Psychological Bulletin
1993, Vol. 114, No. 3, 435–448

Human Contingency Judgments: Rule Based or Associative?

Lorraine G. Allan

The study of the mechanism that detects the contingency between events, in both humans and nonhuman animals, is a matter of considerable research activity. Two broad categories of explanations of the acquisition of contingency information have received extensive evaluation: rule-based models and associative models. This article assesses the two categories of models for human contingency judgments. The data reveal systematic departures in contingency judgments from the predictions of rule-based models. Recent studies indicate that a contiguity model of Pavlovian conditioning is a useful heuristic for conceptualizing human contingency judgments.

The study of the mechanism that detects the contingency between events, in both humans and nonhuman animals, is a matter of considerable research activity. With humans, numerous studies have varied the contingency or covariation between two variables and have examined how such variations influence judgments about the relationship between the two variables. With animals, interest has focused on how such covariations influence conditioned responses (Pavlovian conditioning) or operant behavior (operant conditioning). Two broad categories of explanations of the effects of contingency manipulations on behavior have received extensive evaluation: *rule-based models* and *associative models.* Rule-based models represent organisms as intuitive statisticians who extract contingency information by applying a rule to integrate probabilities or frequencies of events (e.g., Peterson & Beach, 1967). Associative models postulate that apparent contingency learning is really the result of Pavlovian associations formed between all contiguously presented events. The purpose of this article is to assess the status of these two categories of models for human contingency judgments (see also Shanks, in press).

Contingency-Judgment Tasks

With few exceptions (see Well, Boyce, Morris, Shinjo, & Chumbley, 1988), studies concerned with human contingency judgments have involved binary variables. In the simplest case, there are only two binary variables. The subject is presented with information about pairings between the two values of the two variables and then asked about the relationship between the variables. Even in the case of only two binary variables, different tasks have been used. The tasks have varied along a number of dimensions: the question asked, the type of scale used, the presentation mode, the representation of the variables, and the trial procedure.

Experimenters have posed various questions when probing subjects about the relationship between two variables. Subjects have been asked about the contingency or connection between the two variables; about the control, influence, or effectiveness of one variable over the other variable; and about whether one variable predicts or causes the other variable. In this article, *contingency judgment* is used as the generic term.

The perceived relationship between two variables is usually assessed with a rating scale. The rating scale can be unidirectional (one end labeled *no contingency* and the other labeled *perfect contingency*) or bidirectional (one end labeled *perfect negative contingency,* the middle labeled *no contingency,* and the other end labeled *perfect positive contingency*). A bidirectional scale provides information about the sign and the strength of the contingency judgment; the unidirectional scale provides information only about the strength.

Information about the pairings of the values of the two binary variables has been summarized for the subject in a 2 × 2 contingency matrix (described situation; see Shanks, 1991b) or has been presented sequentially in a trial-by-trial format (experienced situation). The sequential information can, of course, be summarized by the 2 × 2 contingency matrix. Figure 1 shows the standard matrix, with I_1 and I_2 as the two values of the input variable (I), and O_1 and O_2 as the two values of the outcome variable (O). The letters in the cells (*a, b, c,* and *d*) represent the joint frequency of one value of I and one value of O.

The two values of I and of O can be represented symmetrically as two events or asymmetrically as event and nonevent. Using the notation introduced by Allan and Jenkins (1980, 1983), we designate the situation where both values of I are events as 2I and the situation where only one value of I is an event as 1I; similarly, 2O designates the situation where both values of O are events, and 1O the situation where only one value of O is an event. The four conditions resulting from the combination of the two representations of I (2I and 1I) with the two representations of O (2O and 1O) are 2I/2O, 1I/2O, 2I/1O, and 1I/1O. In most contingency studies, it is the 1I/1O representation that has been used. For the 1I/1O representation, I_1 and O_1 in Figure 1 denote events, and I_2 and O_2 denote nonevents. In Allan and Jenkins (1980) study, for example, the subject could

---

The preparation of this article was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

I am grateful to Herb Jenkins for introducing me to the world of contingency judgments and to Shep Siegel for many stimulating discussions and for his insightful comments on a draft of this article.

Correspondence concerning this article should be addressed to Lorraine G. Allan, Department of Psychology, McMaster University, Hamilton, Ontario, Canada L8S 4K1. Electronic mail should be addressed to allan@mcmaster.ca.

	O_1	O_2	
I_1	a	b	a+b
I_2	c	d	c+d
	a+c	b+d	a+b+c+d = N

Figure 1. The standard 2 × 2 contingency matrix. (I_1 and I_2 are the two values of the input variable [I], and O_1 and O_2 are the two values of the outcome variable [O]. The letters in the cells [a, b, c, and d] represent the joint frequency of one value of I and one value of O.)

choose to move a joystick (event) or to leave the joystick in its resting position (nonevent). At the end of the response period, one of two pictures occurred. One picture showed a lake scene with the Loch Ness Monster poking its head out of the lake (event); the other picture showed the same lake scene without the Monster (nonevent).

Variants of two basic procedures have been used to present the information sequentially: discrete trial and free operant. In the discrete-trial procedure, each of a series of trials is divided into clearly defined input and outcome periods. The joystick situation described above is an example of a discrete-trial procedure (e.g., Allan & Jenkins, 1980, 1983). Another version of the discrete-trial procedure is a video game developed by Shanks and collaborators (e.g., Dickinson, Shanks, & Evenden, 1984; Shanks, 1985a, 1985b, 1986). In this game, subjects are asked to judge the effectiveness of a new type of shell in destroying tanks. On each trial, a tank moves across the video screen, passing through a gun sight. The subject has the choice of firing or not firing a shell at the tank and then observes whether the tank was destroyed.

Wasserman and collaborators (e.g., Chatlosh, Neunaber, & Wasserman, 1985; Wasserman, Chatlosh, & Neunaber, 1983) introduced the free-operant procedure as an alternative to the discrete-trial procedure. In the free-operant procedure, trials are not defined. Subjects can respond (for example, press a key) whenever they wish rather than just during a predefined period. The presentation of the outcome (for example, a brief illumination of a light) is determined on the basis of a sampling interval (e.g., 1 s). If at least one response occurs during the sampling interval, then the outcome occurs with probability $P(O_1|I_1)$; if no response occurs during the sampling interval, then the outcome occurs with probability $P(O_1|I_2)$.

The task variations described above do have some influence on contingency judgments. Fortunately, data relevant to an evaluation of theoretical accounts of contingency judgments are relatively independent of the task used.

Rule Analyses of Contingency Judgments

The appropriate statistical measure of the dependency of the outcome variable O on the input variable I is ΔP, which is the difference between two independent conditional probabilities (see Allan, 1980). Referring to Figure 1,

$$\Delta P = P(O_1|I_1) - P(O_1|I_2) = a/(a+b) - c/(c+d). \quad (1)$$

Many studies of contingency judgments have concentrated on determining whether humans could accurately judge the size or the sign of the contingency between two binary variables. In such studies, a number of problems are presented to the subject, with the contingency between the input and outcome variables varying over problems. After each problem, the subject is asked to rate the contingency between the variables. Most of these studies used the 1R/1O representation, and usually reported a high correlation between contingency judgments and the actual contingency between the input and output variables (ΔP) (e.g., Allan & Jenkins, 1980, 1983; Alloy & Abramson, 1979; Chatlosh et al., 1985; Dickinson et al., 1984; Neunaber & Wasserman, 1986; Shanks, 1985a, 1987; Wasserman, 1990b; Wasserman et al., 1983; Wasserman, Dorner, & Kao, 1990; Wasserman, Elek, Chatlosh, & Baker, 1993; Wasserman & Shaklee, 1984).

Although the correlation between judgments and ΔP was often high for the 1R/1O representation, systematic departures from ΔP were noted. There were reports, for example, of a *density bias*: Judgments of contingency were not constant for a fixed ΔP, but increased with the frequency of the outcome event [$P(O_1)$] (e.g., Allan & Jenkins, 1980, 1983; Alloy & Abramson, 1979; Baker, Berbrier, & Vallee-Tourangeau, 1989; Chatlosh et al., 1985; Dickinson et al., 1984; Shanks, 1985a; Shanks & Dickinson, 1991). One attempt to deal with the density bias was to correlate judgments with a weighted ΔP, rather than with the unweighted ΔP defined in Equation 1. The unweighted ΔP weights each cell of the matrix equally; the weighted ΔP weights the four cells differentially. The use of a weighted ΔP is supported by the finding that subjects do not evaluate the four cells of the contingency matrix to be of equal importance; they rank the necessity of the four kinds of information as cell *a* > cell *b* > cell *c* > cell *d* (e.g., Kao & Wasserman, 1993; Levin, Wasserman, & Kao, 1992; Wasserman et al., 1990). Weighting ΔP by, for example, assigning a greater weight to $P(O_1|I_1)$ than to $P(O_1|I_2)$ does improve the fit (e.g., Kao & Wasserman, 1993; Levin et al., 1992; Shanks & Dickinson, 1986; Wasserman et al., 1990).

Another attempt to deal with the density bias was reported by Wasserman et al. (1993). They asked subjects to estimate $P(O_1|I_1)$ and $P(O_1|I_2)$ and found these estimates to be inaccurate. They investigated the possibility that contingency judgments were determined by estimates of ΔP based on these inaccurate probability estimates. Wasserman et al. (1993) did not find support, however, for this hypothesis. Rather, contingency judgments were better described by ΔP based on presented probabilities (Equation 1) than by ΔP based on the subjects' estimated probabilities.

Cheng and Novik (1990, 1991, 1992) and Melz, Cheng, Holyoak, and Waldmann (1993) also addressed departures in judgments from ΔP. They proposed that the ΔP rule applies across a focal set of trial types rather than across all trials. For the standard ΔP rule, $P(O_1|I_1)$ is based on all I_1 trials and $P(O_1|I_2)$ on all I_2 trials. The presence of other inputs is ignored in the determination of ΔP. For the focal-set ΔP rule, these other inputs would define the I_1 and I_2 trials on which $P(O_1|I_1)$ and $P(O_1|I_2)$ are based. If only a subset of trials are included in the calculation of ΔP, then the value of ΔP could be very different than if all trials are included.

Cheng and Novik (1992) suggested that the focal-set ΔP rule is relevant to the human contingency literature. They made no reference, however, to the many studies that were conducted in recent years that provided evidence against the ΔP rule (e.g., Baker & Mazmanian, 1989; Baker, Mercier, Vallee-Tourangeau, Frank, & Pan, 1993; Chapman, 1991; Chapman & Robbins, 1990; Dickinson et al., 1984; Shanks, 1985b, 1986, 1989, 1991b; Wasserman, 1990a).

Shanks (1993, in press) provided an eloquent critique of the focal-set ΔP rule. He showed that the focal-set ΔP rule could account for some data that were inconsistent with the simple ΔP rule. In the typical 1I/1O contingency-judgment task, the outcome event (O_1) occurs both in the presence and in the absence of the input event. $P(O_1 | I_1)$ is calculated across all trials on which the input is present (I_1) and $P(O_1 | I_2)$ is calculated across all trials across on which the input is absent (I_2). A modification of this simple contingency-judgment task is to signal all O_1 events that occur in the absence of the input event (Shanks, 1986, 1989). Shanks (1989), for example, used a computer version of the free-operant task. Subjects had the option of pressing or not pressing the space bar on a computer keyboard, and the outcome was the presence or the absence of an event on the computer screen. In the signaled condition, every outcome that occurred in the absence of the action was preceded by another event, the signal, which was a short tone. If ΔP was calculated across all trials, then the value would be the same for the signaled and the unsignaled conditions. Shanks (1986, 1989), however, found that judgments about the effectiveness of responding were greater in the signaled condition than in the unsignaled condition.

Focal-set ΔP would not be determined across all trial types. Rather, it would be determined by contrasting what happens when $P(O_1 | I_1)$ and $P(O_1 | I_2)$ are based on trials that are identical except for the presence and absence of the input event. In the unsignaled condition, ΔP and focal-set ΔP would be identical. In the signaled condition, this would not be the case. Focal-set ΔP would ignore those trials on which the outcome event was signaled, resulting in a smaller value of $P(O_1 | I_2)$ and therefore a larger ΔP.

In his critique of the focal-set ΔP rule, Shanks (1993) showed that although the model could account for data from some contingency experiments (such as signaling), it often failed. Moreover, although there were situations in which it was clear how to define the focal set, the focal-set ΔP rule provided no independent means of determining what the focal set was for a given subject in a given experimental situation:

> It seems that all we can do is to elicit judgments in some situation, and then infer back to what the subject's focal set must have been, assuming he or she is computing contingency according to conditional probabilities. . ..The focal set is therefore just an additional degree of freedom available to the experimenter. (Shanks, 1993)

Alloy and Tabachnik (1984) also addressed departures in judgments from ΔP. They suggested that the situational information available from the 2 × 2 matrix interacted with prior expectations. According to Alloy and Tabachnik, two sources of information are relevant to perceiving contingency: situational information that specifies the objective contingency between events and the subject's prior expectations or beliefs about event contingency (see also Arkes & Harkness, 1983; Baker & Mercier, 1989). Alloy and Tabachnik argued that their theoretical framework could encompass results from animal experiments concerned with manipulations of contingency on strength of conditioning, as well as human contingency judgments.

In a critique of the Alloy and Tabachnik (1984) framework, Goddard and Allan (1988) suggested that their framework led to predictions that were inconsistent with available data, both human and animal. They also showed that even when the framework could account for the data, the explanation was post hoc and often would not have been predicted by the framework. Moreover, there is now abundant evidence that a contiguity model of Pavlovian conditioning is the appropriate way to account for the effects of contingency manipulations with infrahuman subjects (see Papini & Bitterman, 1990).

The Search for the Rule

The strong relationship that is often observed between contingency judgments and ΔP does not necessarily imply that judgments are based on ΔP. Although ΔP is the statistically appropriate summary of the 2 × 2 matrix (see Allan, 1980), it is, of course, not the only summary. ΔP compares conditional probabilities. A summary of the matrix can be based on frequencies. One possibility is ΔD, the difference between the sums of the two diagonal cell frequencies. Referring to Figure 1,

$$\Delta D = (a + d) - (b + c).$$

The two summary numbers, ΔP and ΔD, are perfectly correlated when the two input values are equally likely; that is, $P(I_1) = P(I_2)$ (Allan, 1980). Specifically,

$$\Delta D = N\Delta P \quad \text{if } (a + b) = (c + d).$$

The coordination of ΔD to the cells of the matrix depends on the representation of the input and outcome variables (Allan & Jenkins, 1983). For the 1I/1O representation, ΔD can be described as a comparison of confirming and disconfirming cases. Confirming cases are the joint occurrence of two events (input event paired with outcome event) and of two nonevents (input nonevent paired with outcome nonevent), whereas disconfirming cases are the joint occurrence of an event and a nonevent (input event paired with outcome nonevent, and input nonevent paired with outcome event). Although ΔD can be calculated for the other three representations (2I/1O, 1I/2O, and 2I/2O), the coordination of ΔD to the cells of the matrix is arbitrary, because the cells cannot be represented as confirming cases and disconfirming cases (see Allan & Jenkins, 1983).

ΔD uses information from all four cells of the matrix. Other summaries use less information. Two possibilities examined in the literature are (a) a comparison of cell *a* frequency with cell *c* frequency (F_{a-c}) and (b) a comparison of cell *a* frequency with cell *b* frequency (F_{a-b}). ΔP is perfectly correlated with F_{a-c} when the two input values are equally likely [$P(I_1) = P(I_2)$] and with F_{a-b} when the two outcome values are equally likely [$P(O_1) = P(O_2)$]. Specifically,

$$F_{a-c} = N\Delta P2 \quad \text{if } (a + b) = (c + d)$$

and

$$F_{a-b} = N\Delta P/2 \quad \text{if } (a + c) = (b + d).$$

Great effort has been expended on attempts to determine which summary or rule best describes human contingency judgments. One approach has been to correlate contingency judgments with the various rules to determine which rule provides the best fit to the judgments. This correlational analysis is only informative when the marginal frequencies are different; that is, when $P(I_1)$ does not equal $P(I_2)$ and $P(O_1)$ does not equal $P(O_2)$. Some studies that have imposed this constraint have concluded that ΔP provides a better description than any of the frequency rules (e.g., Wasserman, 1990b; Wasserman et al., 1983), but others have concluded otherwise. For example, judgments in Allan and Jenkins (1983) were best described by ΔD, whereas in Wasserman et al. (1990) F_{a-b} provided the best description for many subjects.

Arkes and Harkness (1983) and Shaklee and collaborators (e.g., Shaklee, 1983; Shaklee & Mimms, 1982; Shaklee & Tucker, 1980; Shaklee & Wasserman, 1986) used a different approach in their attempts to identify the best rule. They constructed a problem set designed to produce a distinct judgment pattern by each of the rules. This rule-analytic technique revealed that the modal rule was not ΔP; it was ΔD in some studies and F_{a-b} in others.

Shanks (1985a, 1987) suggested that another way to evaluate the rules was to track contingency judgments across time. He showed that ΔP and ΔD make different predictions about the manner in which judgments should change with exposure to the relationship between the input and outcome variables. While estimates of conditional probabilities become more accurate with increasing sample size, the mean estimate is independent of sample size. According to the ΔP rule, weighted or unweighted, judgments should be constant across trials. This is also the case for the focal-set ΔP rule. In contrast, ΔD changes across trials, increasing for a positive contingency and decreasing for a negative contingency. For zero contingency, ΔD is constant across trials if the cells are unweighted. If confirming cases are given more weight than disconfirming cases, ΔD increases across trials for zero contingency. Shanks (1985a, 1987) showed that contrary to ΔP rules, judgments did change across trials and that the changes in judgments across trials could not be attributed to changes in confidence with increasing number of trials. However, the acquisition functions did not conform to the predictions of the ΔD rule, weighted or unweighted.

Most studies that have searched for the best rule have used the 1I/1O representation. One of the earliest studies of human contingency judgments, that of Jenkins and Ward (1965), explored the 2I/1O representation, as did later studies by Allan and Jenkins (1980, 1983). For the 2I/1O representation, the two input values are both events. In Allan and Jenkins (1980), for example, the values of the input variable were controlled by the position of a joystick: The subject could choose to move the joystick to the left (an event) or to move it to the right (an event). In the studies using the 2I/1O representation, judged contingency was not well described by any of the rules discussed so far. Rather, judgments increased as the frequency of outcome event, O_1, increased. Instead of comparing the occurrence of the outcome event for the two input values, subjects appeared to compare the overall occurrence of the outcome event to an assumed baseline of no occurrence of the outcome event in the absence of an input event. Subjects appeared to judge contingency as though they assumed that in the absence of an input event, the outcome event would not have occurred.

Allan and Jenkins (1983) compared judgments under all four representations. The input variable was the same as that used by Allan and Jenkins (1980). The outcome variable was movement of a dot on a computer screen. When the outcome values were asymmetric (1O), the dot moved up (an event) or it remained stationary (a nonevent); when the outcome values were symmetric (2O), the dot moved up (an event) or down (an event). Allan and Jenkins (1983) conducted an extensive correlational analysis of their data. Judgments were not well described by ΔP when $P(I_1)$ was not equal to $P(I_2)$. Although ΔD was somewhat better, the coordination of ΔD to the cells of the 2×2 matrix was ad hoc for all representations except 1I/1O, because for the other three representations, the cells could not be represented as confirming and disconfirming cases.

Summary

The search for the best rule has revealed systematic departures in contingency judgments from ΔP. This is the case even when weights are applied to the cells of the matrix or when ΔP is based on estimates of the cell frequencies. Other modifications to the ΔP rule, such as the focal set ΔP rule and the interaction of expectations with ΔP, have been shown to be inadequate. Studies designed to discriminate among the rules have often found that frequency rules describe the data better than ΔP. Although the frequency rules fare better than ΔP, the shape of the acquisition function is not in accord with any of the rules that have been examined. The quest for a rule to describe human judgments of the contingency between binary variables has not yet yielded a satisfactory solution.

Associative Models

In early research on human contingency judgments, the task was conceptualized as involving one input variable and one output variable. More recent research has examined human contingency judgments in situations where multiple input variables are present. The video game described earlier, for example, can be presented with two input variables. In the two-input version of the game, subjects are informed that as the tank moves across the screen, it is moving through a minefield. In this description, firing and mines are the two input variables, and tank destruction is the outcome variable.

The research with multiple input variables has revealed that the judgment of the contingency between one of the input variables and the output variable is influenced by the copresence of the other input variables and by the pairing history of the input variables (e.g., Baker & Mazmanian, 1989; Baker et al., 1993; Chapman, 1991; Chapman & Robbins, 1990; Dickinson et al., 1984; Shanks, 1985b, 1986, 1989, 1991b; Wasserman, 1990a). The nature of the interaction between the input variables is reminiscent of that observed in animal-conditioning experiments and is addressed by associative models. Such data have suggested to a number of researchers that "rather than assuming that subjects mathematically transform real-time events into probabilities and then arithmetically compare those probabilities. . .one might look to elementary associative principles for a viable account" (Wasserman et al., 1993, p. 183).

Early learning theorists assumed that temporal contiguity of a conditioned stimulus (CS) and an unconditioned stimulus (UCS) was sufficient for associative learning. In the 1960s, a number of experiments showed that conditioning was not the inevitable result of CS–UCS pairings (e.g., Kamin, 1969a, 1969b; Rescorla, 1968; Wagner, 1969a). These experiments identified two difficulties for the traditional contiguity view of conditioning. They showed that conditioning was influenced not only by the occurrence of the UCS in the presence of the CS but also by the occurrence of the UCS in the absence of the CS; that is, by ΔP. These experiments also showed that for simultaneously presented stimuli, the conditioning of one stimulus was influenced by the copresence of the other stimuli (e.g., overshadowing) and by the pairing history of the stimuli (e.g., blocking).

Rescorla (1968) postulated a contingency model, similar to the ΔP rule (Equation 1), to describe the influence of contingency manipulations on conditioning. A year later, Rescorla (1969), and also Wagner (1969a), expressed dissatisfaction with the idea that animals "take in large blocks of time, count up numbers of US events, and somehow arrive at probability estimates" (Rescorla, 1969, p. 84). Building on a proposal by Kamin (1969a, 1969b), they developed a new contiguity theory that accounted for many of the results challenging the old (Rescorla & Wagner, 1972; Wagner & Rescorla, 1972). This theory provided the means for animals to "bring together the effects of events separated in time in such a way as to permit all learning to depend on events occurring closely in time" (Rescorla, 1969, p. 88).

Kamin (1969a) proposed that "perhaps for an increment in associative connections to occur, it is necessary that the US instigate some 'mental work' on the part of the animal. This mental work will occur only if the US is unpredictable—if it in some sense 'surprises' the animal" (p. 293). The Rescorla–Wagner (R-W) model formalizes this notion by postulating that the change in the strength of the association between a CS and a UCS is proportional to the degree to which the UCS is unpredicted or surprising. More precisely, the associative or predictive strength (V) of a CS will change on each trial it is presented according to the standard linear operator equation

$$\Delta V = \alpha\beta(\lambda - \Sigma V),$$

where ΔV is the change in predictive strength of the CS, α and β are learning-rate parameters that depend on the salience of the CS and the effectiveness of the UCS, respectively, λ is the maximum amount of predictive strength supported by the UCS, and ΣV is the algebraic sum of the predictive strengths of all stimuli present on that trial. The change in the V value of a particular cue is determined not only by its own V value but also by the associative strength of all cues present on the trial, ΣV. The surprise value of the UCS is the difference between λ and ΣV: The greater the difference between λ and ΣV, the more surprising the UCS is.

The essence of the R-W model is associative competition: There is a limit, λ, to the amount of associative strength that a UCS can support. This limited amount of associative strength is allocated among all stimuli present on the trial: If one stimulus acquires more of the associative strength available, then all other stimuli that are present at the same time must get less. Because a CS never occurs in isolation but is always compounded with contextual (background) stimuli, the UCS is associated with the CS and also with contextual cues. There is always competition; the CS competes with contextual stimuli for associative strength.

The R-W model is one example of current contiguity theories of animal learning (e.g., Gibbon & Balsam, 1981; Jenkins, Barnes, & Barrera, 1981; Mackintosh, 1975; Pearce & Hall, 1980). All of the models share the assumption that the UCS is associated both with the nominal CS and with all other stimuli, including the context. The critical feature of these models for human contingency judgments is that they incorporate a process of *selective attribution*: The attribution of an outcome (i.e., an effect) to a target input (i.e, a cause) depends on the predictive value of other potential inputs (causes) that are simultaneously present.

Adaptive Network Models

Gluck and Bower (1988; see also Shanks, 1990, 1991a) pointed out the similarities between associative models of animal learning and adaptive network models, also known as parallel distributed processing or connectionist networks (McClelland & Rumelhart, 1986; Rumelhart & McClelland, 1986). Adaptive networks consist of processing units or nodes connected by weighted, unidirectional links of activation. The nodes are separated into layers: an input layer, an output layer, and hidden layers. When a stimulus is presented to the network, a set of input nodes is activated. These nodes pass their weighted activation to nodes in the next layer. The resulting pattern of activation in the output layer corresponds to the estimated outcome. The network then receives feedback regarding the desired output. The weights are adjusted to bring the output closer to the feedback. By repeated cycling through output–feedback pairings, the system "learns" the weights that will achieve the closest match.

Gluck and Bower (1988) described a network model that is formally equivalent to the R-W model. The R-W network has two layers, an input layer and an output layer. The weights of the links between the input and output nodes are altered according to the least mean squares (LMS) rule, which is mathematically equivalent to the R-W learning rule. These weights are updated only on trials on which a node is activated; that is, on trials when the stimulus is present.

Although the Gluck–Bower LMS network is formally equivalent to the R-W model, the human-contingency-judgment literature has generally relied on the latter. For that reason, the language of the R-W model will be used in this article.

Associative Models and Human Contingency Judgments

The new contiguity models were developed to explain classical conditioning but have been applied to operant conditioning as well (e.g., Dickinson, 1980; Mackintosh, 1983). Some of the experimental tasks used to evaluate an associative account of human contingency judgments resemble classical conditioning (cue-outcome); others resemble operant conditioning (action-outcome; see Shanks, in press). In a cue-outcome task, the judg-

	UCS	noUCS	
CS + context	a	b	a+b
context	c	d	c+d
	a+c	b+d	a+b+c+d = N

Figure 2. The standard 2 × 2 contingency matrix with labels appropriate for an aminal contingency experiment. (UCS = unconditioned stimulus, no UCS = no unconditioned stimulus, CS = conditioned stimulus.)

ment is about the contingency between a cue (e.g., the presence or absence of a symptom) and an outcome (e.g., the presence or absence of a disease): The cue is the CS, and the outcome is the UCS. In an action-outcome task, the judgment is about the contingency between an action (e.g., pressing or not pressing a key) and an outcome (e.g., the presence or absence of a light flash): The action is the operant, and the outcome is the reinforcer. For simplicity, the classical-conditioning notation is used in this article.

In all tasks used to evaluate an associative account of human contingency judgments, whether action-outcome or cue-outcome, both binary variables have been asymmetric; the two values of each variable are event and nonevent. In the notation introduced earlier, the 1I/1O representation of the variables has been used in such studies.

There is now considerable evidence of similarities between the operations that modulate the strength of conditioning in animals and those that modulate the rating of contingency by humans. The correspondences include the influence of contingency manipulations (e.g., Allan & Jenkins, 1980, 1983; Alloy & Abramson, 1979; Chatlosh et al., 1985; Dickinson et al., 1984; Neunaber & Wasserman, 1986; Shanks, 1985a; Wasserman, 1990b; Wasserman et al., 1983, 1990, 1993); Wasserman & Shaklee, 1984); the effect of signaling (e.g., Shanks, 1986, 1989); the shape of the acquisition (learning) functions (Baker et al., 1989; Shanks, 1985a, 1987); the effect of temporal contiguity (Reed, 1992; Shanks, 1989; Shanks & Dickinson, 1991; Shanks, Pearson, & Dickinson, 1989; Wasserman & Neunaber, 1986); and the role of stimulus interactions, such as blocking (e.g., Chapman, 1991; Chapman & Robbins, 1990; Dickinson et al., 1984; Shanks, 1985b), conditioned inhibition (e.g., Chapman, 1991; Chapman & Robbins, 1990), and relative cue validity (e.g., Baker & Mazmanian, 1989; Baker et al., 1993; Shanks, 1991b; Wasserman, 1990a). These similarities in outcome of animal-learning experiments and human-contingency-judgment experiments suggest a common theoretical account.

Contingency Manipulations

According to the R-W model, contingency manipulations affect conditioning because of the competition for associative strength among various associations formed, both with the CS and with contextual cues. Specifically, the context–UCS associations compete with and block the CS–UCS association. Figure 2 contains the 2 × 2 contingency matrix, with labels appropriate for an animal contingency experiment. On CS trials, the CS is compounded with the context. On noCS trials, only the context is present. On UCS trials, the UCS is available; on noUCS trials, the UCS is absent. The following equations represent the change in associative, or predictive, strength of the CS and of the context for the four trial types specified by the 2 × 2 matrix. V_{CS} and V_X represent the associative strength of the CS and the context respectively; ΔV_{CS} and ΔV_X represent the change in associative strength of the CS and of the context, respectively; α_{CS} and α_X represent the salience of the CS and of the context respectively; β_{UCS} and β_{noUCS} represent the effectiveness of the UCS when it is present and absent, respectively; and λ is the maximum amount of associative strength supported by the UCS, with $\lambda = 0$ when the UCS is absent. On CS–UCS trials (cell *a*), both the CS and the context gain associative strength,

$$\Delta V_{CS} = \alpha_{CS}\beta_{UCS}[\lambda - (V_{CS} + V_X)]$$

and

$$\Delta V_X = \alpha_X\beta_{UCS}[\lambda - (V_{CS} + V_X)].$$

On CS–noUCS trials (cell *b*), both the CS and the context lose associative strength,

$$\Delta V_{CS} = \alpha_{CS}\beta_{noUCS}[0 - (V_{CS} + V_X)]$$

and

$$\Delta V_X = \alpha_X\beta_{noUCS}[0 - (V_{CS} + V_X)].$$

On noCS–UCS trials (cell *c*), the associative strength of the CS is unchanged, and the context gains associative strength,

$$\Delta V_X = \alpha_X\beta_{UCS}[\lambda - (V_X)].$$

On noCS–noUCS trials (cell *d*), the associative strength of the CS is unchanged, and the context loses associative strength,

$$\Delta V_X = \alpha_X\beta_{noUCS}[0 - (V_X)].$$

Learning will continue until both ΔV_{CS} and ΔV_X approach zero. When learning is at asymptote, $V_{CS} = \Delta P$ if it is assumed that $\beta_{UCS} = \beta_{noUCS}$: The associative strength of the CS, as described by the R-W model, is identical to the contingency between the input and outcome variables as described by ΔP (see Chapman & Robbins, 1990). The R-W model, like the ΔP rule, stipulates that asymptotic strength is independent of the marginal probabilities. The marginal probabilities, however, will affect preasymptotic strength according to the R-W model.

Figure 3 contains learning curves for 200 trials, plotted in blocks of 10 trials. The curves were generated by the R-W model for an arbitrary set of parameter values: $\alpha_{CS} = .9$, $\alpha_X = .2$, $\beta_{UCS} = \beta_{noUCS} = .2$, $\lambda = 1$ on UCS trials and $\lambda = 0$ on noUCS trials. Each curve is the average of 50 runs. The associative strength of the CS and of the context is shown for a strong positive contingency [P(UCS | CS) = .9 and P(UCS | noCS) = .1] and for a strong negative contingency [P(UCS | CS) = .1 and P(UCS | noCS) = .9]. For all curves, P(CS) = P(noCS) = .5. Consider, first, the outcome of the simulation at the end of the 200 training trials. When the contingency is positive, the context–UCS association is weak, providing little competition for the CS–UCS association. The CS becomes excitatory; the organism learns that the CS is associated with the presence of the UCS. When the contingency is negative, there is a strong context–UCS association, which drives the associative strength of the CS negative.

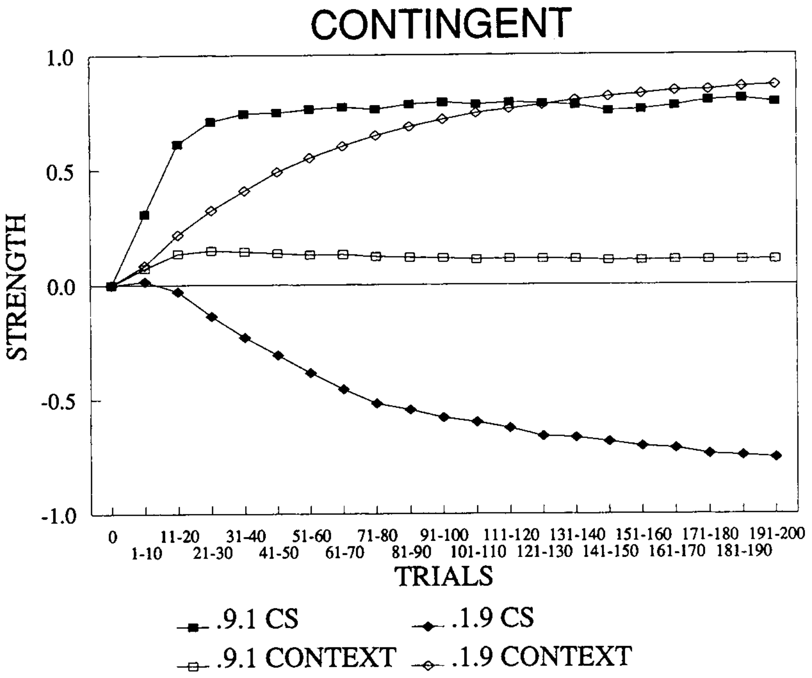

Figure 3. Learning curves for 200 trials plotted in blocks of 10 trials. (The curves were generated by the R-W model for $\alpha_{CS} = .9$, $\alpha_X = .2$, $\beta_{UCS} = \beta_{noUCS} = .2$, $\lambda = 1$ on unconditioned-stimulus [UCS] trials and $\lambda = 0$ on no-unconditioned-stimulus [noUCS] trials. The associative strength of the conditioned stimulus [CS] and of the context is shown for a strong positive contingency [P(UCS | CS) = .9 and P(UCS | noCS) = .1] and for a strong negative contingency [P(UCS | CS) = .1 and P(UCS | noCS) = .9].)

The CS becomes inhibitory; the organism learns that the CS is associated with the absence of the UCS.

The predictions of the R-W model with regard to contingency manipulations have been confirmed many times in the animal-conditioning literature. UCSs interspersed among CS–UCS pairs influence conditioning (e.g., Rescorla, 1968, 1969): When the contingency is positive, conditioning is excitatory; when the contingency is negative, conditioning is inhibitory. As was summarized earlier, parallel results are seen with human contingency judgments. Ratings of contingency are influenced by the size and sign of ΔP: Positive contingencies are judged as positive, and negative contingencies are judged as negative (e.g., Allan & Jenkins, 1980, 1983; Alloy & Abramson, 1979; Baker et al., 1989; Chatlosh et al., 1985; Dickinson et al., 1984; Neunaber & Wasserman, 1986; Shanks, 1985a, 1987; Wasserman, 1990b; Wasserman et al., 1983, 1990, 1993; Wasserman & Shaklee, 1984).

Signaling

In a signaling experiment, as in a contingency experiment, on some trials the UCS is presented without the CS. In a signaling experiment, however, these UCSs are signaled by another CS. That is, on some trials, the UCS is paired with the target CS (CS_T), and on other trials, the UCS is paired with a signal CS (CS_S). According to the R-W model, the CS_S–UCS association should compete with and thereby decrease the strength of the context–UCS associations, resulting in greater associative strength between CS_T and the UCS. A number of investigators have reported, in agreement with the R-W model, that the strength of CS_T is higher in a signaled condition than in a control condition where the extra UCSs are not signaled (e.g., Durlach, 1983; Rescorla, 1984).

Shanks (1986, 1989) has demonstrated the signaling effect in human contingency judgments. Shanks (1986) used a version of the video game described earlier. Subjects were informed that as the tank moved across the screen, it was moving through a minefield. In this situation, firing is the input variable, tank destruction is the outcome variable, and the minefield is the context. In the signaled condition, a jet plane crossed the screen above the tank on trials when the subject did not fire but the tank blew up. That is, all destructions of the tank that occurred when a shell was not fired were signaled by the jet. Judgments about the effectiveness of the shell were greater in the signaled condition than in the control (unsignaled) condition. Signaling tank destruction by the jet protected tank destruction from being attributed to the minefield.

Shanks (1989) demonstrated the signaling effect using a computer version of the free-operant task. Subjects had the option of pressing or not pressing the space bar on a computer keyboard, and the outcome was the presence or the absence of an event on the computer screen. In the signaled condition, every outcome that occurred in the absence of the action was preceded by another event, the signal, which was a short tone. Judgments about the effectiveness of responding were greater in the signaled condition than in the unsignaled condition.

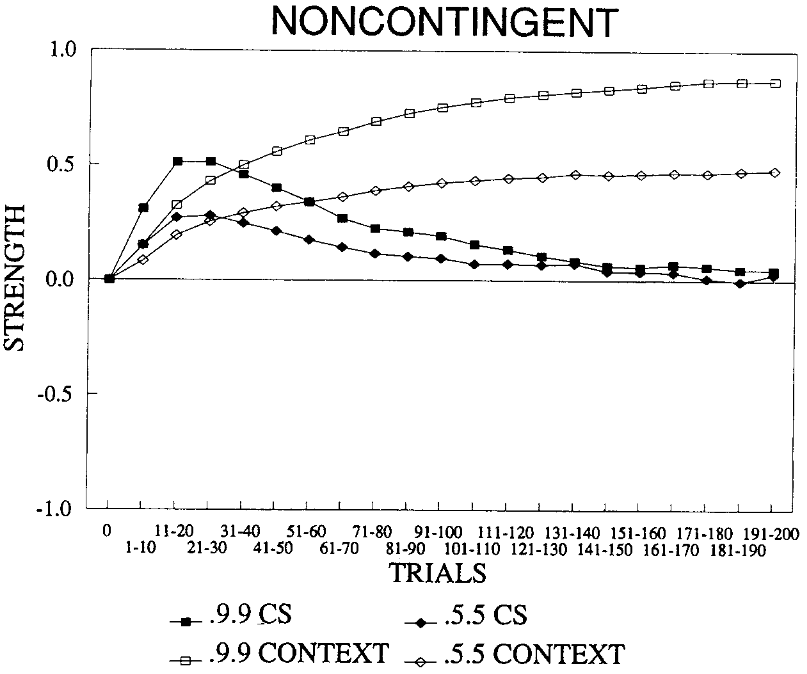

Figure 4. Learning curves generated by the R-W model for two zero contingencies: [P(UCS | CS) = P(UCS | noCS) = .5] and [P(UCS | CS) = P(UCS | noCS) = .9]. (The curves were generated using the same parameter values as the curves in Figure 3. UCS = unconditioned stimulus, CS = conditioned stimulus, noCS = no conditioned stimulus.)

Acquisition

According to the R-W model, for positive and negative contingencies, the strength (excitatory and inhibitory) of the CS–UCS association increases over trials. Simulated learning functions for a positive and a negative contingency are seen in Figure 3. Figure 4 shows the learning functions for two zero contingencies: [P(UCS | CS) = P(UCS | noCS) = .5] and [P(UCS | CS) = P(UCS | noCS) = .9]. The curves in Figure 4 were generated using the same parameter values as the curves in Figure 3. When the contingency is zero, the CS is excitatory on early trials, the associative strength increasing with the probability of the UCS. That is, the R-W model predicts a preasymptotic density bias. The associative strength for a zero contingency is transitory, however, and at asymptote there is no conditioning of the CS.

Incremental learning curves are the norm in animal conditioning, the shape and the asymptote being determined by the contingency between the CS and the UCS. The predictions of the R-W model with regard to the acquisition functions for zero contingency have also been confirmed in the animal-conditioning literature. When the contingency is zero, conditioning is excitatory early in training, the associative strength increasing with the overall frequency of the UCS, but at asymptote there is no conditioning (e.g., Kremer, 1971, 1974; Rescorla, 1972; Rescorla & Wagner, 1972).

In most contingency judgment research involving sequential presentation of the input and outcome events, subjects are asked for only one contingency judgment, after all the information has been presented. In a few studies, subjects have been probed for a contingency judgment a number of times during the sequential presentation of the information (e.g., Baker et al., 1989; Shanks, 1985a, 1987). Acquisition functions similar to the negatively accelerated learning functions predicted by the R-W model were found by Shanks (1985a, 1987) and by Baker et al. (1989) in one experiment but not in another. Judgments increased across trials when the contingency was positive and decreased across trials when the contingency was negative. When the contingency was zero, judgments first increased and then decreased. This positive density bias was greater when the probability of the outcome was high than when it was low.

Shanks (1985a, 1987) and Baker et al. (1989) found that judgments of a zero contingency were negative when the probability of the outcome was low. The R-W model predicts that for a zero contingency, preasymptotic judgments will be more positive when the probability of the outcome is high than when it is low, but the model does not predict negative judgments for a zero contingency. Note, however, that in other contingency studies, zero contingency was not judged as negative even when the probability of the outcome was low (e.g., Chatlosh et al., 1985; Neunaber & Wasserman, 1986; Wasserman et al., 1983).

Rather than probe the subject multiple times during the sequential presentation of the information, Dickinson et al. (1984) varied the number of trials given to different groups of subjects. They also found that for a zero contingency, judgments decreased as the number of trials increased.

Preasymptotic associative strength As was noted earlier, the R-W model, like the ΔP rule, stipulates that for a fixed contingency, asymptotic associative strength is independent of $P(O_1)$ and $P(I_1)$. These marginal probabilities, however, will affect preasymptotic associative strength according to the R-W model.

Allan (1993) showed that preasymptotic dependence of associative strength on $P(O_1)$ and $P(I_1)$ provides an account for data originally reported by Allan and Jenkins (1983; Experiment 3). Allan and Jenkins (1983) varied ΔP, $P(O_1)$, and $P(I_1)$. They found that judgments increased with ΔP but were also dependent on $P(O_1)$ and $P(I_1)$. Allan (1993) presented simulations for these data that were based on the R-W model. The simulations capture the judgment patterns seen in the Allan and Jenkins (1983) data. The R-W model can explain how $P(O_1)$ and $P(I_1)$ influence human judgments of contingency.

Temporal Contiguity

According to the R-W model, decreasing the temporal contiguity between the CS and the UCS effectively decreases the frequency of CS–UCS pairings and increases the frequency of context–UCS pairings. The context–UCS association would be strengthened, thereby weakening the CS–UCS association. Filling the temporal gap would prevent the formation of strong context–UCS associations, thereby protecting the CS–UCS association. Animal studies have shown that temporal contiguity does influence conditioning; conditioning is weakened as the delay between the CS and the UCS is increased (e.g., Williams, 1976). Also, in accord with the R-W model, filling the temporal gap with a stimulus strengthens conditioning in relation to an empty gap (e.g., Reed & Reilly, 1990; Schachtman, Reed, & Hall, 1987).

Wasserman and Neunaber (1986), Reed (1992), and Shanks and collaborators (Shanks, 1989; Shanks & Dickinson, 1991; Shanks et al., 1989) investigated the effect of temporal contiguity on human contingency judgments. They reported that judged contingency decreased as the temporal delay between the input and the outcome was increased. Reed (1992) and Shanks (1989) showed that the effect of a temporal delay was reduced when the gap was filled.

Blocking

The prototypic blocking experiment has two phases. In the first phase, CS_A is paired with the UCS (CS_A–UCS). In the second phase, CS_A is presented simultaneously with a new CS_B, the compound being paired with the UCS (CS_ACS_B–UCS). The prior learning of the CS_A–UCS association competes with and thereby blocks the learning of the new CS_B–UCS association. It is well established in the animal-learning literature that prior CS_A–UCS pairings do attenuate the associative strength of the new component of the compound, CS_B; CS_A blocks CS_B (e.g., Kamin, 1968, 1969a, 1969b).

Blocking in human contingency judgments was demonstrated by Shanks and collaborators (Dickinson et al., 1984; Shanks, 1985b), using the video game. In the blocking condition, the compound trials (firing plus minefield) were preceded by single-element trials (minefield only). That is, subjects were first allowed to see how effective the minefield was without firing the shells. Then the subject was allowed to fire shells. In the control condition, the compound trials were not preceded by minefield-only trials. Subjects in the blocking condition judged the effectiveness of the shells in destroying the tank lower than subjects in the control condition. Learning about the minefield–destruction relationship attenuated subsequent learning about the firing–destruction relationship.

More recently, Chapman and Robbins (1990) demonstrated blocking with a new contingency judgment task and a within-subject design. Subjects viewed a series of trials containing information about a fictitious stock market. Each trial provided information about the price of four stocks and the price of the entire market. The four stocks were P (predictive), N (nonpredictive), B (blocking), and C (control). The trials were divided into two training phases. In Phase 1, there were three types of trials: only Stock P rose in price and so did the market (CS_P–UCS), only Stock N rose in price and the market was stationary (CS_N–noUCS), and none of the stocks rose in price and the market was stationary (noCS–noUCS). In Phase 2, there were three types of compound trials: both Stock P and Stock B rose in price and so did the market (CS_PCS_B–UCS), both Stock N and Stock C rose in price and so did the market (CS_NCS_C–UCS), and none of the stocks rose in price and the market was stationary (noCS–noUCS). At the end of each phase, the subject was asked to judge the extent to which the rise in price of each of the four stocks was predictive of the rise in price of the entire market. Chapman and Robbins (1990) found that Stock B was judged as less predictive than Stock C at the end of Phase 2, even though Stock B and Stock C had the same relationship to the rise of the market. The predictive Stock P blocked the association between Stock B and the market. Chapman (1991) replicated Chapman and Robbins (1990) by using a similar design but a different task concerned with fictitious medical patients.

Waldmann and Holyoak (1992) were critical of the blocking design on the basis that few of the possible trial types are presented. There are eight possible trial types in Phase 2 of a typical blocking experiment: CS_ACS_B–UCS, CS_AnoCS_B–UCS, noCS_ACS_B–UCS, noCS_AnoCS_B–UCS, CS_ACS_B–noUCS, CS_AnoCS_B–noUCS, noCS_ACS_B–noUCS, and noCS_AnoCS_B–noUCS. Waldmann and Holyoak (1992) stated that in the human contingency studies that investigated blocking, only two trial types were presented in Phase 2: CS_ACS_B–UCS and noCS_AnoCS_B–noUCS. This was not an accurate description, however, of the blocking experiments conducted by Shanks and collaborators (Dickinson et al., 1984; Shanks, 1985b). In these experiments, CS_A and noCS_A were mine and no mine, respectively; CS_B and noCS_B were shell and no shell, respectively; and UCS and noUCS were destruction and no destruction, respectively. There were four trial types presented in Phase 2 of these experiments: CS_ACS_B–UCS, CS_ACS_B–noUCS, CS_AnoCS_B–UCS, and CS_AnoCS_B–noUCS. Waldmann and Holyoak (1992) were incorrect about the number and type of trial types presented in these human-blocking studies. Their concern about the importance of missing trial types for theory evaluation was legitimate, however, and should be examined.

Conditioned Inhibition

A well-investigated inhibitory procedure in animal learning is conditioned inhibition (e.g., Rescorla, 1969; Rescorla & Holland, 1977). In conditioned-inhibition training, CS_A is paired with the UCS (CS_A–UCS), and the compound CS_ACS_B is presented without the UCS (CS_ACS_B–noUCS). With this procedure, CS_B becomes inhibitory—it controls a response tendency opposite that of the excitatory CS_A. According to the R-W model, CS_A gains associative strength on CS_A trials. On CS_ACS_B trials, both CS_A and CS_B lose associative strength. Because CS_B never gains strength, the loss of CS_B associative strength on CS_ACS_B trials results in negative associative strength.

Conditioned inhibition in human contingency judgments was demonstrated by Chapman and Robbins (1990), with the fictitious stock market task.[1] There were three stocks: P (predictive), I (inhibition), and N (neutral). In Phase 1, a rise in the price of Stock P was established as predictive of a rise in the market (CS_P–UCS). In Phase 2, CS_P–UCS trials were continued, and two new trials types were added. Stock P and Stock I rose in price and the market was stationary (CS_PCS_I–noUCS), and Stock N rose in price and the market was stationary (CS_N–noUCS). Although I and N had the same relationship to the rise of the market, I was rated as a more negative predictor than N. On CS_PCS_I–noUCS trials, the expected outcome was absent and CS_I became inhibitory. Chapman (1991) replicated conditioned inhibition in contingency judgments with the fictitious medical patients task.

Relative Cue Validity

Wagner, Logan, Haberlandt, and Price (1968; see also Wagner, 1969b) conducted an experiment in which compound CSs were either correlated or uncorrelated with the UCS. For the correlated condition, the compound CS_ACS_X was always paired with the UCS (CS_ACS_X–UCS), and the compound CS_BCS_X was never paired with the UCS (CS_BCS_X–noUCS). For the uncorrelated condition, the compound CS_ACS_X sometimes was paired with the UCS (CS_ACS_X–UCS) and was sometimes not (CS_ACS_X–noUCS), and, similarly, the compound CS_BCS_X sometimes was paired with the UCS (CS_BCS_X–UCS) and sometimes was not (CS_BCS_X–noUCS). Thus, in the correlated condition, CS_A predicted UCS and CS_B predicted noUCS, whereas in the uncorrelated condition, neither CS_A nor CS_B was predictive. Although CS_X had the same relationship to UCS and noUCS in both the correlated and uncorrelated conditions, Wagner et al. (1968) found less conditioning to CS_X in the correlated condition than in the uncorrelated condition. Training with CS_A and CS_B differentially correlated with UCS and noUCS lowered the predictiveness of CS_X compared with training with CS_A and CS_B uncorrelated with UCS and noUCS, even though CS_X had the same relationship to UCS and noUCS in the two conditions. The critical feature in forming an association was not the absolute validity of the cue as a predictor of the outcome, but its relative validity.

Wasserman (1990a), Shanks (1991a, 1991b), and Baker et al. (1993) reported a similar result with human contingency judgments. In the Wasserman (1990a) study, subjects were told to pretend that they were allergists trying to determine the cause of a food-related allergic reaction. They were given descriptions of food combinations and allergic response. In the correlated condition, Foods *A* and *X* (*AX*) always caused an allergic reaction, and Foods B and *X* (*BX*) never did. In the uncorrelated condition, *AX* caused an allergic reaction after half of the meals and did not after the remaining meals, and likewise for *BX*. Thus, in both conditions, Food *X* occurred with an allergic reaction after half of the meals. Subjects rated Food *X* lower in the correlated condition than in the uncorrelated condition. Shanks (1991a, 1991b) used a medical diagnosis situation. In the correlated condition, Symptoms *A* and *X* (*AX*) were paired with the disease, whereas Symptoms B and *X* (*BX*) were never accompanied by the disease. In the uncorrelated condition, Symptoms *AX* and *BX* were paired with the disease on half of the trials and did not accompany the disease on the remaining trials. Subjects rated the relationship between the Symptom *X* and the disease lower in the correlated condition than in the uncorrelated condition.

Baker et al. (1993) explored relative cue validity using a modification of the video game. Subjects were asked to estimate whether camouflaging the tank increased or decreased the likelihood of successfully traversing the minefield. Of interest was the influence of a second predictor, the presence of a spotter plane accompanying the tank, on the estimates of the effectiveness of the camouflage. Under many different contingencies, the presence of a more valid predictor of the outcome reduced the judged effectiveness of a moderate predictor.

Summary

Although associative models were developed primarily to describe results from animal-conditioning experiments, recent studies with human subjects suggest that judgments of contingency can be described by these same models. An associative interpretation of contingency judgments has been supported by a variety of demonstrations that manipulations of the putative CS and UCS have conditioninglike effects. Judgments of contingency are affected by contingency, signaling, and temporal contiguity manipulations. They show changes over trials. They are subject to blocking, conditioned inhibition, and relative cue validity.

There are now many results that indicate that associative learning principles are useful in understanding human contingency judgments. These data strongly support the position taken by Mowrer (1960) over 30 years ago: "We arrive at the conclusion that the causal relationship, as psychically apprehended, is a special case of the more general phenomenon of conditioning or learning by contiguity" (p. 327).[2]

Associative Models: Modifications and Alternatives

The R-W model can account for much of the human contingency data. There are some results, however, that are inconsistent with that model and with the other new associative models as well. In addition, experimental situations that fall outside the boundaries of these models have been explored.

---

[1] The experiment was designed to demonstrate context conditioning in addition to conditioned inhibition. Only those aspects of the design relevant to conditioned inhibition are described.

[2] This quote was brought to my attention by David R. Shanks in an unpublished manuscript.

Trial Order Effects

Trial order is critical in the R-W model. Blocking, for example, should occur if compound CS_ACS_B training follows single CS_A training (referred to as *forward blocking*) but not if compound CS_ACS_B training precedes single CS_A training (referred to as *backward blocking*). In forward blocking, the learning of the CS_A–UCS association in Phase 1 will block the learning of the CS_B–UCS association in Phase 2. In backward blocking, both CS_A and CS_B will gain associative strength in Phase 1. In Phase 2, CS_A will continue to gain associative strength, but the strengthening of the CS_A association will not affect the associative strength of CS_B because associative strength is not altered if the stimulus is not present.

The R-W model also predicts trial order effects in conditioned inhibition. Conditioned inhibition should be more effective if the CS_A–UCS trials are presented before the CS_ACS_B–noUCS trials (forward-conditioned inhibition) rather than in the reverse order (backward-conditioned inhibition). In forward-conditioned inhibition, CS_A gains associative strength on CS_A trials. On CS_ACS_B trials, both CS_A and CS_B lose associative strength. Because CS_B never gains associative strength, the loss on CS_ACS_B trials results in negative associative strength (inhibition). In backward-conditioned inhibition, the associative strength of both CS_A and CS_B would remain unchanged in Phase 1. In Phase 2, CS_A would gain associative strength, but the strengthening of the CS_A association should not affect the strength of the CS_B association because CS_B is not present. CS_B should not become inhibitory in the backward-conditioned inhibition paradigm.

Shanks (1985b) examined trial order in blocking using the video game. In the forward-blocking condition, subjects observed the frequency with which the tank exploded as a result of the mines, before they had the opportunity to fire shells at the tanks. In the backward-blocking condition, the observation period followed the firing period. In the control condition, there was no observation period. The forward- and backward-blocking conditions differed not only in whether the observation period preceded or followed the compound trials but also in the temporal delay between the firing period and the rating of contingency. A second control condition was included to evaluate the contribution of the temporal delay. Shanks (1985b) found that both blocking conditions produced lower ratings of the shell's effectiveness in relation to the appropriate control conditions. Contrary to the R-W model, forward- and backward-blocking procedures had similar effects on contingency judgments.

Chapman (1991) also reported data relevant to trial order effects in human contingency judgments. She used the fictitious health task and compared forward blocking with backward blocking and also forward-conditioned inhibition with backward-conditioned inhibition. Trial order effects were observed for both blocking and conditioned inhibition. The forward procedure was more effective in establishing blocking than was the backward procedure, and the forward procedure produced more inhibition than did the backward procedure. However, contrary to the R-W model, the backward procedures did result in blocking and conditioned inhibition.

Holyoak, Koh, and Nisbett (1989) proposed a nonassociative "theory of classical conditioning based on a parallel, rule-based performance system integrated with mechanisms for inductive learning" (p. 315). This model, like the R-W model, predicted trial order effects in blocking. Holyoak et al. provided a simulation of blocking that was based on their model. According to the simulation, backward blocking would yield identical performance to a control group who received only Phase 2 training. Holyoak et al. emphasized that Kamin's (1968) blocking data were consistent with their model. They did not address the fact that other animal data (e.g., Matzel, Schachtman, & Miller, 1985) and the human-blocking data were not.

The animal-learning data and the human contingency data, which were contrary to the predictions of the R-W model with regard to trial order effects, resulted in suggestions for modifications to the original model.

Rehearsal. To incorporate trial order effects, some investigators (e.g., Baker & Mercier, 1989; Chapman, 1991) postulated that a stimulus that was not presented on a trial might nevertheless be retrieved from memory. Chapman (1991) suggested, for example, that the presentation of CS_A during Phase 2 of the backward-blocking experiment would result in retrieval, on some trials, of memories of similar trial types; that is, of CS_ACS_B. Because CS_A would be present more often than the memories of CS_B, CS_A would gain more of the total strength available, and CS_B would thereby lose strength.

Compounds as unitary events. A different modification to associative models to incorporate trial order effects is to acknowledge a compound stimulus as a unitary event (e.g., Rescorla, 1981; Rescorla & Durlach, 1981; Rudy & Wagner, 1975). Shanks and Dickinson (1987), for example, proposed a model for human contingency judgment that treated a compound (whether of two CSs or of a CS and the context) as a single event with its own associative strength, which was independent of that attached to the components. Although this model could encompass the backward-blocking effects observed with human contingency judgments, it had difficulty with other results, such as relative cue validity. As noted earlier, the need to incorporate "configural cues" was recognized in the animal-learning literature, as was the central role that configural cues had in recently proposed associative models for animal learning (e.g., Wagner & Brandon, 1989).

Experienced Versus Described

In summarizing the research thus far, no distinction was made between studies that presented the input–outcome pairings sequentially(experienced situations) and those that summarized the information in the 2 × 2 matrix (described situations). Some data indicated that similar results were obtained with the two formats (e.g., Baker et al., 1989; Shanks, 1991a, 1991b; Wasserman, 1990a; but see Ward & Jenkins, 1965). At first glance, this might appear problematic for an associative account of human contingency judgments. Associative theories would require the occurrence of temporally distributed events, so that associations could be incremented and decremented trial by trial.

Shanks (1991b) addressed this apparent dilemma and provided a solution. He suggested that organisms possessed an associative learning mechanism that operated in experienced situations, particularly in unfamiliar experienced situations. This mechanism could create new specific causal beliefs, which would then provide a sufficient basis for making judgments in described situations. Specific beliefs would be retrieved from memory to allow judgments to be made in described situations.

Conclusion

Miller, Barnet, and Grahame (in press) recently provided an assessment of the R-W model for animal-learning data. They reviewed the model's predictive successes and failures. They also discussed the model's heuristic value. They concluded "that the model has had and likely will continue to have a distinctly positive influence on the study of simple associative learning in that it has stimulated illuminating research and has contributed to the development of new models." The same could be concluded for the model's heuristic value with regard to human contingency judgments. There are now many results that indicate that associative learning principles are useful in understanding human contingency judgments. Associative models in general, and the R-W model in particular, can account for much of the human contingency data. These models have most certainly stimulated important and exciting research. Not infrequently, these models have prompted the examination of issues unlikely to have been explored outside the framework of associative models.

"Research in the two areas of human and infrahuman learning shares a long history that has focused on elementary associative learning. . .About 20 years ago, however, animal and human learning research became divorced from each other" (Gluck & Bower, 1988, p. 227). The recent trends in human contingency judgments show that a reconciliation is possible. This reconciliation gains additional support from recent research in human perception. Allan and Siegel (e.g., Allan & Siegel, 1986, 1993; Siegel & Allan, 1992; Siegel, Allan, & Eissenberg, 1992) demonstrated the applicability of associative learning principles to an understanding of contingent aftereffects (e.g., the McCollough effect).

References

Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society, 15,* 147–149.

Allan, L. G. (1993). Human contingency judgments: Reanalysis. *Quarterly Journal of Experimental Psychology.* Manuscript submitted for publication.

Allan, L. G., & Jenkins, H. M. (1980). The judgment of contingency and the nature of the response alternatives. *Canadian Journal of Psychology, 34,* 1–11.

Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation, 14,* 381–405.

Allan, L. G., & Siegel, S. (1986). McCollough effects as conditioned responses: Reply to Skowbo. *Psychological Bulletin, 100,* 388–393.

Allan, L. G., & Siegel, S. (1993). McCollough effects as conditioned responses: Reply to Dodwell and Humphrey. *Psychological Review, 100,* 342–346.

Alloy, L. B., & Abramson, L, Y. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General, 108,* 441–485.

Alloy, L. B., & Tabachnik, N. (1984). Assessment of covariation by humans and animals: The joint influence of prior expectations and current situational information. *Psychological Review, 91,* 112–149.

Arkes, H. R., & Harkness, A. R. (1983). Estimates of contingency between two dichotomous variables. *Journal of Experimental Psychology: General, 112,* 117–135.

Baker, A. G., Berbrier, M. W., & Vallee-Tourangeau, F. (1989). Judgments of a 2 × 2 contingency table: Sequential processing and the learning curve. *Quarterly Journal of Experimental Psychology, 41B,* 65–97.

Baker, A. G., & Mazmanian, D. (1989). Selective associations in causality judgments: II. A strong causal relationship may facilitate judgments of a weaker one. *Proceedings of the Eleventh Annual Conference of the Cognitive Science Society* (pp. 538–545). Hillsdale, NJ: Erlbaum.

Baker, A. G., & Mercier, P. (1989). Attention, retrospective processing, and cognitive representations. In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 85–116). Hillsdale, NJ: Erlbaum.

Baker, A. G., Mercier, P., Vallee-Tourangeau, F., Frank, R., & Pan, M. (1993). Selective associations and causality judgments: The presence of a strong causal factor may reduce judgments of a weaker one. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 414–432.

Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 837–854.

Chapman, G. B., & Robbins, S. J. (1990). Cue interaction in human contingency judgment. *Memory & Cognition, 18,* 537–545.

Chatlosh, D. L., Neunaber, D. L., & Wasserman, E. A. (1985). Response–outcome contingency: Behavioral and judgmental effects of appetitive and aversive outcomes with college students. *Learning and Motivation, 16,* 1–34.

Cheng, P. W., & Novik, L. R. (1990). A probabilistic contrast model of causal induction. *Journal of Personality and Social Psychology, 58,* 545–567.

Cheng, P. W., & Novik, L. R. (1991). Cause versus enabling conditions. *Cognition, 40,* 83–120.

Cheng, P. W., & Novik, L. R. (1992). Covariation in natural causal induction. *Psychological Review, 99,* 365–382.

Dickinson, A. (1980). *Contemporary animal learning theory.* Cambridge, England: Cambridge University Press.

Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgement of act–outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology, 36A,* 29–50.

Durlach, P. J. (1983). Effect of signalling intertrial unconditional stimuli in autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 374–389.

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

Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning. *Journal of Experimental Psychology: General, 117,* 227–247.

Goddard, M., & Allan, L. G. (1988). A critique of Alloy and Tabachnik's theoretical framework for understanding covariation assessment. *Psychological Review, 95,* 296–298.

Holyoak, K. J., Koh, K., & Nisbett, R. E. (1989). A theory of conditioning: Inductive learning within rule-based default hierarchies. *Psychological Review, 96,* 315–340.

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

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs, 79*(1, Whole No. 594).

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the production of behavior: Aversive stimulation* (pp. 9-33). Miami, FL: University of Miami Press.

Kamin, L. J. (1969a). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279-296). New York: Appleton-Century-Crofts.

Kamin, L. J. (1969b). Selective association and conditioning. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning* (pp. 42—64). Halifax, Nova Scotia, Canada: Dalhousie University Press.

Kao, S-F., & Wasserman, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*,.

Kremer, E. F. (1971). Truly random and traditional control group procedures in CER conditioning in the rat. *Journal of Comparative and Physiological Psychology, 76*, 441-448.

Kremer, E. F. (1974). The truly random control procedure: Conditioning to the static cues. *Journal of Comparative and Physiological Psychology, 86*, 700-707.

Levin, I. P., Wasserman, E. A., & Kao, S.-F. (1992). Multiple methods for examining the contributions of specific cell information to contingency judgments. *Organizational Behavior and Human Decision Processes.*

Mackintosh, N. J. (1975). 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.* Oxford, England: Clarendon.

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.

McClelland, J. L., & Rumelhart, D. E. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 2. Psychological and biological models.* Cambridge, MA: MIT Press.

Melz, E. R., Cheng, P. W., Holyoak, K. J., & Waldmann, M. R. (1993). Cue competition in human categorization: Contingency or the Rescorla-Wagner learning rule? Comment on Shanks (1991). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1398-1410.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (in press). Assessment of the Rescorla-Wagner model. *Psychological Bulletin.*

Mowrer, O. H. (1960). *Learning theory and the symbolic processes.* New York: Wiley.

Neunaber, D. L., & Wasserman E. A. (1986). The effects of unidirectional versus bidirectional rating procedures on college students' judgments of response-outcome contingency. *Learning and Motivation, 17*, 162-179.

Papini, M. R., & Bitterman, M. E. (1990). The role of contingency in classical conditioning. *Psychological Review, 97*, 396-403.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness conditioned but not unconditioned stimuli. *Psychological Review, 87*, 532-552.

Peterson, C. R., & Beach, L. R. (1967). Man as an intuitive statistician. *Psychological Bulletin, 68*, 29-46.

Reed, P. (1992). Effect of a signalled delay between an action and outcome on human judgment of causality. *Quarterly Journal of Experimental Psychology, 44B*, 81-100.

Reed, P., & Reilly, S. (1990). Context extinction following conditioning with delayed reward enhances subsequent instrumental responding. *Journal of Experimental Psychology: Animal Behavior Processes, 16*, 48-55.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of the 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. (1972). Informational variables in Pavlovian conditioning. In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 6, pp. 42-64). San Diego, CA: Academic Press.

Rescorla, R. A. (1981). Simultaneous associations. In P. Harzem & M. D. Zeiler (Eds.), *Predictability, correlation, and contiguity* (pp. 47-80). New York: Wiley.

Rescorla, R. A. (1984). Signaling intertrial shocks attenuates their negative effect on conditioned suppression. *Bulletin of the Psychonomic Society, 22*, 225-228.

Rescorla, R. A., & Durlach, P. J. (1981). Within-event learning in Pavlovian conditioning. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 81-112). Hillsdale, NJ: Erlbaum.

Rescorla, R. A., & Holland, P. C. (1977). Associations in Pavlovian conditioned inhibition. *Learning and Motivation, 8*, 429-447.

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.

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., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition: Vol. 1. Foundations.* Cambridge, MA: MIT Press

Schachtman, T. R., Reed, P., & Hall, G. (1987). Attenuation and enhancement of instrumental responding by signals for reinforcement on a variable interval schedule. *Journal of Experimental Psychology: Animal Behavior Processes, 13*, 271-279.

Shaklee, H. (1983). Human covariation judgment: Accuracy and strategy. *Learning and Motivation, 14*, 433-448.

Shaklee, H., & Mimms, M. (1982). Sources of error in judging event covariations: Effects of memory demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 2*, 208-224.

Shaklee, H., & Tucker, D. (1980). A rule analysis of judgments of covariation between events. *Memory & Cognition, 8*, 459-467.

Shaklee, H., & Wasserman, E. A. (1986). Judging interevent contingencies: Being right for the wrong reasons. *Bulletin of the Psychonomic Society, 24*, 91-94.

Shanks, D. R. (1985a). Continuous monitoring of human contingency judgment across trials. *Memory & Cognition, 13*, 158-167.

Shanks, D. R. (1985b). Forward and backward blocking in human contingency judgment. *Quarterly Journal of Experimental Psychology, 37B*, 1-21.

Shanks, D. R. (1986). Selective attribution and the judgment of causality. *Learning and Motivation, 17*, 311-334.

Shanks, D. R. (1987). Acquisition functions in contingency judgment. *Learning and Motivation, 18*, 147-166.

Shanks, D. R. (1989). Selectional processes in causality judgment. *Memory & Cognition, 17*, 27-34.

Shanks, D. R. (1990). Connectionism and the learning of probabilistic concepts. *Quarterly Journal of Experimental Psychology, 42A,* 209–237.

Shanks, D. R. (1991a). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 433–443.

Shanks, D. R. (1991b). On similarities between causal judgments in experienced and described situations. *Psychological Science, 2,* 341–350.

Shanks, D. R. (1993). Associative versus contingency accounts of category learning: Reply to Melz, Cheng, Holyoak, and Waldmann (1993). *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 1411–1423.

Shanks, D. R. (in press). Human instrumental learning: A critical review of data and theory. *British Journal of Psychology.*

Shanks, D. R., & Dickinson, A. (1986). The role of selective attribution in causality judgment. In D. J. Hilton (Ed.), *Contemporary science and natural explanations: Commonsense conceptions of causality* (pp. 94–126). Brighton, England: Harvester Press.

Shanks, D. R., & Dickinson, A. (1987). Associative accounts of causality judgment. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 21, pp. 229–261). San Diego, CA: Academic Press.

Shanks, D. R., & Dickinson, A. (1991). Instrumental judgment and performance under variations in action–outcome contingency and contiguity. *Memory & Cognition, 19,* 353–360.

Shanks, D. R., Pearson, S. M., & Dickinson, A. (1989). Temporal contiguity and the judgement of causality by human subjects. *Quarterly Journal of Experimental Psychology, 41B,* 139–159.

Siegel, S., & Allan, L. G. (1992). Pairings in learning and perception: Pavlovian conditioning and contingent aftereffects. In D. Medin, (Ed.), *The psychology of learning and motivation* (Vol. 28, pp. 127–160). San Diego CA: Academic Press.

Siegel, S., Allan, L. G., & Eissenberg, T. (1992). The associative basis of contingent color aftereffects. *Journal of Experimental Psychology: General, 121,* 79–94.

Wagner, A. R. (1969a). Stimulus selection and a "modified contiguity theory." In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 3, pp. 1–41). San Diego, CA: Academic Press.

Wagner, A. R. (1969b). 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., & Brandon, S. E. (1989). Evolution of a structured connectionist model of Pavlovian conditioning (AESOP). In S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Pavlovian conditioning and the status of traditional learning theory* (pp. 149–189). 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.

Wagner, A. R., & Rescorla, R. A. (1972). Inhibition in Pavlovian conditioning: Application to a theory. In R. A. Boakes & M. S. Halliday (Eds.), *Inhibition and learning* (pp. 301–336). San Diego, CA: Academic Press.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121,* 222–236.

Ward, W. C., & Jenkins, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology, 19,* 231–241.

Wasserman, E. A. (1990a). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science, 1,* 298–302.

Wasserman, E. A. (1990b). Detecting response–outcome relations: Toward an understanding of the causal texture of the environment. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 26, pp. 27–82). San Diego, CA: Academic Press.

Wasserman, E. A., Chatlosh, D. L., & Neunaber, D. J. (1983). Perception of causal relations in humans: Factors affecting judgments of response–outcome contingencies under free-operant procedures. *Learning and Motivation, 14,* 406–432.

Wasserman, E. A., Dorner, W. W., & Kao, S. F. (1990). Contributions of specific cell information to judgments of interevent contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16,* 509–521.

Wasserman, E. A., Elek, S. M., Chatlosh, D. L., & Baker, A. G. (1993). Rating causal relations: The role of probability in judgments of response–outcome contingency. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19,* 174–188.

Wasserman, E. A., & Neunaber, D. J. (1986). College students' responding to and rating of contingency relations: The role of temporal contiguity. *Journal of the Experimental Analysis of Behavior, 46,* 15–35.

Wasserman, E. A., & Shaklee, H. (1984). Judging response–outcome relations: The role of response–outcome contingency, outcome probability, and method of information presentation. *Memory & Cognition, 12,* 270–286.

Well, A. D., Boyce, S. J., Morris, R. K., Shinjo, M., & Chumbley, J. I. (1988). Prediction and judgment as indicators of sensitivity to covariation of continuous variables. *Memory & Cognition, 16,* 271–280.

Williams, B. A., (1976). The effects of unsignalled delayed reinforcement. *Journal of the Experimental Analysis of Behavior, 26,* 441–449.

Received February 1, 1993
Revision received May 29, 1993
Accepted May 29, 1993 ■