Learning & Behavior
2005, 33 (2), 131-146

Surprise and change: Variations in the strength of present and absent cues in causal learning

EDWARD A. WASSERMAN
University of Iowa, Iowa City, Iowa

and

LEYRE CASTRO
Universidad de Deusto, Bilbao, Spain

It is said that "absence makes the heart grow fonder." But, when and why does an absent event become salient to the heart or to the brain? An absent event may become salient when its nonoccurrence is surprising. Van Hamme and Wasserman (1994) found that a nonpresented but expected stimulus can actually change its associative status—and in the opposite direction from a presented stimulus. Associative models like that of Rescorla and Wagner (1972) focus only on presented cues; so, they cannot explain this result. However, absent cues can be permitted to change their value by assigning different learning parameters to present and absent cues. Van Hamme and Wasserman revised the Rescorla–Wagner model so that the α parameter is positive for present cues, but negative for absent cues; now, changes in the associative strength of absent cues move in the opposite direction as presented ones. This revised Rescorla–Wagner model can thus explain such otherwise vexing empirical findings as backward blocking, recovery from overshadowing, and backward conditioned inhibition. Moreover, the revised model predicts new effects. For example, explicit information about the absence of nonpresented cues should increase their salience (that is, their negative α value should be larger), leading to stronger associative changes than when no explicit mention is made of cue absence. Support for this prediction is detailed in a new causal judgment experiment in which participants rated the effectiveness of different foods' triggering a patient's allergic reaction. Overall, these and other findings encourage us to view human causal learning from an associative perspective.

> A dog was kept in the stables, and yet, though someone had been in and had fetched out a horse, he had not barked enough to arouse the two lads in the loft. Obviously the midnight visitor was someone whom the dog knew well.

Sherlock Holmes discovered the criminal in one of his celebrated cases (*Silver Blaze*) thanks to his paying attention to the absence of an event: The dog had not barked. This event's nonoccurrence provided the famous fictional detective with relevant and important information. Although countless events are not occurring when something else is occurring, making it more economical to deploy our cognitive resources to process the events that do occur, there are occasions when event absences can be very significant. If two close friends arrive separately at a party, if our plane does not take off after we have already stopped on the runway, or if a normally happy person does not show us her infectious smile, then we are prompted to give extra thought to these episodes. An absent event can in fact be very informative. The nonoccurrence of an event, like its occurrence, can be of very real psychological significance (Hearst, 1991).

Nevertheless, a vast number of events are not occurring at any given moment. It might be intuitively surmised that the absence of an event will be salient only when the event is expected. In an experimental situation, if two stimuli, Cue A and Cue X, are always presented together, and Cue A later occurs, then it is reasonable also to expect the occurrence of Cue X. More formally expressed, when two cues are conjointly presented, an associative link may be established between their central representations. Once this linkage has been formed, the representation of one cue can be activated not only by the presentation of that cue by itself, but also by associative links it has forged with other cues. When Cue A is presented alone, the representation of Cue X will be activated by means of this associative link; because its representation is now activated, the absent cue becomes salient.

Learning and conditioning theories have not traditionally paid much attention to how organisms learn about absent events nor even to whether organisms are sensitive to the nonoccurrence of events except, of course, in the case of the nonoccurrence of expected reinforcers (e.g., Amsel, 1967). However, the modification of re-

L.C. is now at the University of Iowa. We would like to give special thanks to Linda Van Hamme for her exceptional creativity and ingenuity in revising the Rescorla and Wagner model in order to accommodate retrospective revaluation effects. We also thank Lorraine Allan, Mike Aitken, an anonymous reviewer, and Miguel A. Vadillo for their helpful comments on an earlier version of this paper. L.C. was supported by a fellowship from the Spanish Ministry of Education (AP99-14605555). Correspondence should be addressed to E. A. Wasserman or L. Castro, Department of Psychology, E11 Seashore Hall, University of Iowa, Iowa City, IA 52242 (e-mail: ed-wasserman@uiowa.edu or leyre-castroruiz@uiowa.edu).

sponding to a cue—even when this cue is absent—was actually observed in some of the earliest animal conditioning experiments. Sensory preconditioning is one example. After presenting a bell and a light together several times, Brogden (1939) found that, when the bell alone was later paired with shock, the light, although it was never paired with shock, also elicited a conditioned response; that is, the bell–shock pairings led animals to change their responding to an absent cue, the light.

More recent evidence in the animal literature concerning the possibility of learning about absent cues comes from Holland's studies of mediated conditioning. In Holland's (1981) experiments, a tone was paired with food. The tone was later paired with a LiCl injection. When the food was presented again, it had become aversive (see also Ward-Robinson & Hall, 1996). According to Holland, during tone–LiCl pairings, the tone activates a representation of food, which becomes associated with the illness produced by LiCl. The food becomes aversive through this mediated association and its consumption correspondingly decreases. The tone–LiCl pairings evidently prompted the animals to revalue their response to the absent stimulus—food.

Despite these earlier animal studies, it was primarily when associative theories were applied to how humans learn causal relationships that event nonoccurrence became a major challenge for these accounts. How did this story unfold?

Tasks as different as Pavlovian and instrumental conditioning and human causality judgment seem to require organisms to combine information from multiple sources in order to acquire new behavior. Several researchers have noted that the same factors seem to affect behavior in both paradigms: Both were sensitive to temporal contiguity (Shanks & Dickinson, 1991; Shanks, Pearson, & Dickinson, 1989; Wasserman & Neunaber, 1986), to the contingency between events (e.g., Chatlosh, Neunaber, & Wasserman, 1985; Dickinson, Shanks, & Evenden, 1984; Shanks & Dickinson, 1987; Shanks & Dickinson, 1991; Shanks, López, Darby, & Dickinson, 1996; Wasserman, Elek, Chatlosh, & Baker, 1993), and to cue competition; that is, learning depends not only on the relationship between the cue and the outcome, but also on the relationship of other cues with the outcome (e.g., Arcediano, Matute, & Miller, 1997; Chapman, 1991; Chapman & Robbins, 1990; De Houwer, Beckers, & Glautier, 2002; Dickinson et al., 1984; Shanks, 1985).

These empirical parallels encouraged the possibility that the same underlying learning mechanisms might also be operative in studies of associative learning in animals and causal judgment in humans (Allan, 1993; Alloy & Tabachnik, 1984; Dickinson et al., 1984; Gluck & Bower, 1988; Shanks & Dickinson, 1987; Wasserman, 1990b; Young, 1995). Indeed, some authors interpreted conditioning in animals as involving causal perception (i.e., Dickinson, 1980; Killeen, 1981; Mackintosh, 1977). Other authors, such as Kaufman and Bolles (1981), suggested that animals learn about both contiguity and causality, but their responses reflect their knowledge of causal relationships. If one assumes that, during a conditioning procedure, even animals acquire information about the causal structure of their environment, then the correspondence between animal conditioning and human causality learning might not be at all far-fetched.

One of the most representative and influential cases of the application of a conditioning model to explain human causal learning is the Rescorla and Wagner (1972) model (hereafter RW), which was not originally intended to explain learning about absent cues, but to explain the failure of learning about present cues. RW can explain, for example, why when a compound of two cues is paired with an outcome (AX^+) after one of the cues has previously been paired with the same outcome (A^+), a very weak response to Cue X is observed, although Cue X was directly paired with the outcome. This is the well-known blocking effect found by Kamin (1968) with rats, and subsequently reported with human beings as well (e.g., Arcediano et al., 1997; Chapman, 1991; Chapman & Robbins, 1990; De Houwer et al., 2002; Dickinson et al., 1984; Shanks, 1985). The RW model explains the blocking effect as the result of a failure of Cue X to acquire associative strength. According to the RW model, changes in the associative strength of Cue X should proceed according to the following learning rule:

$$\Delta V_X = \alpha\beta(\lambda - \Sigma V_i)$$

so that the change in the associative strength of a Cue X on a particular trial (ΔV_X) is proportional to the difference between the outcome that actually occurs (λ) and the outcome that is predicted by the sum of all present cues (ΣV_i), weighted by the learning parameters α and β, corresponding to the salience of the cue and the outcome, respectively. When there is no discrepancy between what it is predicted and what actually occurs, there is no disparity in the associative strength of the corresponding cue and, therefore, there is no learning. According to RW, in a blocking experiment, at the end of the first phase, the organism has already learned that Cue A perfectly predicts the occurrence of the outcome; that is, the error term ($\lambda - \Sigma V_i$) is equal to 0. When, in the second phase, Cue X is presented along with Cue A, because the occurrence of the outcome is already fully predicted, nothing is learned about Cue X. Thus, RW predicts the blocking effect because changes in the associative strength of Cue X depend on the associative status of other cues presented along with it.

In the last decade, several human studies have consistently found weaker responding to Cue X even when AX^+ trials precede A^+ trials, relative to a condition when the AX^+ trials precede A^- trials—a so-called retrospective revaluation effect (Chapman, 1991; De Houwer et al., 2002; Dickinson & Burke, 1996; Wasserman & Berglan, 1998; Williams, Sagness, & McPhee, 1994). This retrospective revaluation effect poses a considerable problem for RW. According to RW, when AX^+ pairings are first given, the outcome is not predicted by either cue, A or X;

so, participants should learn the A–outcome and X–outcome associations and respond strongly to Cue X in testing. A^+ trials after AX^+ trials should not influence responding to Cue X, because Cue X is not presented during those trials; but, they do. Retrospective revaluation studies indicate that the later pairing of Cue A with the outcome after AX^+ pairings leads to weakened responding to Cue X in testing. Thus retrospective revaluation is a particular problem for RW because this model only allows for changes in the associative value of presented cues—absent cues should not change their associative value.

But, if the premises of the RW model are reconsidered, then the problem can be solved. Van Hamme and Wasserman (1994) observed that the RW model explicitly considers the occurrence and nonoccurrence of the outcome in two different ways: (1) by adjusting the asymptotic level of associative strength that is supportable from λ to 0 (presence vs. absence of the outcome) and (2) by applying different learning-rate parameters on trials with the outcome (β_1) and without the outcome (β_2). However, the RW model does not explicitly consider both the occurrence and nonoccurrence of the antecedent event, the cue. Only a cue that is actually presented is assigned a learning-rate parameter (α); a nonpresented cue is not entered into any associative process and therefore is assigned a zero value.

Van Hamme and Wasserman (1994) suggested that, just as in the case of presented and nonpresented outcomes (the terms β_1 and β_2, respectively), one can assign different learning-rate parameters with nonzero values to presented and nonpresented cues (the terms α_1 and α_2, respectively). As we will see, this theoretical maneuver allows the RW model to explain many otherwise problematic findings.

Updating the Strength of Both Present and Absent Cues: Van Hamme and Wasserman (1994)

Until recently, researchers had not seriously considered the possibility that people who are engaged in evaluating a causal relationship might not only change their ratings of the cues that are *presented* on an informational trial, but they might also change their ratings of cues that are *not* presented. Nevertheless, it seems reasonable to believe that, if the presented cue were to gain strength in light of evidence, then nonpresented cues might immediately and correspondingly lose strength; conversely, if the presented cue were to lose strength in light of evidence, then nonpresented cues might immediately and correspondingly gain strength.

In fact, work on contingency learning has shown that participants' judgments reflect the use of information about *both* the occurrence and the nonoccurrence of *both* the cue and the outcome. Specifically, when participants are provided with: (a) the frequency of occurrence of both a cue and an outcome together, (b) the frequency of occurrence of the cue without the outcome, (c) the frequency of occurrence of the outcome when the cue is absent, and (d) the frequency of nonoccurrence of both the cue and the outcome, participants' utilization of this information has generally been found to reflect the following biased weighting of the four types of information: a > b > c > d (Anderson & Sheu, 1995; Kao & Wasserman, 1993; Levin, Wasserman, & Kao, 1993; Mandel & Lehman, 1998; Wasserman, Dorner, & Kao, 1990).

Table 1
Van Hamme and Wasserman's (1994) Experimental Design

	Condition 1.00 Outcome		Condition 0.50 Outcome		Condition 0.00 Outcome	
AX	8 Yes	0 No	6 Yes	2 No	4 Yes	4 No
BX	0 Yes	8 No	2 Yes	6 No	4 Yes	4 No

Note—AX and BX are compounds of two cues that were paired with the occurrence (Yes) or nonoccurrence (No) of the outcome.

The key purpose of the Van Hamme and Wasserman (1994) project was to see whether participants would update on every trial the associative value of not only the present cues but also the absent cues. During the experimental task (Wasserman, 1990a), participants rated the plausibility of three different foods' (e.g., shrimp, strawberries, and peanuts) being the source of a hypothetical patient's allergic reaction (0 = *definitely not*; 4 = *possibly*; 8 = *definitely the cause of the reaction*). On any given trial, two of three different foods (A, B, or X) were given as possible causes of the allergic reaction, as shown in Table 1. Both AX and BX trials were randomly presented, half of which were paired with the occurrence of the allergic reaction and half with the nonoccurrence of the allergic reaction. The precise compound cue–outcome pairings differed in each of three experimental conditions. In Condition 1.00, each AX trial was paired with the occurrence of the allergic reaction and each BX trial was paired with the nonoccurrence of the allergic reaction; the score of 1.00 represents the difference in the probability of the allergic reaction given AX (1.00) minus the probability of the allergic reaction given BX (0.00). In Condition 0.50, AX was assigned a probability of the allergic reaction of 0.75 and BX was assigned a probability of the allergic reaction of 0.25, a difference of 0.50. And, in Condition 0.00, AX was assigned a probability of the allergic reaction of 0.50 and BX was assigned a probability of the allergic reaction of 0.50, a difference of 0.00.

This experimental design was inspired by the classical and instrumental conditioning work developed by Wagner, Logan, Haberlandt, and Price (1968) and later extended by Wasserman (1974), who showed that, although Cue X had been paired with reinforcement equally often in these experimental conditions, rats', rabbits', and pigeons' conditioned responding to Cue X systematically decreased as the differential predictiveness of AX and BX increased. Presumably, as the A and B cues came to better signal the occurrence and nonoccurrence of reinforcement, Cue X more poorly competed with A and B for the control of conditioned responding.

Indeed, in this human judgment experiment, we found that participants' final ratings of the causal efficacy of

common Cue X systematically fell as the AX−BX difference increased from 0.00 to 0.50 to 1.00. Although the actual relation of Cue X to the allergic reaction did not change across the three conditions, the causal ratings of Cue X decreased as Cue A became a more reliable cause of the occurrence of the allergic reaction and Cue B became a more reliable predictor of its nonoccurrence. This parallel between humans' causal ratings and animals' conditioned responses was certainly important and it greatly increased our confidence that similar psychological processes may be at work in these two very different situations. But, the central interest in this experiment was to see whether participants lawfully changed their ratings of both the presented and nonpresented cues. In order to do so, we had participants rate all *three* foods (A, B, and X) on *every* trial, meaning that one of the two distinctive cues (either A or B) was *not* presented. We then compared participants' ratings on trial $n - 1$ and on trial n to determine the momentary impact of outcome occurrence and nonoccurrence on causal ratings.

The results for the distinctive A and B cues were clear. When a cue was given on trial $n - 1$, reinforcement on that trial increased ratings of that stimulus on trial n by a mean of 1.23; when a cue was given on trial $n - 1$, nonreinforcement on that trial decreased ratings of that stimulus on trial n by a mean of 0.57. These results were completely unproblematic and were expected by associative learning models like that of Rescorla and Wagner (1972). However, when a cue was *not* given on trial $n - 1$, reinforcement on that trial decreased ratings of that stimulus on trial n by a mean of 0.86; when a cue was *not* given on trial $n - 1$, nonreinforcement on that trial increased ratings of that cue on trial n by a mean of 0.22. These results confirmed our hypothesis that nonpresented cues are reevaluated in a way that contrasts with the changes in ratings of presented cues. As well, the changes in ratings for presented cues were of greater absolute value than those for nonpresented cues: 1.23 versus 0.86 and 0.57 versus 0.22. Because associative learning models like RW focus only on the signaling cues that are actually presented, these results were quite problematic and required serious theoretical consideration.

Revised Rescorla–Wagner Model: Van Hamme and Wasserman (1994)

One way to deal with this empirical embarrassment is to reconsider the premises of associative learning theory and to modify it in light of the evidence. As explained above, the RW model does *not* explicitly consider both the occurrence and nonoccurrence of a *signaling* cue. Only a signaling cue that is actually presented is assigned a learning-rate parameter (α); a nonpresented signaling cue is not entered into any associative equations, presumably because nonpresented stimuli have no psychological significance. But, what if nonpresented stimuli *do* have psychological significance? We suggested that, just as in the case of presented and nonpresented *reinforcing* stimuli (the terms β_1 and β_2, respectively), one could assign different learning-rate parameters with nonzero values to presented and nonpresented *signaling* cues (the terms α_1 and α_2, respectively). This theoretical maneuver not only introduces symmetry into the treatment of cues and outcomes, but it also allows the RW model to embrace the very results on causal judgment change scores that at first had appeared to be so embarrassing.

In order to see how the revised RW model accomplishes this theoretical feat, consider both the original and the revised models. As shown in Table 2, there are only two equations in the original model, corresponding to cases in which a cue is given and the outcome follows and to cases in which a cue is given and the outcome does not follow. The revised model adds two equations, corresponding to cases in which a signaling stimulus is not presented and the reinforcing stimulus is, and to cases in which neither the signaling stimulus nor the reinforcing stimulus is presented. If the α_2 parameter is 0, then the revised model reduces to the original model; however, if the α_2 parameter takes on a *negative* value, then changes in the associative strength of nonpresented stimuli will move in the *opposite* direction as changes in the associative strength of presented stimuli. If the absolute value of α_2 is less than that of α_1—amounting to the plausible assumption that presented stimuli are more salient than are nonpresented stimuli—then, after reinforcement, presented stimuli will increase in strength and nonpresented stimuli will decrease in strength, but to a smaller degree. Oppositely, after nonreinforcement, presented stimuli will decrease in strength and nonpresented stimuli will increase in strength, again to a smaller degree. The data from the Van Hamme and Wasserman (1994) experiment perfectly correspond to these predictions. What else can the revised model do?

Trial-by-Trial Backward Blocking and Backward Conditioned Inhibition: Wasserman, Kao, Van Hamme, Katagiri, and Young (1996)

We suspected that the revised model might be able to explain at least two other empirical findings that have challenged the original model: backward blocking and backward conditioned inhibition. Each of these effects entails the apparent retrospective revaluation of a cue, something that is unpredicted by the original model.

In backward blocking, a first phase of training involves AX being followed by the outcome. Then, a sec-

Table 2
Original and Revised Rescorla–Wagner Models

Original Rescorla–Wagner Model
Cue present–outcome present (cell A): $\Delta V_X = \alpha_1\beta_1(\lambda - \Sigma V_i)$
Cue present–outcome absent (cell B): $\Delta V_X = \alpha_1\beta_2(0 - \Sigma V_i)$
Revised Rescorla–Wagner Model
Cue present–outcome present (cell A): $\Delta V_X = \alpha_1\beta_1(\lambda - \Sigma V_i)$
Cue present–outcome absent (cell B): $\Delta V_X = \alpha_1\beta_2(0 - \Sigma V_i)$
Cue absent-outcome present (cell C): $\Delta V_X = \alpha_2\beta_1(\lambda - \Sigma V_i)$
Cue absent–outcome absent (cell D): $\Delta V_X = \alpha_2\beta_2(0 - \Sigma V_i)$

Note—α_1 = learning rate parameter for Cue X present, α_2 = learning-rate parameter for Cue X absent, β_1 = learning-rate parameter for outcome present, and β_2 = learning-rate parameter for outcome absent. V_is are summed only for the cues that are present on a given trial.

ond phase of training is given involving Cue A alone, also followed by the outcome. If causal ratings of Cue X are lower at the end of Phase 2 than at the end of Phase 1, then backward blocking can be said to have occurred. But, this effect cannot be predicted by the original model; associative strength to Cue X cannot change during Phase 2 because it is not given. The revised model does predict backward blocking, because A^+ trials are trials in which the representation of Cue X will be activated; because Cue X is absent, these trials will effectively reduce associative strength to Cue X.

In backward conditioned inhibition, a first phase of training involves AX followed by no outcome. Then, a second phase of training is given involving Cue A alone now followed by the outcome. If causal ratings of Cue X are lower at the end of Phase 2 than at the end of Phase 1, then backward conditioned inhibition can be said to have occurred. This effect cannot be predicted by the original model either; associative strength to Cue X cannot change during Phase 2 because it is not given. The revised model does predict backward conditioned inhibition, again because A^+ trials are trials in which the representation of Cue X will be activated; because Cue X is absent, these trials will effectively reduce associative strength to Cue X.

Chapman (1991) had previously reported both backward blocking and backward conditioned inhibition in a stock market prediction task. Wasserman et al. (1996) sought to both extend the generality of Chapman's results to a different judgment task and to document changes in participants' ratings on a trial-by-trial basis. We therefore examined both forward and backward blocking as well as both forward and backward conditioned inhibition within a clinical setting. We also provided participants with information from all four cells of the contingency table in each type of problem. In order to allow trial-by-trial ratings of all relevant stimuli on every trial, after each piece of information, participants were asked to rate the effect of each of three foods on a person's allergic reaction from −100 (*makes very unlikely*) to 0 (*doesn't affect the likelihood*) to +100 (*makes very likely*) the allergic reaction. Only one or two of these foods were actually given on each trial, again allowing us to see whether both presented and nonpresented cues would be affected by the presentation or nonpresentation of outcomes.

Ratings at the end of Phase 1 and Phase 2 of training showed that both forward blocking and forward conditioned inhibition had occurred. Ratings of Cue X were lower after A^+/AX^+ training (21.76) than after AX^+ training only (68.54), indicating forward blocking; ratings of Cue X were lower after A^+/AX^- training (−33.98) than after AX^- training only (1.93), indicating forward conditioned inhibition. These results are consistent with the original RW model. More interestingly, both backward blocking and backward conditioned inhibition were also obtained. Ratings of Cue X were lower after AX^+/A^+ training (15.09) than after AX^+ training only (68.54), indicating backward blocking; ratings of X were lower after AX^-/A^+ training (−27.22) than after AX^- training only (1.93), indicating backward conditioned inhibition. These results are not predicted by the original RW model, but they are predicted by the revised model.

Also noteworthy are the trial-by-trial ratings in the backward blocking and backward conditioned inhibition conditions. Figure 1 shows participants' ratings of Cue X during the whole set of trials in both the backward blocking and backward conditioned inhibition conditions along with the ratings that were predicted by the original and revised RW models. The values chosen for the predictions of the revised RW model were based on independent research by Kao (1993): α_1, α_2, β_1, and β_2 were 0.70, −0.40, 0.50, and 0.40, respectively. In the backward blocking condition, the original RW model predicts that ratings of Cue X should increase during Phase 1, but remain constant throughout the second phase, because it is not presented during that

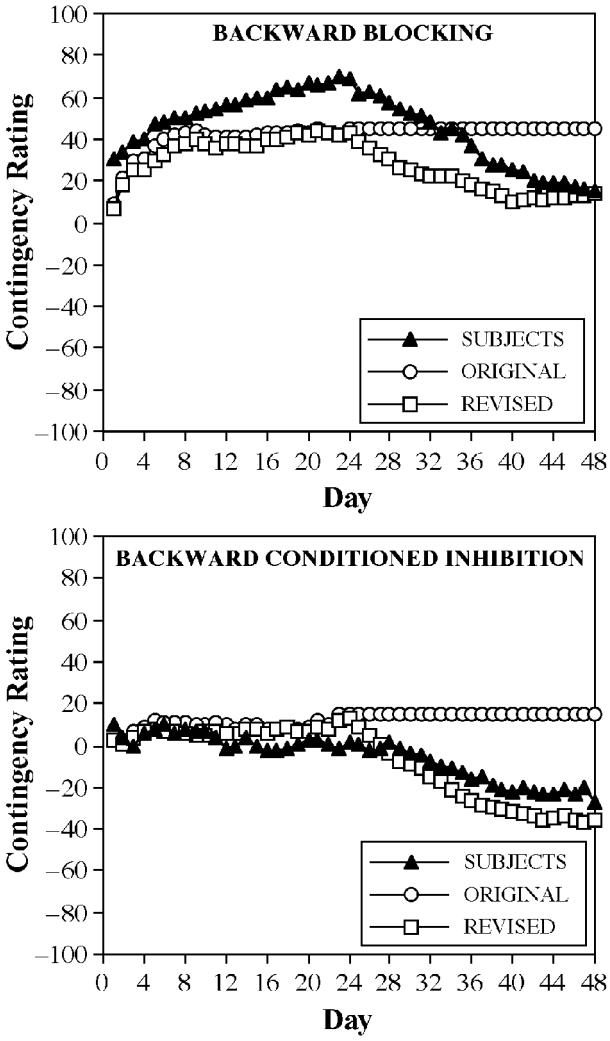

Figure 1. Trial-by-trial ratings of participants in the backward blocking and backward conditioned inhibition conditions of Van Hamme and Wasserman (1994) along with those predicted by Rescorla and Wagner's model and by the revision of their model by Van Hamme and Wasserman.

period; however, participants' ratings of Cue X progressively decreased during Phase 2, as predicted by the revised model. In the backward conditioned inhibition, the original RW model predicts that ratings of Cue X should also remain constant throughout the second phase, in which this cue is not presented; however, participants' ratings of Cue X progressively decreased during Phase 2, again confirming the prediction of the revised model.

Table 3
Wasserman and Berglan's (1998) Experimental Design

Training		
Phase 1	Phase 2	Test
AW^+	A^+	W
BX^+	–	X
CY^+	C^-	Y
DZ^-	D^-	

Note—A, B, and C refer to the competing cues and W, X, and Y refer to the target cues; D and Z are the filler cues. The $^+$ denotes reinforcement, the $^-$ denotes nonreinforcement, and the – indicates that none of the individual cues in the compound BX was presented in that training phase.

The Role of Within-Compound Associations: Wasserman and Berglan (1998)

As we have just seen, human research participants attend to and evaluate both present and absent cues. But, on any given trial, very few events are actually presented, whereas countless other events are not. To which nonpresented cues do participants attend? It is reasonable to suggest that participants attend only to those cues whose memories or representations are active on a particular trial. Van Hamme and Wasserman (1994) suggested how those representations might be activated. First, people can be given instructions to consider particular cues, even though they may not have been presented on a given trial. Second, contextual or discrete cues that had been presented along with the target cues can also trigger representations of those now-absent cues. So, when two cues are paired, a within-compound association will be formed between them, which will then allow the presentation of one of the cues to activate the representation of the other cue.

Dickinson and Burke (1996) explored the possibility that, for retrospective revaluation to occur, participants must form within-compound associations. It should be noted that the formation of a within-compound association is necessary to explain backward blocking (where AW^+ trials precede A^+ trials) or recovery from overshadowing (where CY^+ trials precede C^- trials), but it is not required to explain forward blocking (where A^+ trials precede AW^+ trials) or superconditioning (where C^- trials precede CY^+ trials) (for further discussion of this issue, see Aitken, Larkin, & Dickinson, 2001; Larkin, Aitken, & Dickinson, 1998; Melchers, Lachnit, & Shanks, 2004). Dickinson and Burke found that retrospective revaluation occurred when it was possible for participants to form within-compound associations, but not when it was impossible for them to do so.

Unfortunately, the retrospective revaluation effect that Dickinson and Burke (1996) reported could have involved not only backward blocking but also recovery from overshadowing. When two compound cues are first presented, both followed by the outcome (e.g., AW^+, CY^+) and then in a second training phase only one cue of each compound is presented—one followed by the outcome (e.g., A^+), but not the other (e.g., C^-)—lower ratings of Cue W than Cue Y could be due to a decrement in the strength of Cue W (backward blocking), to an increment in the strength of Cue Y (recovery from overshadowing), or to the sum of both effects. In addition, Dickinson and Burke did not independently assess participants' learning and memory of the putative within-compound associations. These two design shortcomings prompted us to conduct our next study.

First, we aimed to clarify the nature of the retrospective revaluation that was reported by Dickinson and Burke (1996). We hoped to see whether backward blocking, recovery from overshadowing, or both effects were involved in their reported results. As shown in Table 3, a control condition, BX, was included, without training of any of the cues in the compound during Phase 2, so that the causal ratings of Cue X could serve as a baseline for other critical comparisons. If ratings of Cue W were lower than ratings of Cue X at the end of Phase 2, then this result would define backward blocking. If ratings of Cue Y were higher than ratings of Cue X at the end of Phase 2, then this result would define recovery from overshadowing. Cues D and Z were included in order to force participants to discriminate among the compound stimuli in Phase 1.

To independently assess the formation and retention of within-compound associations after the experimental task had been learned, at the end of the experiment, we asked participants to report which cues had been presented with one another during the first phase of training. We hypothesized that those participants who had learned and remembered the compounds would show backward blocking and recovery from overshadowing, whereas those participants who had not learned and remembered the compounds would not show these retrospective revaluation effects.

As predicted by the revised RW model, both backward blocking and recovery from overshadowing occurred. Final ratings of Cue W were lower than final ratings of Cue X (4.75 and 5.63, respectively)—backward blocking. In addition, final ratings of Cue Y were higher than final ratings of Cue X (6.81 and 5.63, respectively)—recovery from overshadowing. Because final ratings were ordered $W < X < Y$, it is likely that both backward blocking and recovery from overshadowing contributed to Dickinson and Burke's (1996) reported results.

Furthermore, those participants who correctly identified the stimulus compounds reliably exhibited the $W < X < Y$ pattern of results ($W = 4.42$, $X = 5.63$, and $Y = 6.92$); but, those participants who did not correctly identify the stimulus compounds did not respond differentially to the target cues ($W = 6.00$, $X = 5.60$, and $Y =$

6.40). These data thus are consistent with the notion that within-compound associations mediate the retrospective revaluation of causal judgments. More recently, the studies of Aitken et al. (2001) and Melchers et al. (2004) have further demonstrated that the formation and preservation of within-compound associations between the cues that were presented in compound are necessary for retrospective revaluation effects to occur.

In summary, the revised RW model correctly predicts trial-by-trial changes in the strength of nonpresented cues in the form of backward blocking, backward conditioned inhibition, and recovery from overshadowing. The revised RW model also correctly predicts the participation of within-compound associations in the retrospective revaluation effects. In the present study, we delved more deeply into the RW model.

The Present Experiment

As we have just seen, recent research suggests that absent cues are behaviorally relevant when they are expected to occur due to within-compound association with other cues that are actually presented. Presented cues activate the representations of nonpresented cues with which they have formed within-compound associations, allowing the nonpresented cues to be involved in the learning process.

In human studies, one might more strongly activate the representations of absent cues through verbal instructions, as Van Hamme and Wasserman (1994) suggested. In the present experiment, we hypothesized that explicit information about the absence of nonpresented cues would increase their salience, leading to stronger retrospective revaluation effects than when no explicit mention is made about the absence of nonpresented cues. This prediction emerges directly from the revised RW model. Increased salience of absent cues translates into a larger negative α value; therefore, increases or decreases in the associative value of absent cues should also be larger, yielding stronger retrospective revaluation effects.

Using the same design as in the prior study, in the present experiment, we tried to modify the salience of absent cues through instructions: we explicitly told participants which of the cues were or were not presented on a particular trial. Tassoni (1995) had earlier used this strategy, but he was not interested in documenting retrospective revaluation effects, only in showing that the quantity of information provided by giving notice about the absence of a cue is greater than is the quantity of information arising from the mere nonoccurrence of a cue. Our clear expectation was that providing explicit information about absent cues would produce stronger retrospective revaluation effects.

METHOD

Participants

A total of 91 students at the University of Iowa received course credit for their voluntary participation. Random assignment resulted in 48 participants in the uninformed group and 43 in the informed group.

Materials

The experiment was programmed using PsyScope Version 1.0.2 (Cohen, MacWhinney, Flatt, & Provost, 1993) on four iMac computers. Eight different foods (*yogurt, mushrooms, carrots, grapes, walnuts, noodles, oranges,* and *chicken*) served as cues. As in Wasserman and Berglan (1998), the cues were counterbalanced following a between-participants Latin square design, which ensured that every food was equally often assigned to each cue role and that each competing cue was equally often paired with each target cue. The cues were written in black type at the top of the screen. During training, the uninformed group observed the names of one or two foods on each trial: the two foods of each pair in Phase 1 and one food of each pair in Phase 2. The informed group observed the names of both foods of each pair in both Phases 1 and 2: Below the name of each food, the word *yes* or *no* was written, indicating whether or not the particular food had been eaten by the hypothetical patient. We reasoned that telling participants that a particular food had not been eaten (Group Informed) would make such information more prominent than when a particular food simply did not appear (Group Uninformed). In each group, the outcomes were *Allergic Reaction* on reinforced trials and *No Allergic Reaction* on nonreinforced trials.

Procedure

From 1 to 4 participants were studied concurrently on each of 4 identically configured computer workstations. All the procedures, instructions, and questionnaires were the same as in Wasserman and Berglan (1998). Each participant sat in front of a workstation and was instructed to fill in the first paper-and-pencil questionnaire, in which they were asked to provide an initial evaluation of the probability that each of the foods could cause an allergic reaction in an ordinary individual. This first rating, as well as the subsequent ones, was made on a scale of 1 to 9. Then, participants were introduced to a scenario in which they played the role of an allergist trying to discover which foods would or would not cause an allergic reaction in a specific individual, Mr. X (see Appendix for instructions). In order to perform that task, participants had to study different daily allergy tests in which Mr. X had or had not suffered an allergic reaction after eating certain foods. The individual food or compound of foods appeared at the top of the screen, and after 3 sec, participants were required to predict whether Mr. X would develop an allergic reaction. Once the participants pressed "1" to indicate *yes* or "3" to indicate *no*, the actual outcome appeared below. After 3 sec, participants could proceed to the next trial.

In the first training phase, participants observed four compounds that were presented 30 times each. Compounds AW, BX, and CY were always paired with the allergic reaction; Compound DZ was never paired with the allergic reaction. The 120 trials in Phase 1 were presented in a block-randomized order; each compound cue was seen six times every 24 trials. After the first training phase, participants again had to evaluate the eight foods, except that they were now asked to use all of the information that they had received to evaluate the likelihood of each food causing an allergic reaction in Mr. X. Participants had no access to their initial ratings.

During the second training phase, Cue A was always paired with the allergic reaction; Cues C and D were never paired with the allergic reaction. In this second phase, each of the cues in isolation was presented to the uninformed group. The informed group, however, observed the competing cues along with the target cues that had been presented in compound with them during Phase 1. Below each of the foods, the words *yes* or *no* appeared, indicating whether or not the patient had eaten it; that is, indicating its presence or absence on any given trial. For example, the two foods corresponding to the compound AW had been presented during Phase 1. Now, during Phase 2, the names of both foods were shown to the informed group; the word *yes* appeared below Cue A, indicating that Cue A had been presented on that trial and the word *no* appeared below Cue W indicating that the Cue W had been absent on that trial. Each of the three types of trials was presented 30 times in a block-randomized order for a total

Table 4
Mean Ratings for All Cues in Each Rating Period in Group Uninformed and Group Informed

	Rating Period 1		Rating Period 2		Rating Period 3	
Cue	*M*	*SE*	*M*	*SE*	*M*	*SE*
			Group Uninformed			
A	2.20	.20	6.84	.29	8.77	.17
W	**2.31**	**.21**	**6.82**	**.31**	**5.31**	**.34**
B	2.46	.19	6.77	.29	5.95	.29
X	**2.48**	**.22**	**6.55**	**.33**	**6.24**	**.29**
C	2.35	.20	6.95	.30	1.08	.04
Y	**2.53**	**.21**	**6.84**	**.32**	**6.97**	**.33**
D	2.91	.27	1.35	.15	1.08	.05
Z	2.37	.17	1.95	.31	2.31	.35
			Group Informed			
A	2.23	.23	6.60	.37	8.52	.29
W	**2.26**	**.22**	**7.36**	**.25**	**2.89**	**.38**
B	2.21	.22	7.13	.29	5.02	.33
X	**2.36**	**.20**	**6.86**	**.33**	**4.71**	**.38**
C	2.31	.21	7.21	.31	2.10	.36
Y	**2.42**	**.23**	**7.10**	**.30**	**6.76**	**.45**
D	2.39	.24	1.31	.18	1.84	.35
Z	2.18	.24	1.05	.03	1.89	.37

Note—Cues W, X, and Y, in boldface, are target cues; Cues A, B, and C are their associated competing cues; Cues D and Z are filler cues. Rating Period 1 provided initial ratings prior to training. Rating Period 2 followed Phase 1 training (AW^+, BX^+, CY^+, DZ^-). Rating Period 3 followed Phase 2 training (A^+, C^-, D^-).

of 90 trials in the second training phase; each type of trial was seen six times every 18 trials. Phase 2 was followed by a third rating period that was identical to the second.

Once participants had completed the third rating period, they were given a postexperimental questionnaire in which they were presented with 16 pairs of foods; their task was to circle the four two-food compounds that had actually been presented during the experiment.

Preanalysis of the Data

To avoid the influence of possible preexperimental biases to the foods, which might produce a ceiling effect prior to the experimental learning stages, we eliminated from the analysis data from participants who gave extremely high initial ratings (a rating equal to or greater than 7 to any of the cues in Rating Period 1). With this criterion, the data from 2 participants in each group were eliminated. We also eliminated the data from 4 participants who invalidly or incompletely marked their questionnaires: 3 participants in the informed group and 1 participant in the uninformed group. In total, the data from 8 participants were eliminated from the analysis. An alpha level of .05 was adopted for all tests of statistical significance. When multiple comparisons were performed, Bonferroni correction was used to set the alpha level.

RESULTS

Causal ratings of all eight cues in each of the three rating periods are provided in Table 4. In both groups, initial ratings, prior to any experimental training, were low and similar for all cues. After Phase 1 training, the ratings of reinforced cues increased, while the ratings of nonreinforced cues remained low in both groups, confirming that training proceeded smoothly. More interesting results came after the crucial Phase 2 training. Table 4 and Figure 2 display the final ratings of Target Cues W, X, and Y. Participants rated the target cues differently although all three had been associated with the same reinforcement schedule throughout the experiment. Moreover, the differences among the target cues were larger in the informed group than in the uninformed group.

A 2 (group: informed vs. uninformed) × 3 (rating period: 1 vs. 2 vs. 3) × 3 (cue: W vs. X vs. Y) analysis of variance (ANOVA) revealed a significant main effect of rating period [$F(2,162) = 279.45$, $MS_e = 4.73$], simply showing that ratings of the target cues, low at the beginning of the experiment, increased due to the reinforcement of the cues, as expected. There was also a significant main effect of cue [$F(2,162) = 15.70$, $MS_e = 3.56$] and a significant rating period × cue interaction [$F(4,324) = 25.52$, $MS_e = 2.07$], showing that ratings of the target cues changed differently during the different rating periods. A significant Group × Rating Period interaction [$F(2,162) = 10.86$, $MS_e = 4.73$] and a significant group × rating period × cue interaction [$F(4,324) = 4.73$, $MS_e = 2.07$] were also found, indicating that changes among the target cues in the three rating periods differed in the informed and uninformed groups.

In the uninformed group, simple effects analyses confirmed that significant differences existed among the cues in Rating Period 3 [$F(2,162) = 7.04$, $MS_e = 4.45$], but not in Rating Period 1 [$F(2,162) < 1$] or 2 [$F(2,162) < 1$]. Figure 2 shows that the final causal ratings of the three cues in Group Uninformed were in the expected order: W < X < Y. Planned comparisons revealed that ratings of Cue W were lower than ratings of Cue Y [$F(1,81) = 11.27$, $MS_e = 5.54$], replicating the typically reported retrospective revaluation effect. The significant difference between ratings of Cue W and Cue X [$F(1,81) = 5.94$, $MS_e = 3.29$] indicated that backward blocking is involved in this effect. The difference between ratings of Cue Y and Cue X, which would indicate recovery from overshadowing, was in the expected direction, but it fell just short of statistical significance in this analysis [$F(1,81) = 2.66$, $MS_e = 4.53$, $p = .10$].

In the informed group, with explicit information about the absence of cues during Phase 2, simple effects analyses also confirmed that significant differences existed among the cues in Rating Period 3 [$F(2,162) = 31.91$, $MS_e = 4.46$], but not in Rating Period 1 [$F(2,162) < 1$] or 2 [$F(2,162) = 1.28$, $MS_e = 1.84$]. Figure 2 shows that the final causal ratings of the three cues in Group Informed were also in the expected order: W < X < Y. Planned comparisons revealed that ratings of Cue W were lower than ratings of Cue Y [$F(1,81) = 51.27$, $MS_e = 5.54$], replicating the retrospective revaluation effect in this group. The difference between ratings of Cue W and Cue X was significant [$F(1,81) = 18.98$, $MS_e = 3.29$], documenting backward blocking. Also significant was the difference between ratings of Cue Y and Cue X [$F(1,81) = 17.65$, $MS_e = 4.53$], documenting recovery from overshadowing.

Moreover, as shown in Figure 2, these effects seemed to be even stronger in the informed group than in the uninformed group. To further analyze the difference in the

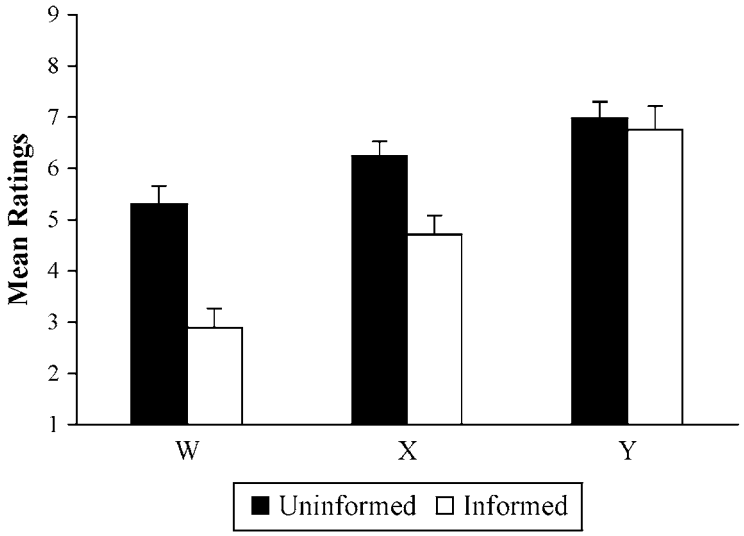

Figure 2. Mean final ratings of Target Cues W, X, and Y in the present experiment for all participants in the uninformed and informed groups. Error bars indicate the standard errors of the means.

magnitude of retrospective revaluation effects, we calculated difference scores among all three target cues: a backward blocking score, by subtracting the mean rating for Cue W from the mean rating for Cue X (1.82 and 0.93, for Groups Informed and Uninformed, respectively), and a recovery from overshadowing score, by subtracting the mean rating for Cue Y from the mean rating for Cue X (2.05 and 0.73, for Groups Informed and Uninformed, respectively). A 2 (group: informed vs. uninformed) × 2 (difference score: backward blocking vs. recovery from overshadowing) ANOVA yielded a main effect of group [$F(1,81) = 9.00$, $MS_e = 5.54$], showing that the difference scores were significantly higher in Group Informed than in Group Uninformed. No interaction was found. Planned comparisons revealed that the difference between groups in the backward blocking score and in the recovery from overshadowing score, although numerically larger in Group Informed than in Group Uninformed, fell short of statistical significance [$F(1,81) = 2.43$, $MS_e = 6.59$, $p = .12$, and $F(1,81) = 3.95$, $MS_e = 9.07$, $p = .05$, for the backward blocking score and for the recovery from overshadowing score, respectively]. Therefore, explicit information about absent cues effectively increased the magnitude of the overall retrospective revaluation effect, but not to the degree necessary to allow us to observe a reliable difference in each of its two components: backward blocking and recovery from overshadowing.

In addition, we considered the results of the postexperimental questionnaire, which allowed us to determine which participants remembered the four compound cues that had been presented during the experiment. Eight participants in the informed group and 15 participants in the uninformed group did not correctly identify the four compounds. Group Uninformed performed less accurately (67% correct) than Group Informed (80% correct), which is reasonable given that participants in Group Informed had been explicitly informed about the two cues of the compound stimuli during Phase 2 (those cues present, indicated with *yes*, or those cues absent, indicated with *no*), whereas participants in Group Uninformed had only been explicitly informed about the presented cues. Because the enhanced retrospective revaluation effect in the informed group might be due not only to enhanced salience of the absent cues but also to enhanced memory of the within-compound associations, we further analyzed the final ratings of the target cues, segregating the data of participants who had correctly identified the four compounds from the data of participants who had not. These reformatted data are displayed in Figure 3.

A 2 (condition: informed vs. uninformed) × 2 (compound memory: incorrect vs. correct) × 3 (cue: W vs. X vs. Y) ANOVA confirmed reliable differences among the ratings of Cues W, X, and Y by participants who correctly identified the compounds in the informed and uninformed conditions, but yielded no differences for participants who failed to correctly identify the compounds. There was a significant main effect of condition [$F(1,79) = 16.25$, $MS_e = 7.07$], a significant main effect of compound memory [$F(1,79) = 6.95$, $MS_e = 7.07$], and a significant main effect of cue [$F(2,158) = 17.65$, $MS_e = 4.30$]. Notably, there was also a significant compound memory × cue interaction [$F(2,158) = 4.64$, $MS_e = 4.30$], showing that cue ratings were different in participants who correctly identified the compounds and in participants who did not correctly identify the compounds. Planned comparisons confirmed that participants who remembered the compounds in the informed condition showed both backward blocking [$F(1,79) = 23.30$, $MS_e = 3.20$] and recovery from overshadowing [$F(1,79) = 16.51$, $MS_e =$

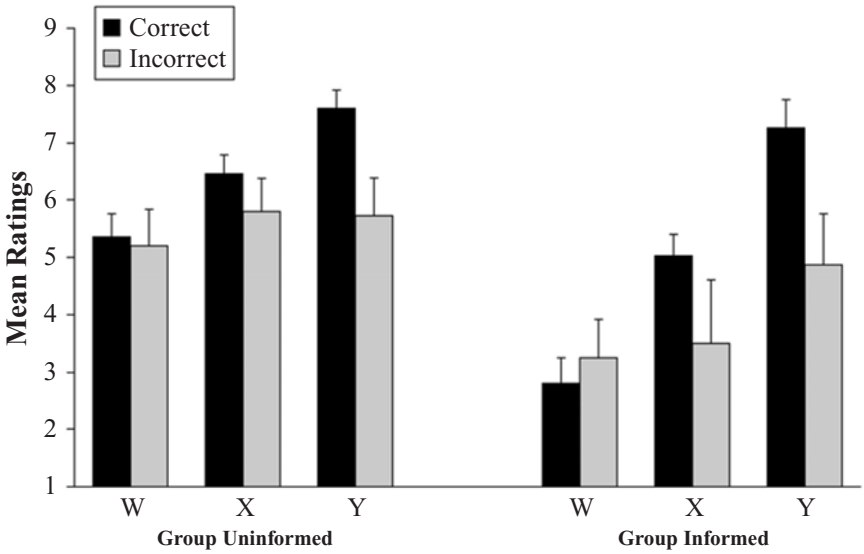

Figure 3. Mean final ratings of Target Cues W, X, and Y in the present experiment for participants in the uninformed and informed conditions separated according to whether or not they correctly identified all four compounds presented during training. Error bars indicate the standard errors of the means.

4.52]. Participants who remembered the compounds in the uninformed condition again showed backward blocking [$F(1,79) = 5.65$, $MS_e = 3.20$]; recovery from overshadowing, which had not reached significance in the previous analysis, was now significant [$F(1,79) = 4.25$, $MS_e = 4.52$].

Especially interesting was the fact that participants who did not correctly identify the compounds did not show any of the retrospective revaluation effects. There was no significant difference between ratings of Cue W and Cue X, in either the informed condition [$F(1,79) = 0.07$, $MS_e = 3.20$] or in the uninformed condition [$F(1,79) = 0.84$, $MS_e = 3.20$]. Nor was there a significant difference between ratings of Cue Y and Cue X in the informed condition [$F(1,79) = 1.66$, $MS_e = 4.52$] or in the uninformed condition [$F(1,79) = 0.01$, $MS_e = 4.52$]. There was not even a significant difference between ratings of Cue W and Cue Y [$F(1,79) = 2.03$, $MS_e = 5.18$, in the informed condition, and $F(1,79) = 0.41$, $MS_e = 5.18$, in the uninformed condition]. These results are consistent with the notion that the formation and retention of within-compound associations in this task is necessary to produce retrospective revaluation effects (see also Aitken et al., 2001; Dickinson & Burke, 1996; Melchers et al., 2004; Wasserman & Berglan, 1998; but see Matute & Pineño, 1998a, 1998b, and Escobar, Pineño, & Matute, 2002).

We also analyzed the difference in the magnitude of the retrospective revaluation effects, now without the participants who did not remember the compounds. We again calculated the difference scores among all three of the target cues to yield a backward blocking score (2.23 and 1.10, for the informed and uninformed conditions, respectively) and a recovery from overshadowing score (2.23 and 1.13, for the informed and uninformed conditions, respectively). A 2 (condition: informed vs. uninformed) $\times$ 2 (difference score: backward blocking vs. recovery from overshadowing) ANOVA yielded a main effect of condition [$F(1,58) = 6.44$, $MS_e = 5.80$], showing that the difference scores were significantly higher in the informed condition than in the uninformed condition. No interaction was found. Planned comparisons revealed that the difference for the backward blocking score and the recovery from overshadowing score, as before, although numerically larger in the informed than in the uninformed condition, fell short of statistical significance [$F(1,58) = 3.13$, $MS_e = 19.26$, $p = .08$, and $F(1,58) = 2.61$, $MS_e = 18.15$, $p = .11$, for the backward blocking score and for the recovery from overshadowing score, respectively]. Most important was the fact that the enhanced overall retrospective revaluation effect in the informed condition compared with the uninformed condition persisted, despite the removal from both conditions of participants who had not correctly identified the compounds (Figure 3). Thus, differential memory alone cannot account for the difference in the magnitude of the overall retrospective revaluation effect in the informed and uninformed groups.

Finally, we should comment on the time course of participants' causal ratings. If we compare ratings after Phase 1 training to final ratings (Table 4), we see that in both groups, causal ratings of Cue W decreased from the end of Phase 1 to the end of Phase 2, consistent with backward blocking.

Ratings of Control Cue X—which should maintain a similar value from the second to the final rating because neither it nor its associated Cue B were presented in

Phase 2—did so in Group Uninformed, but they decreased in Group Informed. It might be that the 90 trials of Phase 2, without Cue X (or Cue B), is responsible for this decrement. Although this decrement did not occur in the present uninformed group, it did occur in Wasserman and Berglan (1998), which also involved an uninformed group. Another explanation might be that training of other cues (Cue A in our case) with the same outcome as previously trained cues (Cue X in our case) can produce an interference effect that results in a decrement in the value of a cue even when neither that cue nor any associated cue is further trained (see Matute & Pineño, 1998a, 1998b, for an extended explanation of this effect).

Causal ratings of Cue Y did not seem to increase from the end of Phase 1 to the end of Phase 2 in any of the groups; such a rise would be expected from recovery from overshadowing. Nonetheless, it should be noted that when the data from participants who did not remember the compounds were excluded, ratings of Cue Y rose from 6.76 in the second rating period to 7.26 in the final rating period in Group Informed, and from 7.10 in the second rating period to 7.60 in the final rating period in Group Uninformed; changes in Cue W and Cue X maintained the same pattern with or without the participants who did not remember the compounds. It could also be that, because Cue Y is already rated quite highly after Phase 1 training, a ceiling effect occurred that allowed only a weak increment in causal ratings for this cue.

DISCUSSION

Our present success in obtaining backward blocking and recovery from overshadowing, both with and without explicit information about absent cues, joins other experimental research showing increments and decrements in the strength of absent cues when other cues are presented. Most of the earlier retrospective revaluation studies lacked a control condition that would allow experimenters to determine whether the observed behavioral differences were due to backward blocking, to recovery from overshadowing, or to both effects (Chapman, 1991; Chapman & Robbins, 1990; Dickinson & Burke, 1996; Williams et al., 1994). Wasserman and Berglan (1998) showed, using a control condition without training in Phase 2, that both effects were involved in retrospective revaluation. Our uninformed group followed Wasserman and Bergman's design and it yielded the same results. Furthermore, both backward blocking and recovery from overshadowing occurred, with even greater magnitude, in the informed group. Processing an event as "absent" can thus be enhanced by explicitly noting its absence; evidently explicit information about the absence of an event is even greater than is information arising from the mere nonoccurrence of an event.

Still, some recent studies have found that recovery from overshadowing can be a more robust retrospective revaluation effect than is backward blocking. Larkin et al. (1998), using the same control condition without training in Phase 2, obtained recovery from overshadowing but not backward blocking (see also Lovibond, Been, Mitchell, Bouton, & Frohardt, 2003). Although the revised RW model predicts both backward blocking and recovery from overshadowing, other theoretical accounts predict only recovery from overshadowing; they predict that backward blocking should be weak, nonexistent, or intertwined with other effects.

According to Wagner's (1981) SOP model, stimuli are represented as nodes consisting of several elements. The elements in a node can be in an inactive state (I) or in one of two possible activation states: a primary activation state (A1) or a secondary activation state (A2). The elements of the representation of a particular Cue X change from being inactive to the maximal activation level, the A1 state, only when the cue is presented; thereafter, they decay into the secondary activation state, A2; and, finally, they again become inactive. The elements of the representation of a cue can also change from the inactive state to the A2 state when another cue is presented that had previously been paired with Cue X. When the elements of a representation of Cue X are in A1, it can enter into associations with other active representations. An excitatory association between representations will be formed when the elements of both representations are in the A1 state; on the contrary, an inhibitory association will be formed when the elements of a representation are in A1 and the elements of another representation are in A2.

Dickinson and Burke (1996) extended Wagner's SOP model to allow the development of an excitatory association from a representation in A1 or A2 to another concurrent representation in the same activation state; on the other hand, an inhibitory association should be formed when two concurrent representations are in different states (see also Aitken & Dickinson, 2005). Thus, after training of Compound CY^+ in Phase 1, presentation in Phase 2 of Cue C^- should activate the representations of both Cue Y and reinforcement in A2; since both representations are in A2, an excitatory association between them should be strengthened, leading to an increment in responding to Cue Y in testing. By contrast, after AW^+ training, when Cue A^+ is presented alone in Phase 2, it should activate both Cue W and reinforcement in A2. Because reinforcement is also present, some elements of its representation will also be in A1. Cue W will be involved in two learning processes: one excitatory, a consequence of Cue W's activation in A2 and the activation in A2 of representative elements of the reinforcer, and another inhibitory, a consequence of Cue W's activation in A2 and the activation in A1 of representative elements of the reinforcer. So, backward blocking should be a weak effect that may or may not occur, depending on which of these two processes is stronger.

Other researchers, who propose that retrospective revaluation is the result of statistical and inferential processes, also doubt that backward blocking is a genuine effect. For example, according to Cheng (1997), in order to evaluate the causal efficiency of a cue, participants have to deter-

mine to what extent the cue influences the probability of the occurrence of the outcome. In the case of backward blocking, Cue W is presented along with Cue A predicting the outcome, and then Cue A alone is established as a cause of the outcome. Because during this training the outcome is always present, it is not possible to know whether Cue W alone could have the capacity of increasing the probability of occurrence of the outcome; this "ceiling effect" would preclude any conclusion about the causal status of Cue W. In other words, in the case of backward blocking, if Cue A is a cause, then Cue W could still be a cause; people simply cannot rationally make this decision. The low ratings that participants may give to Cue W may reflect their uncertainty about its causal status instead of their certainty about Cue W being noncausal.

The case of recovery from overshadowing is different. When Cue Y is presented along with Cue C predicting the outcome and then participants are informed that Cue C alone does not cause the outcome, participants can logically infer that Cue Y has to be the cause of the outcome. In this case, there is no uncertainty and participants should show a genuine effect. However, Cheng (1997) would not be able to explain why we found an even stronger backward blocking effect when participants were explicitly told that Cue W is absent than when Cue A alone was presented. Uncertainty about what would happen if Cue W is presented alone should be the same in both uninformed and informed groups. Cheng would have a similar problem explaining why informed absence also enhanced the strength of recovery from overshadowing.

In addition to Larkin et al. (1998), De Houwer et al. (2002) and Lovibond et al. (2003) did not find backward blocking under standard training procedures, although these researchers were able to get backward blocking when they minimized possible ceiling effects on reinforcer magnitude. However, not only the present study and that of Wasserman and Berglan (1998), but also the studies of Shanks (1985) and Le Pelley and McLaren (2001), have obtained backward blocking with the control condition in which none of the cues is trained in Phase 2. In fact, in the present experiment, it was backward blocking that tended to be stronger than recovery from overshadowing. In Group Uninformed, although the trend existed for the whole set of participants, recovery from overshadowing was statistically significant only when the data from participants who did not correctly identify all of the compounds were eliminated, whereas backward blocking was significant in both informed and uninformed groups, both with and without the data from the participants who did not correctly identify the compounds.

One possible reason for earlier unsuccessful efforts to obtain backward blocking might be the necessarily low salience of absent cues. Because of this low salience, a considerable amount of training might be needed to form or to modulate associations involving absent cues. As Larkin et al. (1998) noted, in the case of backward blocking, an absent cue must be involved in an inhibitory process, which can take longer than the excitatory process involved in recovery from overshadowing. Studies that have yielded backward blocking have generally involved a large number of trials in Phase 2 (around 30 trials of each type, although Le Pelley & McLaren, 2001, used only 8 trials), whereas studies that have not yielded backward blocking have generally involved a small number of trials (10 or fewer trials of each type). Determining whether it is the number of trials or some other variable that is responsible for some researchers' difficulty in obtaining backward blocking awaits further research.

Conclusions

The nonoccurrence of an event, like its occurrence, can be informative, indeed. Learning theories should be concerned not only with events that are actually *present*, but also with those that are *absent*. We have found that a cue's strength can be changed even when this cue is absent. Moreover, the change in an absent cue's strength seems to be in the opposite direction of the change in a present cue's strength. When two cues are first presented together in compound and one of the cues is later presented alone, the strength of the absent cue will decrease if the present cue alone is paired with the outcome, whereas the strength of the absent cue will increase if the present cue alone is not paired with the outcome. These changes can be observed even when judgments are required after every trial, thereby allowing us to conclude that modifications in the strength of absent cues are updated in a trial-by-trial fashion. Finally, we have here reported that the salience of a nonpresented cue can be enhanced through explicit information about its absence. It seems to us that an associative learning theory, like the revised RW model, is well suited to explain all of these facts of human causal judgment. This success is notable because the revised RW model provides a mechanistic process account of behavioral effects that some theorists believe can only be explained by complex statistical formulations.

Statistical models of human causal learning explain the different phenomena observed in causality judgments as the result of an inductive reasoning process that takes place only after people have been provided with all of the relevant contingency information. People are said to extract information about event covariation and the frequencies or probabilities of the events, and later to apply a rule to integrate this information (e.g., Cheng, 1997). Although we can compare people's final performance to evaluate associative and statistical models, as many studies have done (see, e.g., Baker, Murphy, Vallée-Tourangeau, & Mehta, 2001; Lober & Shanks, 2000; Perales & Shanks, 2003; Wasserman et al., 1996), it should be noted that these two kinds of theories actually focus on different stages of the learning process. In Marr's (1982) terminology, statistical accounts are normative models that specify the objectives that people have to satisfy when solving a particular task; in other words, these accounts are concerned with *what* is computed. On the other hand,

associative accounts are normally algorithmic models that describe *how* the computations are carried out (see López, Cobos, Caño, & Shanks, 1998; Shanks, 1995).

Statistical models are not concerned with the nature of the processes underlying the acquisition of contingency information. Acquisition processes and other phenomena that may take place during this stage of training cannot therefore be explained (e.g., the unfolding of learning curves; the nature of trial order effects; or the fact that retrospective revaluation depends on the cues having been presented together). Statistical models are concerned with a particular input and what the corresponding output must be, not with the processes that the input undergoes or the factors that influence how that input produces that output. But, if learning theory cannot explain those processes, then what can? Any final account of contingency learning and judgment need not be an associative model; but, up to now, associative models have most successfully explained the phenomena of learning and contingency judgment.

Learning is an adaptive process, one that is driven by surprise. We infer its existence from changes in behavior; but, learning is more than this behavior change. As the revised RW model indicates, we learn only when there is a discrepancy between what happens and what we predict ought to have happened; we learn when something is surprising, be it the occurrence or the nonoccurrence of an outcome, be it the occurrence or the nonoccurrence of a cue. Thus, we reaffirm that an associative account like the revised RW model can be a valuable theoretical tool if we are to explain causal learning.

We have argued that human research participants evaluate both present and absent cues in light of evidence they receive about outcome occurrence and nonoccurrence. This claim is not difficult to accept because we can verbally instruct humans to consider and to rate a set of cues, even cues that are not given. But, is it at all reasonable to believe that something akin to this process goes on with nonhuman animals, who cannot be so instructed? We believe so, but only if there is some other cue that is presented on a trial that might trigger a memory or representation of the absent cue. Such a trigger or retrieval stimulus might be either a contextual cue or a discrete cue with which the target cue has previously been associated.

In animal conditioning research, changes in responding to absent cues due to additional training with other associated cues have been reported; the first such effect was recovery from overshadowing (Kaufman & Bolles, 1981; Matzel, Schachtman, & Miller, 1985). Backward blocking has proven to be more difficult to obtain, unless the associated cues are of low biological significance (Miller & Matute, 1996b). Cues of high biological significance appear to be resistant to retroactive interference, which may be reasonable from an evolutionary point of view (for further discussion about possible species and task differences in backward blocking, see Miller & Matute, 1996a, 1996b).

Backward blocking and recovery from overshadowing are interpretively embraced by the revised RW model. But, other recent data on retrospective revaluation cannot be explained by our model. Denniston, Savastano, Blaisdell, and Miller (2003; see also Denniston, Savastano, & Miller, 2001) found that rats change their responding to absent cues even when they do not have a direct within-compound association with an additionally trained cue. In their study, the Compound AB was followed by the outcome in Phase 1, and then, in Phase 2, the Compound BX was also followed by the outcome. Then, in Phase 3, Cue A was either presented on its own without the outcome or was not presented at all. When tested with Cue X, rats' responding was weaker when Cue A had been followed by no outcome than when Cue A had not been presented. Although Cue X had never been presented with Cue A—and therefore an A–X within-compound association could not have been formed—further training with Cue A influenced responding to Cue X, presumably because Cues A and X had each been paired with Cue B.

The revised RW cannot explain this effect because there is no a within-compound association between Cue A and Cue X. And, even if one were to propose that presentation of Cue A activates the representation of Cue B, which in turn activates the representation of Cue X, allowing an associative change in this absent cue, the obtained effect is still unexplainable by the revised RW model. According to the revised RW model, when Cue A is not followed by the outcome, if absent Cue X were somehow activated, then it should increase its associative strength; however, the opposite result is actually observed. This so-called "second-order" retrospective revaluation effect has also been found with human participants (De Houwer & Beckers, 2002a; Melchers et al., 2004); De Houwer and Beckers (2002b) have further documented "second-order" backward blocking and "second-order" recovery from overshadowing.

A final point deserves discussion. At the beginning of our paper, we mentioned sensory preconditioning (e.g., Brogden, 1939) and mediated conditioning (e.g., Holland, 1981) as established effects in which learning about absent cues is observed. But these effects are not retrospective revaluation effects. In sensory preconditioning and mediated conditioning, the change in strength of the absent cue is in the *same* direction as the change in strength of the present cue; but, in the case of retrospective revaluation effects, the change in strength of the absent cue is in the *opposite* direction from the change in strength of the present cue. These opposite changes in the associative values of absent cues pose a critical challenge to all learning models because models that can explain one directional effect cannot explain the opposite one.[1]

Some authors have tried to resolve this disparity. In his flavor conditioning studies, Dwyer (1999, 2001) suggested that the timing of the cue–outcome relationship may produce opposite effects. If the representation of the absent cue is retrieved while the outcome is present, then

the associative strength of the absent cue might decrease; conversely, if the outcome is presented after the retrieval of the absent cue has occurred, then the associative strength of the absent cue may increase. Other researchers have proposed that the number of training trials may play a critical role in producing opposite associative changes. After training Cue A followed by the outcome and the Compound AB followed by the nonoccurrence of the outcome, Yin, Barnet, and Miller (1994) found second-order conditioning when few AB trials were given, either after or interspersed with A^+ trials; but, they found conditioned inhibition when the number of AB trials was large and interspersed with the A^+ trials. In other words, strong responding to Cue B was obtained with a small number of trials, but Cue B became an inhibitory stimulus after a large number of trials (see also Rashotte, Marshall, & O'Connell, 1981; Rescorla, 1973; Santoveña, Álvarez, Fernández, Pérez, & Loy, 2002; Stout, Escobar, & Miller, 2004). We earlier conjectured that a small number of trials might account for past failures to obtain backward blocking (Larkin et al., 1998). It would be worthwhile to see whether a large number of trials facilitates backward blocking or whether the opposite effect would be obtained with a small number of trials.

Clearly, there are still many questions awaiting an answer. If animal conditioning research takes into account the challenges and advances that are coming from human learning in the same way as human learning has benefited from animal learning discoveries and models, then the fruitful interchange of ideas and theories might continue and enhance our knowledge about how the cognitive system works.

REFERENCES

Aitken, M. R. F., Dickinson, A. (2005). *Learning & Behavior*, **33**, 147-159.

Aitken, M. R. F., Larkin, M. J. W., & Dickinson, A. (2001). Re-examination of the role of within-compound associations in the retrospective revaluation of causal judgements. *Quarterly Journal of Experimental Psychology*, **54B**, 27-51.

Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin*, **114**, 435-448.

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.

Amsel, A. (1967). Partial reinforcement effects on vigor and persistence: Advances in frustration theory derived from a variety of within-subjects experiments. In K. W. Kenneth & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 1, pp. 1-65). New York: Academic Press.

Anderson, J. R., & Sheu, C.-F. (1995). Causal inferences as perceptual judgments. *Memory & Cognition*, **23**, 510-524.

Arcediano, F., Matute, H., & Miller, R. R. (1997). Blocking of Pavlovian conditioning in humans. *Learning & Motivation*, **28**, 188-199.

Baker, A. G., Murphy, R. A., Vallè-Tourangeau, F., & Mehta, R. (2001). Contingency learning and causal reasoning. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 255-306). Montreal: McGill.

Brogden, W. J. (1939). Sensory pre-conditioning. *Journal of Experimental Psychology*, **25**, 323-332.

Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, & 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. J., & Wasserman, E. A. (1985). Response–outcome contingency: Behavioral and judgmental effects of appetitive and aversive outcomes with college students. *Learning & Motivation*, **16**, 1-34.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, **104**, 367-405.

Cohen, J., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behavior Research Methods, Instruments, & Computers*, **25**, 257-271.

De Houwer, J., & Beckers, T. (2002a). A review of recent developments in research and theories on human contingency learning. *Quarterly Journal of Experimental Psychology*, **55B**, 289-310.

De Houwer, J., & Beckers, T. (2002b). Second-order backward blocking and unovershadowing in human causal learning. *Experimental Psychology*, **49**, 27-33.

De Houwer, J., Beckers, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, **55A**, 965-985.

Denniston, J. C., Savastano, H. I., Blaisdell, A. P., & Miller, R. R. (2003). Cue competition as a retrieval deficit. *Learning & Motivation*, **34**, 1-31.

Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. In R. R. Mowrer & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 65-117). Hillsdale, NJ: Erlbaum.

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

Dickinson, A., & Burke, J. (1996). Within-compound associations mediate the retrospective revaluation of causality judgements. *Quarterly Journal of Experimental Psychology*, **49B**, 60-80.

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

Dwyer, D. M. (1999). Retrospective revaluation or mediated conditioning? The effect of different reinforcers. *Quarterly Journal of Experimental Psychology*, **52B**, 289-306.

Dwyer, D. M. (2001). Mediated conditioning and retrospective revaluation with LiCl then flavour pairings. *Quarterly Journal of Experimental Psychology*, **54B**, 145-165.

Escobar, M., Pineño, O., & Matute, H. (2002). A comparison between elemental and compound training of cues in retrospective revaluation. *Animal Learning & Behavior*, **30**, 228-238.

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

Hearst, E. (1991). Psychology and nothing. *American Scientist*, **79**, 432-443.

Holland, P. C. (1981). Acquisition of representation-mediated conditioned food aversion. *Learning & Motivation*, **12**, 1-12.

Kamin, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami symposium on the prediction of behavior: Aversive stimulation* (pp. 9-31). Miami: University of Miami Press.

Kao, S. F. (1993). *Information integration and associative learning as accounts of human contingency judgment*. Unpublished doctoral dissertation, University of Iowa.

Kao, S. F., & Wasserman, E. A. (1993). Assessement of an information integration account of contingency judgment with examination of participative cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1363-1386.

Kaufman, M. A., & Bolles, R. C. (1981). A nonassociative aspect of overshadowing. *Bulletin of the Psychonomic Society*, **18**, 318-320.

Killeen, R. R. (1981). Learning as causal inference. In M. L. Commons & J. A. Nevins (Eds.), *Quantitative analyses of behavior: Vol. 1. Discriminative properties of reinforcement schedules* (pp. 89-112). Cambridge, MA: Ballinger.

Larkin, M. J. W., Aitken, M. R. F., & Dickinson, A. (1998). Retrospective revaluation under positive and negative contingencies. *Jour-*

nal of Experimental Psychology: Learning, Memory, & Cognition, **24**, 1331-1352.

Le Pelley, M. E., & McLaren, I. P. M. (2001). Retrospective revaluation in humans: Learning or memory? *Quarterly Journal of Experimental Psychology*, **54B**, 311-352.

Levin, I. P., Wasserman, E. A., & Kao, S. F. (1993). Multiple methods for examining biased information use in contingency judgments. *Organizational Behavior & Human Decision Processes*, **55**, 228-250.

Lober, K., & Shanks, D. R. (2000). Is causal induction based on causal power? Critique of Cheng (1997). *Psychological Review*, **107**, 195-212.

López, F. J., Cobos, P. L., Caño, A., & Shanks, D. R. (1998). The rational analysis of human causal and probability judgment. In M. Oaksford & N. Chater (Eds.), *Rational models of cognition* (pp. 314-352). Oxford: Oxford University Press.

Lovibond, P. F., Been, S.-L., Mitchell, C. J., Bouton, M. E., & Frohardt, R. (2003). Forward and backward blocking of causal judgment is enhanced by additivity of effect magnitude. *Memory & Cognition*, **31**, 133-142.

Mackintosh, N. J. (1977). Conditioning as the perception of causal relations. In R. E. Buttes & J. Hintikka (Eds.), *Foundational problems in the special sciences* (pp. 214-250). Dordrecht, The Netherlands: Reidel.

Mandel, D. R., & Lehman, D. R. (1998). Integration of contingency information in judgments of cause, covariation, and probability. *Journal of Experimental Psychology: General*, **127**, 269-285.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: Freeman.

Matute, H., & Pineño, O. (1998a). Cue Competition in the absence of compound training: Its relation to paradigms of interference between outcomes. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 38, pp. 45-81). San Diego: Academic Press.

Matute, H., & Pineño, O. (1998b). Stimulus competition in the absence of compound conditioning. *Animal Learning & Behavior*, **26**, 3-14.

Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1985). Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learning & Motivation*, **16**, 398-412.

Melchers, K. G., Lachnit, H., & Shanks, D. R. (2004). Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory. *Quarterly Journal of Experimental Psychology*, **56B**, 25-53.

Miller, R. R., & Matute, H. (1996a). Animal analogues of causal judgment. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.): *The psychology of learning and motivation: Vol. 34. Causal learning* (pp. 133-166). San Diego: Academic Press.

Miller, R. R., & Matute, H. (1996b). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General*, **125**, 370-386.

Perales, J. C., & Shanks, D. R. (2003). Normative and descriptive accounts of the influence of power and contingency on causal judgement. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, **56A**, 977-1007.

Rashotte, M. E., Marshall, B. S., & O'Connell, J. M. (1981). Signaling functions of the second-order CS: Partial reinforcement during second-order conditioning of the pigeon's keypeck. *Animal Learning & Behavior*, **9**, 253-260.

Rescorla, R. A. (1973). Second-order conditioning: Implications for theories of learning. In F. J. McGuigan & D. B. Lumsden (Eds.), *Contemporary approaches to conditioning and learning* (pp. 127-150). New York: Wiley.

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

Santoveña, A., Álvarez, I., Fernández, A., Pérez, J. C., & Loy, I. (2002). *Contrary associative effects: Second order conditioning and conditioned inhibition.* Paper presented at the Fourteenth Meeting of the Spanish Society of Comparative Psychology, Universidad de Sevilla, Spain.

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

Shanks, D. R. (1995). *The psychology of associative learning*. Cambridge: Cambridge University 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: 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., López, F. J., Darby, R. J., & Dickinson, A. (1996). Distinguishing associative and probabilistic contrast theories of human contingency judgment. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation: Vol. 34. Causal learning* (pp. 265-311). San Diego: Academic Press.

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

Stout, S., Escobar, M., & Miller, R. R. (2004). Trial number and compound stimuli temporal relationship as joint determinants of second-order conditioning and conditioned inhibition. *Learning & Behavior*, **32**, 230-239.

Tassoni, C. J. (1995). The least mean squares network with information coding: A model of cue learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 193-204.

Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning & Motivation*, **25**, 127-151.

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., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, **76**, 171-180.

Ward-Robinson, J., & Hall, G. (1996). Backward sensory preconditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, **22**, 395-404.

Wasserman, E. A. (1974). Stimulus-reinforcer predictiveness and selective discrimination learning in pigeons. *Journal of Experimental Psychology*, **103**, 284-297.

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: Academic Press.

Wasserman, E. A., & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgment: The role of within-compound associations. *Quarterly Journal of Experimental Psychology*, **51B**, 121-138.

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, & Cognition*, **16**, 509-521.

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

Wasserman, E. A., Kao, S. F., Van Hamme, L. J., Katagiri, M., & Young, M. E. (1996). Causation and association. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of learning and motivation: Vol. 34. Causal learning* (pp. 207-264). San Diego: Academic Press.

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.

Williams, D. A., Sagness, K. E., & McPhee, J. E. (1994). Configural

and elemental strategies in predictive learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 694-709.

Yin, H., Barnet, R. C., & Miller, R. R. (1994). Second-order conditioning and Pavlovian conditioned inhibition: Operational similarities and differences. *Journal of Experimental Psychology: Animal Behavior Processes*, **20**, 419-428.

Young, M. E. (1995). On the origin of personal causal theories. *Psychonomic Bulletin & Review*, **2**, 83-104.

NOTE

1. Because the output of the SOP model reflects the net effect of a set of opponent learning processes, the possibility exists that this model might predict mediated conditioning or sensory preconditioning under some circumstances and retrospective revaluation under other circumstances. However, which specific factors can lead to these opposite results needs first to be found and detailed before any model can explain them.

APPENDIX
Experimental Instructions

Initial instructions, presented before Rating Period 1

This is an exercise to gauge how people form opinions. Before we begin the experiment, we would like to get your initial opinions of some foods. This is to check whether you happen to hold any opinions about the foods that we will be using in these tests. We would like for you to imagine that an *ordinary individual* who is *completely unknown* to you has eaten certain foods. In this questionnaire, you will see a list of foods. We will ask you to rate the likelihood that *this individual* would have a nasty allergic reaction after eating each of those foods. This person may not have an allergic reaction to any foods. Imagine that you had to bet on whether or not *this* individual would have an allergic reaction after eating each of the foods. We would like for you to remember the following two facts before continuing:

1) You should base your opinions on the fact that *most* people do *not* suffer from allergic reactions to any foods.

2) You have *no* reason to believe that *this* individual is any different from *most* people.

Instructions presented after Rating Period 1

Now we would like you to imagine that you are an allergist, that is, you are someone who tries to discover the cause of allergic reactions in people. You have just been presented with a new patient, "Mr. X," who suffers from allergic reactions. In an attempt to discover which foods cause him to have allergic reactions, you arrange for him to eat various foods for a meal on each day and you observe if he has an allergic reaction or not. The results of the daily allergy tests will be shown to you on a series of screens. You will see a separate screen for each day of the allergy test. On each screen you will be told what the patient ate for dinner that day and if there was an allergic reaction. Sometimes you will be shown the name of one food and sometimes two. Please read the food names carefully and remember that your task as an allergist is to determine which food or foods are causing an allergic reaction.

After seeing each day's foods, you will be asked to predict whether or not each meal caused an allergic reaction in your patient. You simply press the "1" key on the keyboard if you believe that your patient will suffer an allergic reaction and you press the "3" key if you think the patient will not suffer an allergic reaction. After you make your prediction, the computer will inform you whether or not the patient actually suffered from a reaction. Obviously, at first you will have to guess because you will not know anything about your patient, but hopefully you will begin to learn which foods cause him an allergic reaction and which do not. You might view this experiment as a game and try to score as many points (correct predictions) as you can. You will see the number of correct predictions that you have made near the bottom of the screen during the daily allergy tests. Later in the experiment, you will again be asked to rate the foods you previously rated. But, in those future ratings, we would like for you to rate the likelihood that each food would cause "Mr. X," your patient, to have an allergic reaction. You should use all of the knowledge that you have acquired during the daily allergy tests when you make your ratings in the future.

Instructions presented before Rating Periods 2 and 3

Now we would like you for you to rate the foods again, but now for "Mr. X," your patient. You will see a rating scale identical to the one you used previously. When you make your ratings this time, please consider all of the information that you have received throughout the experiment, not just the information from the last day.