given training in bar pressing to obtain food; an initial session followed in which a 30-s tone was presented four times, ending each time with a shock. This would establish the rule $T \Rightarrow S$.

The animals then were divided into two groups, which received different procedures for inhibitory conditioning. For both groups, each subsequent session involved four presentations of the tone paired with shock, intermixed with four presentations of the tone in combination with a flashing light without shock. These events would generate the inhibitory rules $L + T \Rightarrow \bar{S}$ and $L \Rightarrow \bar{S}$ by exception formation, establishing the light as an inhibitory cue.

Group 1 received no other presentations of CSs. However, Group 2 also received four intermixed presentations of the light alone without shock. From the perspective of the Rescorla-Wagner theory, these were *extinction* trials that should have diminished the inhibitory power of the light, thus slowing down the acquisition of inhibition to the light-tone compound for Group 2 relative to Group 1. In contrast, from the point of view of our theory, these are additional occasions for strengthening of the $L \Rightarrow \bar{S}$ rule. Because this rule contributes an inhibitory influence when the light-tone compound is presented, its

strengthening when the light is presented alone should actually accelerate early acquisition of inhibition to the compound. Note, however, that the light rule will share reward with the other light-plus-tone exception rule when the compound is presented. Greater strength of the former rule will eventually lead to diminished strength of the latter. Asymptotic suppression to the compound cue should be equal for the two groups. Accordingly, our model predicts that Group 2 will show less suppression than Group 1 to the compound early in acquisition, but that the two groups will show comparable suppression later in training.

Figures 8 and 9, respectively, present the data from the actual experiment by Zimmer-Hart and Rescorla (1974) and from our simulation of it. These data were obtained on test trials involving three reinforced presentations of the tone and two nonreinforced presentations of the light-tone compound. Suppression to the tone presented alone was asymptotic for both groups over the entire test period. The results for the light-tone presentations favor our model over the Rescorla-Wagner account. Group 2, which experienced separate presentations of the inhibitory light CS without reinforcement, exhibited increased inhibition to the light-tone compound during early trials (reported by Zimmer-Hart & Rescorla, 1974, to be marginally significant by a nonparametric test); on later trials, the two groups showed similar levels of suppression. The greater inhibition in Group 2 is somewhat more persistent in the simulation results than in the actual data, but the simulation indicates that the groups should indeed converge on a common asymptote.

Note that our theory accounts for the well-established fact that inhibition is acquired much more readily when the cue intended to become an inhibitor is presented in compound with an excitor ($A+$, $AX-$), rather than alone ($A+$, $X-$). At least some compound $AX-$ presentations are required to establish the rule $X \Rightarrow \bar{S}$ as one that censors the excitatory rule $A \Rightarrow S$. Because the reward for predicting absence of shock supports a much lower asymptotic strength than does that for predicting shock, the censoring mechanism is crucial for allowing robust

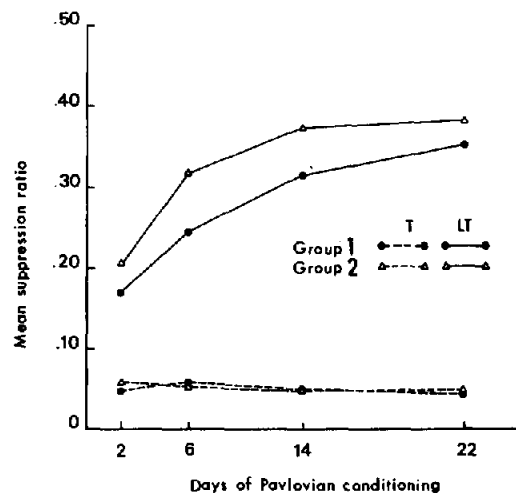


Figure 9. Simulation of the results of Zimmer-Hart and Rescorla (1974), depicted in Figure 8. (Data are based on 50 simulated subjects in each condition.)

inhibition of a strong excitatory rule. The censoring mechanism permits the inhibitor to modulate the effect of the excitor, as suggested by Rescorla (1985). In addition, the inhibitory rule is much more likely to control behavior when the compound cue is presented because it then operates in place of the strong excitatory rule, which dominates the bidding competition; when the inhibitor is presented alone, the rule $X \Rightarrow \bar{S}$ may be bested by the background rule $\# \Rightarrow \text{Press}$. Thus, increased inhibition due to nonreinforced separate presentations of the inhibitory cue will only be found under limited circumstances.

Partial transfer of inhibition. It should be readily apparent that our model correctly predicts that transfer of inhibition will be partial when an inhibitory cue X originally conditioned in the compound AX is presented in the compound BX , in which B is also an excitatory cue. The conditioned-inhibition paradigm typically will trigger generation of both the exception rule $X + A \Rightarrow \bar{S}$ and its generalization $X \Rightarrow \bar{S}$. The latter, but not the former, will be matched when a BX compound is presented. We would therefore predict that in an experiment modeled after that of Zimmer-Hart and Rescorla (1974), animals trained using the procedure of Group 2 would show greater transfer of inhibition than those trained using the procedure of Group 1 if the light were later paired with a different excitatory cue (because in Group 2 the general rule $L \Rightarrow \bar{S}$, which enables transfer, would acquire greater strength than the more specific rule $L + T \Rightarrow \bar{S}$ because of the presentations of the light alone in the absence of shock). As this prediction illustrates, our theory allows the possibility that apparently similar performance (as was obtained for Groups 1 and 2 by the end of training) may arise from differing underlying representations. Such differences in representations may potentially be revealed by indirect measures such as transfer tests.

Explicit Representation of CS Consequences

An important formal distinction between the present approach and that of the Rescorla-Wagner model and its variants is that our theory postulates rules that explicitly represent the consequences of the CS. Rules can predict not only US but \bar{US} . Rules of the latter type were crucial in allowing the theory to account for data concerning conditioned inhibition. This approach readily can be generalized to allow the explicit prediction of US s that differ in magnitude and quality. For example, one rule might predict a single shock (S), whereas another might predict a double shock (SS). Let us consider a number of phenomena that provide evidence distinguishing alternative assumptions regarding what is learned about CS consequences.

Overexpectation. The Rescorla-Wagner model predicts that if two CSs are individually conditioned to a US up to asymptotic strengths (i.e., $A+$, $B+$) and subsequently presented repeatedly as a compound with reinforcement (i.e., $AB+$), then conditioning to the individual cues will actually decrease. This is because the sum of V_A and V_B will exceed the asymptote supportable by the US , causing the associative strengths of each CS to be reduced toward the asymptote. This *overexpectation* effect has been demonstrated in a number of studies (e.g., Kamin & Gaioni, 1974; Kremer, 1978; Rescorla & Wagner, 1972).

In terms of our model, the concept of overexpectation does not really arise. Although rules have continuous strength val-

ues, on any given processing cycle a specific expectation is generated—either the US is expected or it is not. Nonetheless, our model also predicts the overexpectation phenomenon. When the two cues are combined in the second phase, the two rules, $A \Rightarrow S$ and $B \Rightarrow S$, which formerly each received full reward for their successes, will begin to divide the reward, with each paying as large a bid as ever. The result will be a net loss of strength until a lower asymptote is reached at which the payout of each rule is equal to its new, lower, expected reward. Intuitively, a given magnitude of the US justifies a certain probability of a CER. In the first phase of training, each rule individually comes to elicit the appropriate probability of a CER; in the second phase, the two rules together will at first lead to an inappropriately high probability, a problem that the ensuing strength reduction will correct. From our perspective, the phenomenon is not really “overexpectation of the US ,” but rather “excessive emitting of a CER in expectation of the US .”

Although the models make converging predictions for the above situation, the Rescorla-Wagner model makes a highly counterintuitive prediction, which our model does not share, in a variant of the overexpectation paradigm. Suppose in the second phase the two conditioned cues, A and B , are presented in a compound with a third neutral cue, X , accompanied by reinforcement (i.e., $ABX+$). The strengths of all of the associations will be reduced to correct the overexpectation of the US ; not only should A and B become less excitatory, but X , which began at $V_X = 0$, should actually become inhibitory. This inhibitory conditioning is predicted despite the fact that X is always paired with excitatory cues and followed by the US .

The main evidence supporting this prediction comes from an experiment by Kremer (1978). He found that when the second phase of training included $ABX+$ trials intermixed with continuing $A+$ and $B+$ trials (to maintain the high V_{CS} values of A and B), the X cue indeed showed a small amount of inhibition later when first presented alone with the US , and was relatively slow to gain excitatory power.

Our model predicts no such effect because there is never any failed expectation of shock in this paradigm, and hence no trigger for forming any rule to predict absence of the US . But, although Kremer's (1978) study seems to favor the prediction of the Rescorla-Wagner model, closer examination calls its interpretation into question. The main comparison Kremer made was between the aforementioned group that received $ABX+$ trials after initial conditioning to A and to B individually, and a control group that received $AB+$ trials in the second phase. Both groups were then presented with $X+$ trials in the third phase, and the rate of obtaining suppression to X provided the measure of initial inhibition. The $ABX+$ group was slower to acquire suppression to X in the third phase than was the $AB+$ control.

It is evident, however, that this comparison confounds presentation of X in the $ABX+$ compound of the second phase with simple preexposure to X . As we saw earlier, there is a great deal of evidence indicating that a preexposed CS, which is no longer unusual, is more difficult to condition. Moreover, according to our theory, Kremer's $ABX+$, $A+$, $B+$ procedure on the second phase should be particularly damaging to the associability of X . The rule $X \Rightarrow S$ will be formed by chaining. However, not only will this rule be overshadowed by the strong existing A

and B rules, it will actually be driven down to zero strength. This is because both the A and B rules continue to receive full reward when these cues are presented individually, whereas the X rule invariably shares the payout and reward. The latter rule is thus both redundant and less general, a situation that makes its asymptotic strength zero (see our earlier discussion of the simulation of the results of Rescorla, 1968). Thus by the third phase, the X cue is not only familiar, but has already produced a shock-predicting rule that was tested and failed miserably.

Furthermore, Kremer's (1978) experiment included a third condition in which the second phase included ABX+ trials but not A+ or B+ trials. The Rescorla-Wagner model predicts that this variation should produce some lesser degree of inhibitory conditioning to X. Our theory, however, predicts that the excitatory $X \Rightarrow S$ rule will be formed; and because the A and B rules are not rewarded separately, the X rule will be overshadowed but not driven out of the system. The latter rule will therefore be available to gain strength once it is paired separately with shock in the third phase. And in fact, Kremer found that this third condition yielded excitatory conditioning to X.

In sum, our theory can account for the most clear-cut findings from the overexpectation paradigm (reduction in the CER elicited by previously asymptotic cues caused by their presentation as a compound). Where our theory makes a prediction differing from that of the Rescorla-Wagner model, the evidence is less clear but may well be consistent with our proposal.

Superconditioning. Another variation of the blocking paradigm provides evidence that has been used to argue that the expected consequence of a compound is simply the sum of V_{CS} s of its constituent cues. Suppose that in the first phase of learning, a cue A acquires inhibitory power (e.g., as the result of a B+, AB- procedure), and in a second phase, the compound AX is presented paired with the US. According to the Rescorla-Wagner model, the negative value of V_A will increase the discrepancy between the sum $V_A + V_X$ and the asymptote, resulting in especially effective excitatory conditioning to X. Experimental tests have provided evidence for such superconditioning (Rescorla, 1971b; Wagner, 1971).

This evidence must be interpreted carefully. The term *superconditioning* suggests that the cue X is conditioned more effectively in a compound with an inhibitory cue than in isolation. Although Wagner (1971) reported such an effect, attempts to replicate it have not succeeded (Navarro, Hallam, Matzel, & Miller, 1989). The study by Rescorla (1971b) provided only a weaker result. Rescorla showed that a cue X received least conditioning in a compound with an excitatory cue (the familiar case that produces blocking), somewhat more in a compound with another neutral cue (the case that typically produces mutual overshadowing), and most successfully in a compound with an inhibitory cue (the case of superconditioning). No comparison was made between the latter procedure and conditioning to the cue X presented separately. The term *superconditioning* is thus quite misleading for such results; all that has been demonstrated is that prior inhibitory conditioning to the A cue reduces (or perhaps eliminates) the overshadowing otherwise suffered by the X cue when presented in the compound.

These results can readily be explained in terms of rule generation and competition. Excitatory conditioning to A produces a strong rule $A \Rightarrow S$ that blocks strengthening of the rule $X \Rightarrow S$.

If A and X are both neutral when conditioning of the compound begins, both the A and X rules will be formed and compete for reward on an equal footing, neither achieving the strength it would have reached had it been the only operative rule. In contrast, if prior experience has established the inhibitory rule $A \Rightarrow \bar{S}$, the existence of this strong rule will prevent formation of the contradictory rule $A \Rightarrow S$. Instead, only the rule $X \Rightarrow S$ will be formed when the AX compound is paired with shock. Intuitively, the occurrence of the unfamiliar cue X provides a far more plausible hypothesis to explain the pairing of shock with the compound than does occurrence of the familiar cue A, which is known to predict absence of shock. If only the X rule is entertained, it will not have to share reward, and hence will be strengthened as readily as (but not more readily than) if X had been presented as a separate stimulus. Our account thus predicts that true superconditioning—greater conditioning to a cue presented in a compound than in isolation—cannot be obtained. At present, relatively little evidence appears to support this prediction of the Rescorla-Wagner model.

Unblocking. Another phenomenon that provides possible support for our approach comes from studies in which the qualitative nature of the US is varied. Kamin (1968) demonstrated that the usual blocking effect is not obtained if the compound cue is followed by a stronger shock than that presented after the blocking cue alone. Such unblocking can be readily explained by most theories, including that of Rescorla and Wagner (1972) and the present one, inasmuch as the more intense US will allow additional conditioning to occur when the compound is presented. The present theory, however, is better equipped to account for cases of unblocking that are sometimes obtained when the intensity of the US is reduced for the compound cue. For example, suppose that cue A is followed by a double shock (A++). After conditioning of A has approached asymptote, the compound AB++ is presented. The usual blocking of conditioning to B will occur. But now suppose that the AB compound is paired not with the double shock but with a single shock (AB+). A number of studies have found that this unblocking procedure produces a greater amount of conditioning of the B cue (e.g., Dickinson & Mackintosh, 1979; Mackintosh, Dickinson, & Cotton, 1980). The Rescorla-Wagner model cannot account for unblocking because the theory allows the magnitude of the reinforcement only to affect V_{CS} . Thus, the highly excitatory cue A, when presented in conjunction with a neutral cue B, should actually overpredict the single shock, which if anything, should result in B becoming inhibitory (cf. Kremer, 1978).

In contrast, the result is entirely in accord with our approach. In the unblocking paradigm, the animal first acquires the rule $A \Rightarrow SS$. When the AB compound is presented with a single shock, the resulting predictive failure will trigger generation of the new rules $B + A \Rightarrow S$ and $B \Rightarrow S$, which will be strengthened by repetitions of AB+. These exception rules will protect the original expectancy of double shock in the presence of A alone. The latter rule will cause suppression when the B cue is presented alone.

Extensions of the Theory

In the preceding section we presented a broad set of detailed applications of our theory to CER phenomena. In addition to

providing successful fits to phenomena accounted for by earlier proposals (e.g., the primacy of statistical predictability of the US as a determinant of conditioning, blocking), we were able to account for a wide range of results that are problematic for some alternative models (e.g., the detrimental effect of CS pre-exposure on conditioning, failure to extinguish inhibition by separate nonreinforcements, unblocking with a weaker US). No other proposed model yields the entire range of predictions we have outlined.

Our general aim has been to develop a model that integrates mechanisms of hypothesis formation and strength revision within a comprehensive performance system. Theories that invoke the notion of hypothesis generation often have been unduly restrictive, in our view, in assuming that hypotheses are entertained and tested serially, and ultimately rejected if any exceptions are found. By representing hypotheses as rules that can operate in parallel, take on continuous strength values, interact as defaults and exceptions, and actively compete to control behavior and to gain reward for predictive successes, a great deal of theoretical power is gained. Although the present theory makes extensive use of mechanisms for strength revision, heuristics for rule generation are also central. We suspect that the explanatory power of learning models based entirely on strength revision, including current connectionist models, will prove to have limits. A major limitation of current connectionist models of learning is that the proposed algorithms for adjusting connection weights become computationally expensive when the number of interconnections grows large (Hinton, 1987). Heuristics that propose plausible candidate rules can function to drastically reduce the effective size of the search space in which strength-revision procedures operate.

We are optimistic that the present theory can provide insights into forms of learning more complex than classical conditioning. For example, studies of operant conditioning have revealed virtually all of the major phenomena we have discussed in this article, including the effects of statistical covariation, blocking, overshadowing, learned irrelevance, and conditioned inhibition, suggesting important commonalities with simple associative learning (Schwartz, 1978). It follows that a successful theory of classical conditioning will likely contribute to the understanding of operant conditioning as well.

It is possible that some of the constructs used in the theory can be usefully related to work on the biological basis of memory. In particular, research on the effects on conditioning of damage to the hippocampal region has revealed a complex pattern of savings and deficits. Simple excitatory conditioning is typically spared in hippocampally ablated animals; furthermore, the detrimental effects of CS pre-exposure and blocking are actually reduced (Solomon, 1977; Solomon & Moore, 1975). Such results are consistent with our separation of a procedural memory for associations and a long-term declarative memory for feature familiarity. If we assume, following Squire (1982), that the hippocampus supports declarative memory, then its destruction will result in selective loss of the ability to acquire familiarity information. Cues will therefore remain unusual for the hippocampally ablated animal despite extensive pre-exposure, reducing the decrement in conditionability that would normally result.

Although simple excitatory conditioning is preserved after

hippocampotomy, more complex forms of learning such as reversal conditioning (Berger & Orr, 1983) and serial feature-positive discriminations (i.e., A followed by B is reinforced, whereas B alone is not; Ross, Orr, Holland, & Berger, 1984) are impaired. A possibility, albeit speculative, is that these impairments result from disruption of the censoring mechanism involved in more complex default hierarchies. This possibility may be compatible with Hirsh's (1974, 1980) theory of hippocampal functions.

Although the theory presented here can account for a broad range of empirical data, there are many phenomena with which it cannot deal. As mentioned earlier, the model is not equipped to deal with response magnitudes or fine-grained temporal aspects of conditioning (cf. Donegan & Wagner, 1987; Sutton & Barto, 1981). In particular, the strong forward-causation heuristic embodied in the model renders it incapable of addressing the temporal complexities of excitatory and inhibitory backward conditioning, as exhibited in the phenomena demonstrated by Wagner and Larew (1985). More generally, many conditioning phenomena (e.g., partial reinforcement, effects of repeated extinction and reacquisition) appear to depend on the acquisition of information about rates of event occurrences (Gallistel, in press). The current model lacks any serious treatment of the representation of time.

Another interesting class of phenomena that is problematic for the theory involves effects of extinction of an excitatory CS on subsequent responses to another CS. In particular, extinction of an excitatory CS₁ can diminish the effectiveness of an inhibitory CS₂ originally conditioned in conjunction with CS₁, even though CS₂ has not been presented during the extinction trials for CS₁ (Kaplan & Hearst, 1985; Miller & Schachtman, 1985). Similarly, extinguishing an excitatory CS₁ that originally overshadowed an excitatory CS₂ presented in compound with CS₁ can produce an increase in the potency of CS₂ (Kaufman & Bolles, 1981; Matzel, Schachtman, & Miller, 1985). Such findings, which are problematic for virtually all extant conditioning theories, call into question our principle that only rules that are actually tested have their strengths revised. It may be that rules that were acquired together are more intimately coupled, so that changes in the strength of one rule can sometime indirectly affect the strength of another.

We will now suggest some theoretical extensions that may allow our approach to encompass a much wider range of inductive phenomena in lower animals and possibly in humans.

Inferential Heuristics

The present model emphasizes the unusualness heuristic, which serves as the basis for selection of cues for rule formation in the three specific heuristics that are implemented. The present implementation of the unusualness heuristic is oversimplified, in that a simple count of number of onsets serves to establish a cue as familiar, and hence no longer eligible for use in rule formation. The criterion should clearly be more probabilistic; increased familiarity should diminish the chance that a cue will be used to form a rule, but not eliminate it altogether. In addition, forgetting due to interpolated activity should serve to restore a cue to a less familiar state. The most attractive general approach to defining unusualness is that proposed by Wagner

(1978): A cue is no longer unusual if it is itself predictable from its context. This criterion accounts for the fact that the negative impact of preexposure on conditioning is context-specific; a cue preexposed in one context will condition readily if presented in a different context (e.g., Lubow, Rifkin, & Alek, 1976). Note that in order to implement a rule-based system that determines unusualness by assessing the predictability of CSs, it is necessary for the system to generate rules that predict CSs as well as USs. This capacity is itself a crucial extension of the theory, to which we will return.

Thus, an elaborated model would provide a more adequate implementation of the unusualness heuristic. It would also provide a small set of additional general heuristics for rule formation. Despite the extensive efforts of investigators to identify some single principle that governs CS processing, it seems much more likely that multiple factors operate. In addition to the unusualness heuristic, a prime candidate is a predictiveness heuristic, based on a proposal by Mackintosh (1975b). A number of studies have shown that preexposure to a cue is less detrimental to its later conditionability if it is followed by some reinforcer during its initial presentation. Indeed, arbitrary stimuli can be rank-ordered from most to least conditionable as follows: novel cues, preexposed cues that have predicted a different US, preexposed cues that have predicted nothing, and preexposed cues that were previously shown to be worthless as predictors of the US (Dickinson, 1976; Hall & Pearce, 1979). These findings suggest that known predictors are favored over other familiar cues for inclusion in new rules.

There is also evidence that more specific knowledge about plausible predictors can be acquired through experience. Mackintosh (1973) demonstrated that learned irrelevance is to some degree reinforcer-specific: Uncorrelated presentations of a CS and a reinforcer are more detrimental if the animal is later required to learn that the CS predicts that particular reinforcer than to learn it predicts another. Baker and Mackintosh (1979) provided evidence that the detrimental impact of uncorrelated presentations on later conditioning reflects more than the sum of the retardation attributable to preexposure of the CS and the US individually. Such results suggest that animals are able to acquire models of their environment that include relatively specific hypotheses about what kinds of events signal other particular kinds of events.

Another candidate for a general principle involved in rule formation is a similarity-based heuristic: It is easier to learn rules that link stimuli with salient similarities than to link dissimilar stimuli (Rescorla & Furrow, 1977; Seligman, 1970). In addition, some organisms may come "hard wired" with strong biases to form hypotheses linking particular types of events, such as gustatory events and illness in the case of rats. These specific inferential heuristics may underlie the spectacular demonstrations of conditioning over long time lags that are found with certain types of biologically related cues (Garcia et al., 1968). By postulating heuristics for rule formation with varying degrees of generality, it may be possible to integrate findings regarding conditioning of both arbitrary and nonarbitrary associations within a unified general framework.

Rule Sequences

Another important area in which the theory requires extension involves providing mechanisms for learning useful se-

quences of rules. In the present implementation, all rules directly specify an effector action. But even in relatively simple classical conditioning paradigms, there is clear evidence that animals can learn rules relating one CS to another, as well as rules relating a CS to a US. Experiments on second-order conditioning indicate that once a first CS has been conditioned to a US, a second CS paired with the first may acquire either a direct association to the US or an indirect association via the first CS (P. C. Holland, 1980; Rescorla, 1984; Rescorla & Durlach, 1981). The first type of second-order conditioning is approximated by our chaining heuristic; however, the second form requires rule sequences of the form $CS_2 \Rightarrow CS_1$; $CS_1 \Rightarrow US$. Rules that relate one CS to another are also required to implement a more adequate version of the unusualness heuristic.

In fact, our framework is particularly well suited to account for the acquisition of rule sequences. The system of bidding and payout of bids by winning rules used in the present model is directly related to an algorithm developed by J. H. Holland (1986) for back-chaining reward through sequence of rules. In the present model, the payout is simply subtracted from the strength of winning rules, thus serving to stabilize asymptotic strengths of rules. In J. H. Holland's bucket-brigade algorithm, the bid paid out by a rule is subtracted from its own strength and added to the strength of the rule or rules that acted on the prior cycle to enable the rule to be matched. For example, the rule $A \Rightarrow B$ might fire on one cycle and establish the prerequisite for the rule $B \Rightarrow C$ to be matched on the next cycle. The second rule in the sequence would pay its bid to the precursor rule.

The bucket-brigade algorithm allows the system to pass rewards implicitly from rules that are directly rewarded by the environment to earlier acting rules that provided prerequisites, even though the prerequisite rules may never receive direct reinforcement. This computational mechanism may permit extension of the present theory to account for higher order conditioning, and also for acquisition of the more elaborate rule sequences required to model instrumental conditioning (see J. H. Holland et al., 1986, chap. 5).

Implications for Human Learning

Finally, in modeling human cognition, rule-based theories of elementary learning processes provide a common theoretical vocabulary that can be applied both to autonomous inductive learning, such as learning categories from examples, and to instructional learning. For example, the rules that a teacher imparts may be viewed as entering into a competition to represent the world with rules the student had previously induced or had been taught (J. H. Holland et al., 1986; Holyoak & Nisbett, 1987). Although caution is required in relating theories of animal learning to the more abstract learning capabilities of humans, it is possible that default hierarchies recur in various forms of learning that allow organisms to cope with hierarchical relations among contingencies.

References

- Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.
 Baker, A. G. (1974). Conditioned inhibition is not the symmetrical op-

- posite of conditioned excitation: A test of the Rescorla-Wagner model. *Learning and Motivation*, 5, 369-379.
- Baker, A. G. (1976). Learned irrelevance and learned helplessness: Rats learn that stimuli, reinforcers and responses are uncorrelated. *Journal of Experimental Psychology: Animal Behavior Processes*, 2, 130-141.
- Baker, A. G., & Baker, P. A. (1985). Does inhibition differ from excitation? Proactive interference, contextual conditioning, and extinction. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 151-183). Hillsdale, NJ: Erlbaum.
- Baker, A. G., & Mackintosh, N. J. (1977). Excitatory and inhibitory conditioning following uncorrelated presentation of CS and UCS. *Animal Learning and Behavior*, 5, 315-319.
- Baker, A. G., & Mackintosh, N. J. (1979). Preexposure to the CS alone, the US alone or CS and US uncorrelated: Latent inhibition, blocking by context, or learned irrelevance? *Learning and Motivation*, 10, 278-294.
- Berger, T. W., & Orr, W. B. (1983). Hippocampectomy selectively disrupts discrimination reversal conditioning of the rabbit nictitating membrane response. *Behavioural Brain Research*, 8, 49-68.
- Bush, R. R., & Mosteller, F. (1955). *Stochastic models for learning*. New York: Wiley.
- DeVito, P. L., & Fowler, H. (1986). Effect of contingency violations on the extinction of a conditioned fear inhibitor and a conditioned fear excitor. *Journal of Experimental Psychology: Animal Behavior Processes*, 12, 99-115.
- Dickinson, A. (1976). Appetitive-aversive interactions: Facilitation of aversive conditioning by prior appetitive training in the rat. *Animal Learning and Behavior*, 4, 416-420.
- Dickinson, A., & Mackintosh, N. J. (1979). Reinforcer specificity in the enhancement of conditioning by posttrial surprise. *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 162-177.
- Donegan, N. H., & Wagner, A. R. (1987). Conditioned diminution and facilitation of the UR: A sometimes opponent-process interpretation. In I. Gormezano, W. F. Prokasy, & R. F. Thompson (Eds.), *Classical conditioning* (pp. 339-369). Hillsdale, NJ: Erlbaum.
- Einhorn, H. J., & Hogarth, R. M. (1986). Judging plausible cause. *Psychological Bulletin*, 99, 3-19.
- Frey, P. W., & Sears, R. J. (1978). Model of conditioning incorporating the Rescorla-Wagner associative axiom, a dynamic attention process, and a catastrophe rule. *Psychological Review*, 85, 321-340.
- Gallistel, C. R. (in press). *The organization of learning*. Cambridge, MA: MIT Press.
- Gamzu, E. R., & Williams, D. R. (1973). Associative factors underlying the pigeon's key pecking in auto-shaping procedures. *Journal of Experimental Analysis of Behavior*, 19, 225-232.
- Garcia, J., McGowan, B. K., Ervin, F., & Koelling, R. (1968). Cues: Their relative effectiveness as reinforcers. *Science*, 160, 794-795.
- Hall, G., & Pearce, J. M. (1979). Latent inhibition of a CS during CS-US pairings. *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 31-42.
- Hearst, E. (1972). Some persistent problems in the analysis of conditioned inhibition. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 5-39). London: Academic Press.
- Hinton, G. (1987). *Connectionist learning procedures* (Tech. Rep. CMU-CS-87-115). Carnegie Mellon University, Computer Science Department.
- Hirsh, R. (1974). The hippocampus and contextual retrieval of information from memory: A theory. *Behavioral Biology*, 12, 421-444.
- Hirsh, R. (1980). The hippocampus, conditional operations, and cognition. *Physiological Psychology*, 8, 175-182.
- Holland, J. H. (1975). *Adaptation in natural and artificial systems*. Ann Arbor: University of Michigan Press.
- Holland, J. H. (1986). Escaping brittleness: The possibilities of general purpose machine learning algorithms applied to parallel rule-based systems. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 2, pp. 593-623). Los Altos, CA: Kaufmann.
- Holland, J. H., Holyoak, K. J., Nisbett, R. E., & Thagard, P. R. (1986). *Induction: Processes of inference, learning and discovery*. Cambridge, MA: Bradford Books/MIT Press.
- Holland, J. H., & Reitman, J. (1978). Cognitive systems based on adaptive algorithms. In D. Waterman & F. Hayes-Roth (Eds.), *Pattern-directed inference systems*. New York: Academic Press.
- Holland, P. C. (1980). Second-order conditioning with and without unconditioned stimulus presentation. *Journal of Experimental Psychology: Animal Behavior Processes*, 6, 238-250.
- Holland, P. C. (1985). The nature of conditioned inhibition in serial and simultaneous feature negative discriminations. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 267-297). Hillsdale, NJ: Erlbaum.
- Holland, P. C., & Gory, J. (1986). Extinction of inhibition after serial and simultaneous feature negative discrimination training. *Quarterly Journal of Experimental Psychology*, 38B, 245-265.
- Holyoak, K. J., & Nisbett, R. E. (1987). Induction. In R. J. Sternberg & E. E. Smith (Eds.), *The psychology of human thought* (pp. 50-91). New York: Cambridge University Press.
- Hull, C. L. (1943). *Principles of behavior*. New York: Appleton-Century-Crofts.
- James, J. H., & Wagner, A. R. (1980). One-trial overshadowing: Evidence of distributive processing. *Journal of Experimental Psychology: Animal Behavior Processes*, 6, 188-205.
- Jenkins, H. M. (1985). Conditioned inhibition of keypecking in the pigeon. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 118-147). Hillsdale, NJ: Erlbaum.
- Kamin, L. J. (1965). Temporal and intensity characteristics of the conditioned stimulus. In W. K. Prokasy (Ed.), *Classical conditioning: A symposium* (pp. 9-31). New York: Appleton-Century-Crofts.
- Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation*. Miami, FL: University of Miami Press.
- Kamin, L. J., & Gaioni, S. J. (1974). Compound conditioned emotional response conditioning with differentially salient elements in rats. *Journal of Comparative and Physiological Psychology*, 87, 591-597.
- Kaplan, P. S., & Hearst, E. (1985). Excitation, inhibition, and context: Studies of extinction and reinstatement. In P. D. Balsam & A. Tomie (Eds.), *Context and learning* (pp. 195-224). Hillsdale, NJ: Erlbaum.
- Kaufman, M. A., & Bolles, R. C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society*, 18, 318-320.
- Kelley, H. H. (1973). The processes of causal attribution. *American Psychologist*, 28, 107-128.
- Konorski, J. (1948). *Conditioned reflexes and neuron organization*. London: Cambridge University Press.
- Konorski, J. (1967). *Integrative activity of the brain: An interdisciplinary approach*. Chicago: University of Chicago Press.
- Kremer, E. F. (1971). Truly random and traditional control procedures in CER conditioning in the rat. *Journal of Comparative and Physiological Psychology*, 74, 441-448.
- Kremer, E. F. (1978). The Rescorla-Wagner model: Losses in associative strength in compound conditioned stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, 4, 22-36.
- Lubow, R. E., Rifkin, B., & Alek, M. (1976). The context effect: The relationship between stimulus preexposure and environmental preexposure determines subsequent learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 2, 38-47.
- Luce, R. D. (1963). Detection and recognition. In R. D. Luce, R. R.

- Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 103-189). New York: Wiley.
- Mackintosh, N. J. (1971). An analysis of overshadowing and blocking. *Quarterly Journal of Experimental Psychology*, 23, 118-125.
- Mackintosh, N. J. (1973). Stimulus selection: Learning to ignore stimuli that predict no change in reinforcement. In R. A. Hinde & J. S. Hinde (Eds.), *Constraints on learning* (pp. 75-100). New York: Academic Press.
- Mackintosh, N. J. (1974). *The psychology of animal learning*. New York: Academic Press.
- Mackintosh, N. J. (1975a). Blocking of conditioned suppression: Role of the first compound trial. *Journal of Experimental Psychology: Animal Behavior Processes*, 1, 335-345.
- Mackintosh, N. J. (1975b). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82, 276-298.
- Mackintosh, N. J. (1983). *Conditioning and associative learning*. New York: Oxford University Press.
- Mackintosh, N. J., Dickinson, A., & Cotton, M. M. (1980). Surprise and blocking: Effects of the number of compound trials. *Animal Learning and Behavior*, 8, 387-391.
- Mackintosh, N. J., & Reese, B. (1979). One-trial overshadowing. *Quarterly Journal of Experimental Psychology*, 31, 519-526.
- Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning and Motivation*, 16, 398-412.
- Miller, R. R., & Schachtman, T. R. (1985). Conditioning context as an associative baseline: Implications for response generation and the nature of conditioned inhibition. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition*. Hillsdale, NJ: Erlbaum.
- Navarro, J. I., Hallam, S. C., Matzel, L. D., & Miller, R. R. (in press). Superconditioning and overshadowing. *Learning and Motivation*.
- Newell, A. (1972). Production systems: Models of control structures. In W. G. Chase (Ed.), *Visual information processing* (pp. 463-526). New York: Academic Press.
- Nisbett, R. E., Fong, G. T., Lehman, D. R., & Cheng, P. W. (1987). Teaching reasoning. *Science*, 238, 625-631.
- Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford University Press.
- Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, 94, 61-73.
- Pearce, J. M., & Hall, G. (1979). Loss of associability by a compound stimulus comprising excitatory and inhibitory elements. *Journal of Experimental Psychology: Animal Behavior Processes*, 5, 19-30.
- Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87, 532-552.
- Pearce, J. M., Nicholas, D. J., & Dickinson, A. (1982). Loss of associability by a conditioned inhibitor. *Quarterly Journal of Experimental Psychology*, 33B, 149-162.
- Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, 66, 1-5.
- Rescorla, R. A. (1969). Pavlovian conditioned inhibition. *Psychological Bulletin*, 72, 77-94.
- Rescorla, R. A. (1971a). Summation and retardation tests of latent inhibition. *Journal of Comparative and Physiological Psychology*, 75, 77-81.
- Rescorla, R. A. (1971b). Variations in the effectiveness of reinforcement and nonreinforcement following prior inhibitory conditioning. *Learning and Motivation*, 2, 113-123.
- Rescorla, R. A. (1972a). "Configural" conditioning in discrete-trial bar pressing. *Journal of Comparative and Physiological Psychology*, 79, 307-317.
- Rescorla, R. A. (1972b). Informational variables in Pavlovian conditioning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 6, pp. 1-46). New York: Academic Press.
- Rescorla, R. A. (1973). Evidence for a "unique stimulus" account of configural conditioning. *Journal of Comparative and Physiological Psychology*, 85, 331-338.
- Rescorla, R. A. (1984). Comments on three Pavlovian paradigms. In D. L. Alkon & J. Farley (Eds.), *Primary neural substrates of learning and behavioral change*. New York: Cambridge University Press.
- Rescorla, R. A. (1985). Conditioned inhibition and facilitation. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 299-326). Hillsdale, NJ: Erlbaum.
- Rescorla, R. A., & Durlach, P. J. (1981). Within-event learning in Pavlovian conditioning. In N. S. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 81-112). Hillsdale, NJ: Erlbaum.
- Rescorla, R. A., & Furrow, D. R. (1977). Stimulus similarity as a determinant of Pavlovian conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 3, 203-215.
- Rescorla, R. A., & Holland, P. C. (1982). Behavioral studies of associative learning in animals. *Annual Review of Psychology*, 33, 265-308.
- 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 theory and research* (pp. 64-99). New York: Appleton-Century-Crofts.
- Revusky, S. H. (1971). The role of interference in association over a delay. In W. K. Honig & P. H. R. James (Eds.), *Animal memory* (pp. 155-213). New York: Academic Press.
- Rosenblatt, F. (1962). *Principles of neurodynamics*. New York: Spartan Books.
- Ross, R. T., Orr, W. B., Holland, P. C., & Berger, T. W. (1984). Hippocampotomy disrupts acquisition and retention of learned conditional responding. *Behavioral Neuroscience*, 98, 211-225.
- Rudy, J. W., & Wagner, A. R. (1975). Stimulus selection in associative learning. In W. K. Estes (Ed.), *Handbook of learning and cognitive processes* (Vol. 2, pp. 269-303). Hillsdale, NJ: Erlbaum.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning internal representations by error propagation. In Rumelhart, D. E., McClelland, J. L., & the PDP Research Group, *Parallel distributed processing: Explorations in the microstructure of cognition* (Vol. 1, pp. 318-362). Cambridge, MA: Bradford Books/MIT Press.
- Rumelhart, D. E., McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition* (Vols. 1-2). Cambridge, MA: Bradford Books/MIT Press.
- Schwartz, B. (1978). *Psychology of learning and behavior*. New York: Norton.
- Seligman, M. E. P. (1970). On the generality of the laws of learning. *Psychological Review*, 77, 406-418.
- Siegel, S., & Domjan, M. (1971). Backward conditioning as an inhibitory procedure. *Learning and Motivation*, 2, 1-11.
- Solomon, P. R. (1977). Role of the hippocampus in blocking and conditioned inhibition of the rabbit's nictitating membrane response. *Journal of Comparative and Physiological Psychology*, 91, 407-417.
- Solomon, P. R., & Moore, J. W. (1975). Latent inhibition and stimulus generalization of the classically conditioned nictitating membrane response in rabbits (*Oryctolagus cuniculus*) following dorsal hippocampal ablation. *Journal of Comparative and Physiological Psychology*, 89, 1192-1203.
- Sołtysik, S. S. (1985). Protection from extinction: New data and a hypothesis of several varieties of conditioned inhibition. In R. R. Miller

- & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 369-394). Hillsdale, NJ: Erlbaum.
- Squire, L. R. (1982). The neuropsychology of human memory. *Annual Review of Neuroscience*, 5, 241-273.
- Squire, L. R. (1987). *Memory and brain*. New York: Oxford University Press.
- Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: Expectation and prediction. *Psychological Review*, 88, 135-170.
- Sutton, R. S., & Barto, A. G. (1987). A temporal-difference model of classical conditioning. *Proceedings of the Ninth Annual Meeting of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.
- Wagner, A. R. (1969). Stimulus validity and stimulus selection in associative learning. In N. J. Mackintosh & W. K. Honig, (Eds.), *Fundamental issues in associative learning* (pp. 90-122). Halifax, Nova Scotia, Canada: Dalhousie University Press.
- Wagner, A. R. (1971). Elementary associations. In H. H. Kendler & J. T. Spence (Eds.), *Essays in neobehaviorism: A memorial volume to Kenneth W. Spence* (pp. 187-213). New York: Appleton-Century-Crofts.
- Wagner, A. R. (1978). Expectancies and the priming of STM. In S. H. Hulse, H. Fowler, & W. K. Honig (Eds.), *Cognitive processes in animal behavior* (pp. 177-209). Hillsdale, NJ: Erlbaum.
- Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5-47). Hillsdale, NJ: Erlbaum.
- Wagner, A. R., & Larew, M. B. (1985). Opponent processes and Pavlovian inhibition. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 223-265). Hillsdale, NJ: Erlbaum.
- Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, 76, 171-180.
- Widrow, G., & Hoff, M. E. (1960). Adaptive switching circuits. *Institute of Radio Engineers, Western Electronic Show and Convention, Convention Record* (Pt. 4), 96-104.
- Winston, P. H. (1986). Learning by augmenting rules and accumulating sensors. In R. S. Michalski, J. G. Carbonell, & T. M. Mitchell (Eds.), *Machine learning: An artificial intelligence approach* (Vol. 2, pp. 45-61). Los Altos, CA: Kaufmann.
- Woodbury, C. B. (1943). The learning of stimulus patterns by dogs. *Journal of Comparative Psychology*, 35, 29-40.
- Zimmer-Hart, C. L., & Rescorla, R. A. (1974). Extinction of Pavlovian conditioned inhibition. *Journal of Comparative and Physiological Psychology*, 86, 837-845.

Appendix

Simulation of Conditioning Phenomena

The simulation results described earlier were obtained using an implementation of the theory in a Pascal program. In the program, both conditions and actions of rules are represented as bit strings of the three-letter alphabet 1, 0, #, following the classifier notation of J. H. Holland (1986). The # symbol is a wild card in the sense that it can match either a 1 or a 0, thus allowing rules to vary in their generality by increasing the size of the set of matching messages.⁴ All rules have an associated strength parameter, s . Messages have the same bit-string structure as conditions and actions of rules, making the matching process computationally straightforward.

The program includes lists representing four memory stores: (a) long-term memory for rules, (b) long-term memory for occurrence frequencies of features, (c) a short-term buffer for unusual input messages, and (d) a message buffer holding the messages currently being matched and compared.

A conditioning trial is modeled by several cycles of the system, with N cycles corresponding to a minute of simulated time. When a message describing the current state of the environment is received, a comparison with the prediction message posted on the previous cycle determines how much reward is given to the rules that acted on the previous cycle. Specifically, if shock was predicted and occurred, a large reward, R_1 , is given; if absence of shock was predicted and shock did not occur, a lesser reward, R_2 , is given; and if absence of shock was predicted and shock occurred, a negative reward (i.e., punishment), R_3 , is given. Otherwise, the reward is zero.

The message representing the current environment is compared with the message representing the previous environmental state to determine whether an unusual event has occurred. A change in a bit value from 0 to 1 signals the onset of a feature; a check of a long-term feature list then serves to determine if the feature is unfamiliar (defined as having occurred less than F times), in which case the current event is defined as unusual and entered in the short-term event store.

On each cycle, response selection proceeds in the following four steps:

1. Each matched rule posts a bid, b_i , in accord with Equation 2, $b_i = k * s_i$.
2. For each matched rule, a check is made to see if any exception rules are also matched. (Default/exception pairs are marked when exception rules are generated by the appropriate heuristic; see ahead.) The bids of any such exception rules are summed, and with a probability equal to the ratio of the summed bid to a constant criterion T , the corresponding default rule is marked to be censored in Step 4 that follows.
3. All matched rules with consistent effector actions then operate together by summing their bids. Any default rule that in Step 2 is marked to be censored bids alone.
4. For each of the resulting rule sets, the probability of controlling the response is equal to the summed bid of a set divided by the sum of the bids of all matched rules (a simple version of the Luce, 1963, choice rule). If a rule marked to be censored is selected as the winner by this procedure, the corresponding exception rule(s) are substituted for the censored default. That is, the exception rules replace the default just in case the default would otherwise have controlled the response.

The rules created by the learning mechanisms all compete with an a priori rule that calls for bar pressing. The strength of this operant response rule is held constant at strength s_B . The rules that determine the response become the winning set, W . The action called for by W is performed, creating a new predicted message, and the indicated effector action is taken. The next cycle then begins.

The strengths of rules are revised in accord with Equation 3,

$$\Delta s_i = (R - \sum_{j=1}^n b_j) / n \quad \text{for } i \in W, 0 \text{ otherwise.} \quad (3)$$

⁴ J. H. Holland (1986) uses the # symbol on the right-hand side of rules to indicate pass through (i.e., an instruction to copy the corresponding bit from the left-hand side). For our purposes it is sufficient to interpret # as a wild card on both sides of rules.

Given that $b_j = k * s_j$, Equation 3 can be rewritten as follows:

$$\Delta s_i = (R/k - \sum_{j=1}^n s_j)k/n \quad \text{for } i \in W, 0 \text{ otherwise.} \quad (3a)$$

Equation 3a corresponds closely to Equation 1 of the Rescorla-Wagner model for positive values of V_{CS} , if we map rules onto associations. (Associations with negative values of V_{CS} correspond to rules with positive strength values that predict \bar{US} .) The asymptote of conditioning in the Rescorla-Wagner model, λ_{US} , corresponds to R_{US}/k in Equation 3a. The two equations differ in parameters governing rate of learning, particularly in that Equation 3a revises the strengths of rules in W by an amount inversely proportional to the number of rules in W , and leaves the strengths of rules not in W unchanged. The two equations yield essentially equivalent predictions, however, with respect to the pattern of asymptotic strengths. In each case, the expected strength change equals 0 (i.e., strength is asymptotic) whenever a parameter depending on the US (λ_{US} in Equation 1, R_{US}/k in Equation 3a) is equal to the sum of relevant strength values (of associations or rules, respectively).

The program contains three inferential heuristics for generating new rules.

1. *Covariation detection.* (a) If a shock occurs unexpectedly and is preceded by or concurrent with an unusual event stored in the short-term buffer, then construct a new rule or rules with the condition set to the unfamiliar feature(s) of the unusual event and the action set to crouching behavior and expectation of a shock. In all of the experiments simulated, the unusual features preceded shock when covariation detection was triggered. (b) If no unusual event is stored, and no other heuristic applies, then with some probability p ($p < 1$) a general rule is constructed with only $\#s$ in its condition and the same action as discussed above.

2. *Exception formation.* If a rule with strength greater than a criterion C makes an erroneous prediction about the presence or absence of a shock, then check if an unusual event that occurred prior to or concurrent with the prior cycle (when the failed rule was matched) is stored in the short-term buffer. If so, form exception rules by (a) adding the unusual feature(s) to the condition of the failed rule, and substituting the appropriate action, and (b) using the unusual features alone to form the condition. The failed rule is preserved as a default, tagged with the newly created exception rules. Rules formed by exception formation are based on an unusual event that occurred prior to the expected US .

For both covariation detection and exception formation, the initial strength of a new rule is set equal to I_1 if the prediction is occurrence of shock, and to I_2 if the prediction is nonoccurrence of shock ($I_2 < I_1$).

3. *Chaining.* If an unusual event is stored in the buffer and a strong rule (i.e., $s > C$) is included in W , then form a new rule that uses the unfamiliar feature(s) of the unusual event in the condition and that has the same action as its parent rule. The initial strength value of a rule constructed by chaining is set to a proportion of the strength of the parent rule from which the new rule was constructed. The strength of the new rule is set to a higher proportion of the parent rule's strength (I_3) if the onset of the unusual event preceded the old CS (i.e., preceded

the cycle on which the parent rule was matched) than if it was concurrent with it (I_4).

Several additional constraints govern all three heuristics for rule generation. If the selected unusual event has multiple unfamiliar features, all possible rules based on the unfamiliar features and their combinations are constructed. Whenever multiple rules are generated at once by the same heuristic, the initial strength that would be assigned to a single new rule is divided equally among them. (The experiments we have simulated involve events with no more than two unfamiliar features, yielding a maximum of three new rules.) If a heuristic proposes a rule that is already in the system, the new candidate rule is rejected.

The parameters that the program allows to be specified are listed below. The parameter values used in the reported simulations are given in parentheses. Parameters R_3 and I_1 were defined in terms of R_1 ; and R_2 , I_2 , and T were defined in terms of k . R_1 and k were allowed to vary by a factor of not more than 2 in the reported simulations. Extremely similar results were obtained in another set of runs in which k was held constant (at .004) and only R_1 varied.

The parameters are defined as follows:

k : Proportion of strength equal to bid (varied).

R_1 : Reward for correctly predicting shock (varied).

R_2 : Reward for correctly predicting absence of shock (.2 * k).

R_3 : Negative reward for incorrectly predicting absence of shock ($-R_1$).

I_1 : Initial strength of rule predicting shock (covariation detection and exception formation; R_1).

I_2 : Initial strength of rule predicting absence of shock (covariation detection and exception formation; $4 * k$).

I_3 : Initial strength of rule formed by chaining (as proportion of parent rule's strength) when CS_2 precedes CS_1 (.9).

I_4 : Initial strength of rule formed by chaining (as proportion of parent rule's strength) when CS_2 is concurrent with CS_1 (.01).

s_B : Constant strength of bar-pressing rule (.1).

p : Probability of forming rule by covariation detection in the absence of any unusual event (.5).

C : Criterial strength for triggering exception formation or chaining (.5).

T : Constant used to determine probability that exception rules censor default ($30 * k$).

F : Number of presentations required for a cue to become familiar (10).

N : Number of processing cycles per simulated minute of trial (18).

The values of R_1 and k varied as follows across the simulated experiments (in order of presentation in the article): (a) Rescorla (1968): $R_1 = .2$, $k = .003$; (b) Kamin (1968) blocking experiment: $R_1 = .3$, $k = .004$; (c) Rescorla (1972b): $R_1 = .3$, $k = .004$; (d) Kamin (1968) conditioned-inhibition experiment: $R_1 = .19$, $k = .006$; (e) Zimmer-Hart and Rescorla (1974): $R_1 = .19$, $k = .006$.

Received March 20, 1987

Revision received July 7, 1988

Accepted July 8, 1988 ■