

## Causal order does not affect cue selection in human associative learning

DAVID R. SHANKS

*University College London, London, England*

and

FRANCISCO J. LOPEZ

*University of Malaga, Malaga, Spain*

Waldmann and Holyoak (1992) presented evidence in support of the claim that cue selection does not emerge in "diagnostic" human learning tasks in which the cues are interpretable as effects and the outcomes as the causes of those effects. Waldmann and Holyoak argued that this evidence presents a major difficulty for associationist theories of learning and instead supports a "causal model" theory. We identify a number of flaws in Waldmann and Holyoak's experimental procedures and report three new experiments designed to test their claim. In Experiment 1, cue selection was observed regardless of causal order and regardless of whether the cues were abstractly or concretely specified. In Experiments 2 and 3, cue selection was again observed when subjects predicted causes from effects. We conclude that our results are consistent with simple associationist theories of learning but contradict Waldmann and Holyoak's causal model theory.

In an influential animal conditioning experiment, Saavedra (reported by Wagner, 1969) demonstrated that the extent to which a cue and an outcome come to be associated is a function not simply of their repeated pairing but rather of the degree to which the cue is a reliable predictor of the outcome. Saavedra used a rabbit eyelid conditioning procedure in which a brief neutral stimulus such as a light is paired with a mild shock to the eye, as a result of which it becomes a conditioned stimulus (CS) capable of eliciting by itself the conditioned response of blinking. In one group (the contingent group), the light was paired on some trials with another neutral stimulus, a tone, and this compound was followed by a shock. On other trials, the tone occurred on its own and was unreinforced. If the light was Cue A and the tone Cue B, and shock was the outcome, then the animals in this group received intermixed AB→outcome and B→no-outcome trials, with shock being contingent on Cue A's presence. At the end of the training phase, Cue A was tested on its own, and (unsurprisingly) it elicited a substantial conditioned response, indicating that it had become strongly associated with the shock.

A second group of animals (the noncontingent group) again received AB→outcome trials, but for these animals trials with Cue B by itself were reinforced rather than un-

reinforced; thus these animals received intermixed AB→outcome and B→outcome trials. Since shock could occur on trials with B alone, shock was not contingent on Cue A's presence. When this group was tested with Cue A by itself, little conditioned responding was elicited, indicating that Cue A had become only weakly (if at all) associated with the shock.

The importance of Saavedra's result (and of a number of other, similar, results obtained around the same time; see Mackintosh, 1983) lies in the fact that even though Cue A was paired with the outcome an equal number of times in the two groups, its association with the outcome differed. The reason is that in the contingent group, A was a good predictor of the unconditioned stimulus (US) compared to Cue B, whereas in the noncontingent group, A was an entirely redundant stimulus in that it provided no information that was not already conveyed by B. Thus, a process of cue selection appears to operate in associative learning: A cue will be selected only for association with another event if it is a good relative predictor of that event.

This sort of cue selection effect played a major role in the late 1960s in the development of formal associationist learning theories such as that of Rescorla and Wagner (1972), which is equivalent to the "delta" rule used to update the weights in many current connectionist models of human learning (e.g., Gluck & Bower, 1988; McClelland & Rumelhart, 1985). In brief, Rescorla and Wagner's (1972) theory explains Saavedra's result by assuming that there is a ceiling or limit to the amount of association strength that can be supported by any given outcome event. Early in training, Cues A and B both accrue some associative strength for the outcome on an AB→outcome trial, and on a B→outcome trial in the noncontingent condition Cue B will gain further strength. Before long, the combined strengths

---

The research described here was supported in part by a project grant from the United Kingdom Biotechnology and Biological Sciences Research Council to D.R.S. and a predoctoral fellowship from the Junta de Andalucia to F.J.L. We thank Julian Almaraz, Anthony Dickinson, Pablo Fernandez, Mark Gluck, Doug Hintzman, Keith Holyoak, Michael Waldmann, John Wixted, and an anonymous reviewer for their helpful comments. Correspondence should be addressed to D. R. Shanks, Department of Psychology, University College London, Gower St., London WC1E 6BT, England (e-mail: david.shanks@psychol.ucl.ac.uk).

of A and B on the compound trials will exceed what can be borne by the outcome, and at that point, the cues necessarily begin to lose strength. However, Cue B is able to make up what it loses on compound trials by acquiring strength increments on B-alone trials. Cue A is not able to do this, with the result that it keeps losing strength until it has none at all. Effectively, Cue B "blocks" it from retaining any associative strength at asymptote. In the contingent condition, by contrast, the B→no-outcome trials force Cue B to lose some of the strength it has acquired on the compound trials. With continued exposure to AB trials, Cue A gains more and more strength, but because the total strength that can be attached to the outcome is limited, the strength of Cue B will be driven down, eventually reaching an asymptotic value of zero. Hence the result obtained by Saavedra is explained.

Given the extensive success of connectionist models that use the delta rule (or variants of it) in modeling human learning in a variety of different domains (see, e.g., Cohen, Dunbar, & McClelland, 1990; Gluck & Bower, 1988; McClelland & Rumelhart, 1985; Seidenberg & McClelland, 1989), it is perhaps not surprising that cue selection effects similar to Saavedra's have now been observed in a range of human learning situations, including Pavlovian (e.g., Martin & Levey, 1991), instrumental (e.g., Shanks, 1985; Wasserman, Elek, Chatlosh, & Baker, 1993), and category learning (e.g., Shanks, 1991). For example, Wasserman et al. (1993) asked subjects to judge the relationship between pressing a button (the cue) and a light coming on (the outcome) in a situation in which the context was an alternative potential cause of the outcome (equivalent to Cue B). When a design analogous to Saavedra's was used, with the context by itself either causing the outcome (noncontingent condition) or not causing the outcome (contingent condition), subjects gave significantly lower ratings for noncontingent responses than for contingent ones.

Although there is no consensus concerning the validity of the Rescorla–Wagner model itself (see Mackintosh, 1975; Pearce, 1987, and Pearce & Hall, 1980, for alternative associative models), and despite the fact that cue selection effects exist that have resisted explanation (e.g., Busemeyer, Myung, & McDaniel, 1993a, 1993b; Shanks, 1985), it has been widely argued (e.g., Dickinson, 1980; Hall, 1991; Shanks, 1994) that the appropriate way to understand cue selection phenomena is in terms of the sorts of competition mechanisms embodied in associationist theories. In sum, much of the recent interest in associationist or connectionist theories has been stimulated by their ability to account for a range of cue selection effects.

### Causal Models and Causal Order

In contrast to this associationist approach, Waldmann and Holyoak (1992) have argued that cue selection effects emerge for entirely different reasons. Waldmann and Holyoak's article is part of a large and very interesting research program being pursued by these authors and their colleagues aimed at challenging the enterprise of understanding human and animal learning in terms of associationist principles (see, e.g., Cheng, 1993; Cheng & Holyoak, 1995; Cheng &

Novick, 1992; Holyoak, Koh, & Nisbett, 1989; Melz, Cheng, Holyoak, & Waldmann, 1993; see also Shanks, 1995). Waldmann and Holyoak adopt a more mentalistic "causal model" approach to learning, in which

people use meaningful world knowledge, often of a highly abstract sort, to guide their learning about new domains. One major example of abstract world knowledge is knowledge about the basic characteristics of causal relations, such as the temporal precedence of causes to their effects. (Waldmann & Holyoak, 1992, p. 224)

According to this account, subjects use real-world knowledge to interpret the cues and outcomes in an experiment, and in particular use knowledge about such things as causal precedence: Regardless of the order in which the events are perceived, subjects know about the directionality of causation.

Having brought an interpretative causal mental model to bear on the task, subjects then determine the extent to which the cause and effect are related. In contrast to the notion that associative knowledge is derived by incrementing or decrementing a mental bond, Waldmann and Holyoak proposed instead that subjects compute the degree of statistical contingency,  $\Delta P$ , between the events, where contingency is defined as

$$\Delta P = P(E/C) - P(E/-C) \quad (1)$$

Here,  $P(E/C)$  is the probability of the effect given the cause and  $P(E/-C)$  is the probability of the effect in the absence of the cause. The derived value of  $\Delta P$  then represents the subject's judgment about the relationship between C and E.

Like the Rescorla–Wagner model, this account of learning is able to explain certain cue selection phenomena. With respect to Saavedra's experiment, such an approach readily yields the appropriate result. In both the contingent and noncontingent conditions, the probability of the effect (shock) given the target cause (Cue A) is 1.0. However, in the contingent condition the probability of shock in the absence of Cue A is zero, whereas in the noncontingent condition it is greater than zero; in fact, assuming that the probabilities are computed only across the trials and not across the intertrial intervals,  $P(E/-C)$  is 1.0 and hence  $\Delta P$  is zero. At any rate, the statistic  $\Delta P$  can readily be seen to differ between the two conditions in the required direction (see Melz et al., 1993, for further details).

As an aside, it is worth mentioning that there are certain elaborations to the way in which  $\Delta P$  is calculated (e.g., the notion of conditional independence; see Cheng & Holyoak, 1995) that allow Waldmann and Holyoak (1992) to explain how subjects distinguish between genuine and spurious causes. For instance, in a noncontingent condition it is possible that Cue A is a genuine cause that would, if presented on its own, be followed by the outcome. Alternatively, its pairing with the outcome on AB trials may be spurious. Associative theories such as that of Rescorla and Wagner have little to say about this distinction, whereas Waldmann and Holyoak's theory specifies how subjects distinguish between these possibilities in real life. However,

the elaborations that allow this to be accomplished will not be described here because they are not relevant to the present concerns (see Cheng & Holyoak, 1995, for details).

How may we discriminate between the associationist and causal model explanations of cue selection? Waldmann and Holyoak (1992) noted that different predictions can be derived from the theories in so-called diagnostic learning tasks. Suppose the subject is required to learn some cue–outcome relationship, but the appropriate causal interpretation is that the event labeled as the outcome would (in the real world) have caused the event labeled as the cue rather than vice versa. For instance, in a number of studies (e.g., Shanks, 1991), subjects have learned to predict diseases on the basis of symptoms, but of course in the real world it is diseases that cause symptoms. This sort of “diagnostic” learning can be distinguished from “predictive” learning where the cues are the causes and the outcomes the effects.

Waldmann and Holyoak (1992) predicted that while cue selection may emerge in predictive contexts, it should not emerge in diagnostic conditions. In the noncontingent  $AB \rightarrow \text{outcome}$ ,  $B \rightarrow \text{outcome}$  design, for example, the mental model theory assumes that the subject interprets these trial types as  $E_1, E_2 \leftarrow \text{cause}$  and  $E_2 \leftarrow \text{cause}$ , where  $E_1$  and  $E_2$  are different effects. Despite the fact that the order in which the events are experienced is opposed to their true causal order, on this theory subjects will learn relationships from the cause to the effects, and the presence of  $E_2$  should not alter in any way the judged causal relationship between the cause and  $E_1$ . This is because the presence of multiple effects does not change the computed value of  $\Delta P$  for a given cause–effect relationship. As Waldmann and Holyoak said (1992, p. 224), “People have a strong disposition to learn directed links from causes to their effects, rather than vice versa, even in situations in which they receive effect information prior to cause information,” and, “Different effects, like different dependent measures obtained in an experiment, do not compete with one another; rather, each effect, as well as any interaction among the effects, provides information about the consequences of the cause” (p. 226).

For an associative account, however, the real-world interpretation of the events is immaterial. If the subject is required to predict outcomes from cues, the cues represent the input to the system and the outcome represents the target to be predicted, regardless of causal order. Thus, even though cue selection in predictive situations (where the cues are interpreted as causes and the outcomes as effects) can be explained by either theory, the causal model account predicts no cue selection in diagnostic situations where the opposite interpretation is made, while the associationist theory does. Note that in the remainder of this article we will use the terminology “cause–effect” (CE) and “effect–cause” (EC) rather than that used by Waldmann and Holyoak (1992) (“predictive” and “diagnostic,” respectively) to describe the causal interpretation of the events, because in both types of tasks subjects are asked to make predictions (in the sense of stating ahead of time what they think will happen). In this sense, the terms CE

and EC are less ambiguous as descriptions of the status of the cues and outcomes.

What is the evidence of cue selection in EC tasks? On the face of it, Waldmann and Holyoak’s theory would appear to be incorrect since cue selection has been obtained in a number of EC situations. For instance, in a direct comparison between contingent and noncontingent conditions, Shanks (1991) obtained a cue selection effect in a task in which subjects learned relationships between the symptoms and diseases that a variety of patients had. Similarly, Chapman (1991) observed cue selection with the same medical diagnosis task. Moreover, Waldmann and Holyoak (1992, Experiment 2) themselves obtained an apparent cue selection effect in an EC task, and the magnitude of the effect was comparable to that obtained in a CE task. In their experiment, the actual cues and outcomes in the CE and EC tasks were similar, but the tasks differed in terms of how people interpreted the events. In the CE context, subjects had to learn relationships between certain features of a person’s appearance (perspiration, skin, posture, weight) and the outcome (an emotional response in observers of that person). In the EC context, the same cues were interpreted as symptoms (i.e., effects) of a virus.

Both tasks took place across two phases. In Phase 1, a predictive Cue P was consistently paired with the outcome in that it was present every time the outcome was present and absent whenever the outcome was absent. Another two cues were also presented, Cue C, which occurred on every trial regardless of whether the outcome was present, and Cue U, which was uncorrelated with the outcome. During Phase 2 the same cues were again present but a target, Cue R, was added. This redundant cue was present only on those trials in which Cue P was also present. Waldmann and Holyoak (1992) reasoned that since Cues P and R had each been paired with the outcome, any difference between them in terms of subjects’ ratings at the end of Phase 2 would be evidence of cue selection. The relevant result was that an attenuation of the ratings of Cue R occurred in *both* the CE and the EC contexts so that, apparently, cue competition had arisen in both cases.

Rather than taking this result as evidence favoring the associative theory, Waldmann and Holyoak (1992) made the reasonable point that the attenuation of the ratings for Cue R relative to Cue P in the EC context may not have been due to cue competition per se (competition between effects in this case) but instead to the fact that subjects may, on the basis of real-world knowledge, have attributed the effects to extra-experimental causal factors. For instance, in a medical diagnosis context with multiple symptoms (effects) of a single disease (cause), if one of the symptoms is a runny nose, the subject may assume that another disease (e.g., a cold) must be present in addition to the specifically mentioned cause. The critical point is that if a cue in an EC design has a concrete label that evokes prior knowledge, the target cause may implicitly be competing with other causes.

To elaborate, Waldmann and Holyoak (1992, p. 231) stated that

no attenuation would be expected if there were only one possible theory explaining the evidence; however this is seldom the case in realistic diagnostic tasks. There is potential competition if multiple causal theories might account for a given effect. More specifically, if subjects bring to bear prior knowledge of alternative possible causes of an abnormal body sign, they may lower their rated predictiveness of the sign as an indicant of the fictitious disease being taught in the experiment. Thus, even though subjects learned in Phase 2 that being underweight (Cue R) is an effect of the virus, they presumably also knew that there are many other reasons a person might be underweight. These alternative causes may have competed with the newly acquired cause as a possible explanation of the single cue *underweight*, causing subjects to give the single cue a relatively low predictiveness rating.

To eliminate this alternative explanation of the lower ratings for the redundant cue in the EC context, Waldmann and Holyoak (1992) performed another experiment (Experiment 3) in which any effects attributable to prior knowledge concerning the cues were eliminated. In this experiment they used an unfamiliar context for which subjects were assumed to be unable to introduce prior knowledge of alternative potential causes, and again ratings in CE and EC contexts were compared. Subjects were told to imagine they were in a bank and had to learn relationships between the state of activation of the alarm system and various buttons (CE condition) or indicator lights (EC condition). In this case, Waldmann and Holyoak obtained a significant cue competition effect in the CE condition, but not in the EC condition, in accordance with the causal model theory and contradictory to an associationist analysis. They observed a similar difference between CE and EC tasks in their first experiment, but we delay discussion of that experiment until the General Discussion.

Waldmann and Holyoak's (1992) results are, on the face of it, highly problematic for associationist explanations of cue selection. In associationist accounts it is assumed that when the subject is asked to predict an outcome, the cues represent the input to a network and the outcome represents the output, with prediction responses being determined by unidirectional weights from the cues to the outcomes. Cue selection should emerge whenever multiple cues predict a single outcome, regardless of the possible causal order interpretation that can be put on the events. Waldmann and Holyoak's (1992, Experiment 3) result contradicts this prediction, and the present experiments were undertaken in an attempt to reproduce and explore the generality of their findings. Experiments 1 and 2 were designed to test the claim that cue competition does not occur in EC tasks, and in Experiment 3 we attempted to distinguish between the associationist and causal model theories.

### EXPERIMENT 1

Waldmann and Holyoak (1992, Experiments 2 and 3) concluded from their results that cue competition arises in CE but not EC contexts. There are two slightly different ways of stating the problem that this result presents to associative theories. One way is to concentrate on the inter-

action Waldmann and Holyoak obtained in their Experiment 3 between causal order and cue (Cues P and R received different ratings in the CE condition but not in the EC condition), and the other way is to concentrate solely on the absence of a cue competition effect in the EC condition. While both of these results are problematic for associative theories, we believe that attention should be principally focused on the latter characterization.

First, the interaction Waldmann and Holyoak (1992, Experiment 3) obtained arose in an experiment in which rather different cover stories were given to subjects in the CE and EC tasks. Remember that the key result was that Cue P was rated higher than were Cue R in the CE task but not in the EC task. Thus far, we have not described the exact details of the cover story for that experiment, but those details are worth examining because they suggest a possible reason for the difference in ratings. In the CE task, cues were said to be different buttons in the control rooms of a bank, and subjects had to detect which button or buttons (cues) switched on the alarm (outcome). In Phase 1, subjects were instructed that Peter was working in Room A. In that phase, Button P was a reliable predictor of the outcome. Before the beginning of Phase 2, subjects were told that Mary was in another control room (Room B) and that, by accident, whenever Peter pressed Button P in Room A Mary pressed Button R in Room B. The outcome was that Button R was rated lower than Button P at the end of Phase 2. Note that because the buttons were being pressed by different people in different rooms, there is no reason why subjects should have believed that there was any necessary relationship between pressings of Buttons P and R; in fact, the simultaneous pressings were explicitly described as accidental.

In the EC task, cues were said to be lights in a control room that indicated whether the alarm was on or off. During Phase 1, only Room A was available. Subjects had to detect which light (cues) indicated the status of the alarm (outcome). A particular cue (Cue P) was a reliable predictor of the status of the alarm. Before the beginning of Phase 2, subjects were told that there was another control room (Room B) with different indicator lights. During Phase 2, whenever Light P was on in Room A, Light R was also on in Room B. At the end of Phase 2, ratings for Cues P and R did not differ reliably, implying an absence of cue selection. However, it is not very clear what specific mental representation subjects might have had about this task given the cover story provided. In particular, since subjects were told at the outset of Phase 2 that information would now be provided about a signal light in Room B, they might have reasoned that the light had been present all along and had been lit in Phase 1 whenever Light P was lit, but that they were simply not informed about it. If this was the case, then subjects might have reasoned that Light R was perfectly correlated with Light P via some causal link, in which case equivalent judgments should have been given to them.

In sum, subjects may have inferred that in the EC version of the task, but not in the CE version, Cues P and R were linked to each other. That being the case, subjects

would be justified in giving them equivalent ratings. Thus, in our view the interaction Waldmann and Holyoak (1992, Experiment 3) obtained may have arisen because the cover stories could be read in a way that was inconsistent with the aim of the experiment. In the present experiment we attempted to see if such an interaction occurs in a situation in which ambiguity in the cover stories is eliminated.

A second reason for focusing mainly on the absence of a cue selection effect in the EC condition is that this is—given prior demonstrations of cue competition in human associative learning (see, e.g., Martin & Levey, 1991; Shanks, 1985, 1991; Wasserman et al., 1993)—probably the most surprising finding of Waldmann and Holyoak's (1992) study. Thus in the discussion that follows, we concentrate on the apparent absence of competition in the EC condition rather than on the difference in the magnitude of competition in the EC and CE conditions.

We begin by discussing some significant problems with Waldmann and Holyoak's (1992) experiments that may have contributed to the absence of the key competition effect in Experiment 3. The most significant problem is that because of inappropriate experimental design, their studies do not in fact allow any conclusions to be drawn concerning cue competition effects at all. To illustrate this important point, the following discussion simplifies Waldmann and Holyoak's design by excluding some of the irrelevant cues. The experiment took place in two phases. Ignoring Cues C and U, Phase 1 consisted of P→outcome trials while Phase 2 contained PR→outcome trials. Waldmann and Holyoak compared ratings given to P with those given to R at the end of Phase 2 and concluded that cue competition had taken place if a difference occurred between these ratings. But a moment's reflection reveals that these are not appropriate ratings for comparison to determine whether cue competition has occurred. Apart from anything else, subjects have had unequal exposure to these cues: Unless learning is very fast, even a noncompetitive learning theory is likely to predict higher ratings for P, simply because the subject has witnessed more pairings of P with the outcome than of R.<sup>1</sup>

Demonstrating cue selection (see Kamin, 1968) requires a separate comparison condition, as in Saavedra's (cited in Wagner, 1969) experiment. The simplest of these would involve a between-subjects comparison between ratings for Cue R in the original group versus those in a group exposed only to the Phase 2 trials. Better still is a within-subjects design (see, e.g., Rescorla, 1983) in which subjects see P→no-outcome and P'→outcome<sub>2</sub> trials in Phase 1 followed by PR→outcome<sub>1</sub> and P'R'→outcome<sub>2</sub> trials in Phase 2, the relevant contrast being between Cues R and R'. A difference between these two ratings would represent a genuine cue competition effect uncontaminated by differences in exposure to the cues (R and R') being contrasted. It is known that cue competition in CE tasks does occur with this design (Shanks, 1985).

Three further interpretational problems with Waldmann and Holyoak's (1992) experiments should be mentioned. First, as will be apparent from our description of their experiments, Waldmann and Holyoak used a blocked- rather

than an intermixed-trials design. Subjects saw a block of P→outcome trials followed by a block of PR→outcome trials. In Saavedra's experiment, in contrast, the AB→outcome and B→outcome trials were intermixed. While each of these designs is capable of demonstrating cue selection (indeed, the blocked-trials design was used by Kamin, 1968, in his famous experiments), it is well known that an attenuation of learning to Cue A after separate B→outcome and AB→outcome blocks of trials (relative to learning with just the AB→outcome trials) can be quite hard to obtain with human subjects (see Davey & Singh, 1988; Lovibond, Siddle, & Bond, 1988; Martin & Levey, 1991) for reasons that have little to do with cue selection per se. The evidence suggests that human subjects are particularly prone to treat the B and AB trial types as unrelated and thus to regard the two stages of the blocked-trials design as unconnected, leading them in some circumstances to learn normally about Cue A (see Lovibond et al., 1988). If this is the case, then Waldmann and Holyoak may have been using an experimental design ill-suited to observing cue selection. A mixed-trials design may be more sensitive.

Second, in attributing the presence versus absence of cue competition in EC situations to whether or not prior knowledge was available, Waldmann and Holyoak (1992) were forced to make a cross-experiment comparison between Experiments 2 and 3. Because the tasks used were very different, it is possible that procedural or subject differences between the experiments may have produced the result, rather than the extent of prior knowledge engaged by the task. Third, the only result Waldmann and Holyoak obtained that represents a difficulty for associative theories is a null result, namely the absence of a reliable difference in ratings of Cues P and R in the EC condition of Experiment 3. With only 12 subjects in the EC group, each making one judgment for Cue R and one for Cue P, it is possible that the failure to observe a difference between the ratings for the two cues is attributable to a lack of sufficient statistical power.

In the present study we tried to reproduce Waldmann and Holyoak's (1992) results in a single experiment using basically the same intermixed-trials design that Shanks (1991, Experiment 1) used to study cue competition. In four different groups, we manipulated the two factors that Waldmann and Holyoak found relevant to the asymmetry issue: the context in which the task is presented (CE vs. EC order) and the relevance of prior knowledge (concrete vs. abstract cues). Thus, the two tasks were identical, except that the cues and outcomes could be understood as causes and effects (CE task) or as effects and causes (EC task), respectively. In the CE task, the cues represented different foods and the outcomes represented different allergic reactions. Subjects had to learn the relationships between these different foods and the different types of allergic reaction that they caused. The EC task used a medical diagnosis task. Cues were understood as symptoms and outcomes as diseases from which the different patients suffered. In addition to varying causal order, we also manipulated the extent to which prior knowledge might be relevant. In the concrete versions of the task, the different cues

were given actual names (e.g., avocado, dizziness). In the abstract versions cues were labeled by letters. If symptoms are labeled in this abstract way in the EC version, it is difficult to think how they can evoke alternative causes as possible explanations of their occurrence.

The trial types presented to the groups were identical. In the contingent conditions, the compound Cues CD and GH were paired with Outcomes 2 and 4, respectively, whereas D and H were unaccompanied by these outcomes when they occurred on their own. Thus, the relative predictiveness of the target Cues C and G remained high in these conditions. In the noncontingent conditions, the compound Cues AB and EF were paired with Outcomes 1 and 3, respectively, whereas cues B and F were also paired with Outcomes 1 and 3, respectively, when they occurred by themselves. Thus, the relative predictiveness of the target Cues A and E was attenuated.

## Method

**Subjects and Apparatus.** A total of 64 University College and University of Malaga undergraduate students of both sexes were paid for their participation in the present experiment. They were tested individually in a sound-proofed testing cubicle. The experiment was presented on a personal computer with the subject responding via the keyboard. Note that for the University of Malaga subjects the instructions and cue/outcome names were translated into Spanish.

**Procedure.** At the start of the experiment, subjects were given a brief verbal description of the experiment telling them that they would have to learn to predict various outcomes. They then read the experimental instructions on the computer screen (see Appendix).

There was a total of 120 learning trials. On each trial, the program selected at random 1 of the 10 trial types shown in Table 1, with the constraint that there should be 12 trials of each type. The cues appeared on the screen in a list, and for trial types containing more than one cue, the order of cues in the list was chosen at random. The subject selected a category response and typed the appropriate key, and then corrective feedback was given.

The filler trial types IJK and LMN were included to allow an additional test of Waldmann and Holyoak's (1992) causal model theory. In the concrete/EC version of the task, the rate of learning should be slower than in the three other groups. According to the theory, an alternative cause (or causes) can be evoked in this condition due to subjects' prior knowledge, and this implicit cause (or causes) will compete with the actual cause with which those cues have been explicitly paired. This, in turn, will make the relationship between the filler cues and their associated outcomes harder to detect, and hence should retard learning.

The subjects were allocated at random to one of four groups ( $n = 16$  per group, 8 from University College and 8 from the University of Malaga). For half the subjects (concrete/CE and abstract/CE groups), the cues corresponded to causes and the outcomes to effects. The cover story said that on each trial a patient was being given different types of foods as cues, and the task was to predict the type of allergic reaction the patient would develop. In the EC groups

(concrete/EC and abstract/EC), the cover story stated that on each trial a patient's symptoms would be described and the task was to predict the accompanying disease. Orthogonally, for half the subjects (abstract/CE and abstract/EC groups) the task was abstract while for the other half it was concrete (concrete/CE and concrete/EC). In the abstract information groups the foods or symptoms were labeled from A to N. In the concrete information groups, the foods (from A to N: cheese, chocolate, milk, eggs, fish, banana, olive oil, avocado, orange squash, butter, onion, coffee, meat, garlic) and symptoms (from A to N: discolored gums, skin rash, puffy eyes, slurred speech, swollen glands, stomach cramps, dizziness, bloody nose, fever, sore arms, dulled hearing, coughing, perspiration, stiff joints) were given specific names.

In the abstract versions, the allergies were labeled Type 1 to Type 6 and the diseases were labeled Disease 1 to Disease 6. In the concrete versions, the allergies were again named from Type 1–6, but the diseases (1–6) were: Marshall-Isaacs disease, Coralgia, Phipp's syndrome, Burlosis, Terrigitis, and Dempes disorder.

The allocation of cue and outcome names was counterbalanced in each group. In the concrete/CE group, for instance, Cue A was cheese and Cue C was milk for half the subjects, while for the remaining subjects these assignments were switched. Orthogonally, for half the subjects Cue E was fish and Cue G olive oil, while for the remaining subjects these were reversed. Similar counterbalancing took place with respect to Outcomes 1 and 2 and Outcomes 3 and 4.

After the training phase, subjects were given a questionnaire in which they had to rate on a scale from 0 to 100 how strongly each of the Cues A, C, E, G, I, and L was associated with each of the six different outcomes: "On a scale from 0 to 100, how strongly is [cue] associated with [outcome]? Zero indicates no relationship, 100 indicates a very strong relationship."

## Results and Discussion

The subjects were able to make correct predictions by the end of the training phase of the experiment. Table 2 shows the percentage of correct responses across the different experimental groups on the last two trials of each trial type. Because there were six categories, plus the no-outcome category, chance performance is 1/7 or 14% correct. The table shows that the subjects had learned correctly to categorize each of the different cue types, with the percentage of correct responses varying from 79.7% to 96.9%.

The data of interest are the ratings subjects gave for the predictive relationship between a cue and the correct outcome: Thus for Cue A the relevant rating is for Outcome 1, for Cue C it is Outcome 2, and so on. We collapsed the A and E ratings into a noncontingent rating and the C and G ratings into a contingent rating and carried out a 2 (contingency: contingent vs. noncontingent)  $\times$  2 (form: concrete vs. abstract)  $\times$  2 (causal order: CE vs. EC) analysis of variance (ANOVA) on these ratings. The means are shown in Figure 1. The only significant result from the analysis was the within-subjects contingency effect [ $F(1,60) = 34.23$ ,  $MS_e = 608.3$ ,  $p < .001$ ], reflecting the fact that the ratings for the contingent cues ( $M = 75.2$ ,  $SEM = 3.84$ ) were higher than those for the noncontingent ones ( $M = 49.7$ ,  $SEM = 3.68$ ). The between-subjects factor, form, was not significant [ $F(1,60) = 1.04$ ,  $MS_e = 1,153.6$ ], nor was the main effect of order [ $F(1,60) = 2.52$ ,  $MS_e = 1,153.6$ ].

Importantly, we were unable to replicate Waldmann and Holyoak's (1992) finding of a difference in the magnitude of the cue competition effect in the abstract/EC condition relative to the other conditions: The contingency  $\times$  form

Table 1  
Trial Types in Experiment 1

Condition	Cues→Outcomes
Noncontingent	AB → 1 / EF → 3 B → 1 / F → 3
Contingent	CD → 2 / GH → 4 D → 0 / H → 0
Fillers	IJK → 5 / LMN → 6

Note—A–N are the cues and 1–6 are the outcomes. 0, no outcome.

**Table 2**  
**Mean Percentages of Correct Responses Across the Last Two Trials**  
**for Each Trial Type and Each Group in Experiment 1,**  
**and for Subjects in Experiment 2**

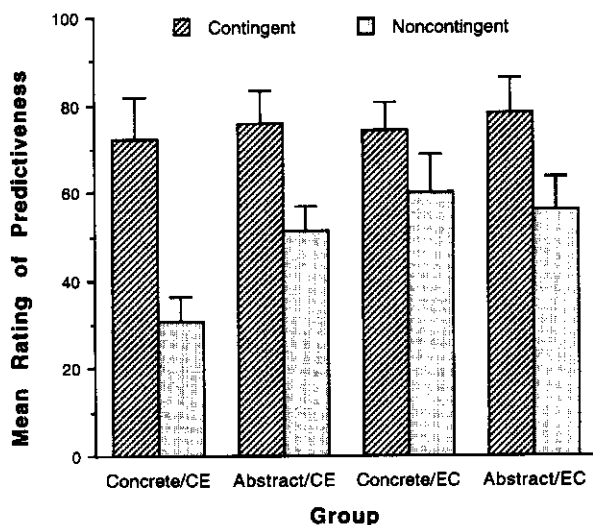
Trial Type	Experiment 1				Experiment 2
	Conc/CE	Conc/EC	Abst/CE	Abst/EC	Abst/EC
AB→1/EF→3	85.9	87.5	79.7	84.4	87.5
B→1/F→3	90.6	89.1	85.9	93.8	90.6
CD→2/GH→4	89.1	79.7	79.7	81.3	84.4
D→0/H→0	90.1	96.9	93.8	90.6	94.5
IJK→5/LMN→6	85.9	92.2	81.3	84.4	86.7

Note—A–N are the cues and 1–6 are the outcomes (see Table 1). 0, no outcome. Conc, concrete; Abst, abstract; CE, cause–effect; EC, effect–cause.

× order interaction was not reliable [ $F(1,60) = 2.00$ ,  $MS_e = 608.3$ ]. No other interactions were significant [all  $F_s(1,60) < 2.88$ ].

The contingent ( $M = 78.3$ ,  $SEM = 7.85$ ) and noncontingent ( $M = 56.3$ ,  $SEM = 7.29$ ) ratings in the crucial abstract/EC group were also compared. Contrary to Waldmann and Holyoak's (1992) findings, the difference between these ratings was significant [ $t(15) = 2.52$ ,  $p < .05$ ]. As Figure 1 shows, the competition effect was if anything slightly larger in the abstract/EC group than in the concrete/EC group, exactly opposite to Waldmann and Holyoak's prediction. It does not seem that ratings were affected to any great extent by real-world knowledge that the concrete cues may have brought to mind.

We interviewed the subjects at the end of the experiment to see whether they had correctly perceived the causal order of the events that they were presented. All the subjects in the CE groups stated that in the real world the cues would be causes and the outcomes effects, whereas all those in the EC groups stated that the cues would be effects and the outcomes the causes.



**Figure 1.** Mean associative rating ( $\pm SEM$ ) for the contingent and noncontingent cues in each group for Experiment 1. CE, cause–effect interpretation of the cues and outcomes; EC, effect–cause interpretation.

One final analysis was carried out. As noted previously, Waldmann and Holyoak's (1992) theory predicts that subjects should find it harder to learn the classification in the concrete/EC group than in the other groups, because in this group extra-experimental knowledge will allow other causes to be brought to mind that will implicitly compete with the actual cause. To test this, we divided the 24 filler trials into six blocks of trials (each block consisting of two presentations per filler) to derive a learning curve for each of the four groups. The percentages of correct responses per block are shown in Figure 2 and indicate no differences between the groups. We performed a trials × form × order ANOVA on these filler results. The within-subjects factor of trial showed the only significant effect [ $F(5,300) = 34.44$ ,  $MS_e = 0.05$ ,  $p < .001$ ]. The trial × form × order interaction was not significant ( $F < 1$ ), disconfirming the prediction of a slower rate of learning in the concrete/EC group relative to that in the others. No other main effects or interactions were significant (all  $F_s < 1.34$ ).

These results are exactly what would be predicted by an associative account of learning such as the Rescorla–Wagner theory. On such a theory, the cues—regardless of their real-world interpretation—represent the inputs to an associative network and the outcomes represent the target outputs of the network. Cue competition is predicted in each of the four groups since in each case the relative predictiveness of the target cue is greater in the contingent than in the noncontingent conditions. The results replicate prior instances of cue selection but also show that selection occurs even when the appropriate interpretation of the cues and outcomes is as effects and causes, respectively.

On the other hand, our results fail to support Waldmann and Holyoak's (1992) claim that cue selection does not occur in abstract/EC tasks, and we have also failed to obtain the difference in the magnitude of cue competition between abstract/EC and abstract/CE tasks that they obtained in their Experiment 3. Although our design and procedure differ in many ways from those used by Waldmann and Holyoak, we believe that our result is the more convincing one, because, as argued above, Waldmann and Holyoak used an inappropriate design that in fact does not speak to the issue of cue selection. Moreover, the results from the abstract version of the EC task allowed us to eliminate the hypothesis that subjects' prior beliefs explain why cue selection appears in EC tasks. In the abstract version,

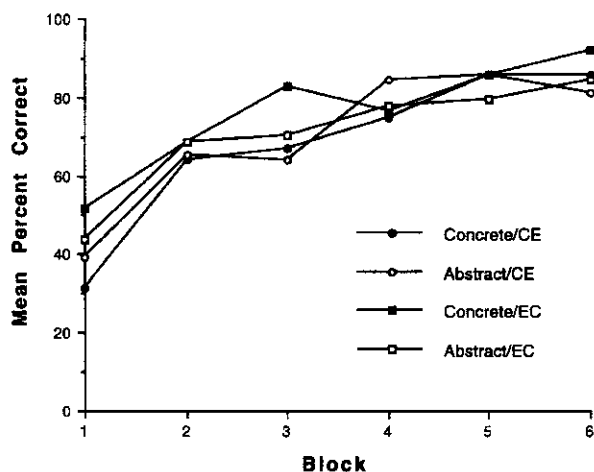


Figure 2. Mean percentage of correct predictions across blocks of four trials (two per filler) in each group for Experiment 1. CE, cause-effect interpretation of the cues and outcomes; EC, effect-cause interpretation.

in which the cues were designated by letters, subjects' prior beliefs about what might cause those cues cannot have played a role.

### EXPERIMENT 2

In Experiment 2 we attempted to replicate the cue selection effect observed in the crucial abstract/EC condition of Experiment 1. To assess the generality of that effect, we used a task with an entirely new cover story.

#### Method

**Subjects.** Thirty-two undergraduate students of both sexes from the University of Malaga took part in this experiment, which was conducted in Spanish.

**Procedure.** The procedure was identical to that used for the abstract/EC group of Experiment 1 (Table 1), except that the cover story for the task was changed. In this experiment, subjects had to detect relationships between different indicator lights in the control room of a chemical plant and different problems that could occur in the plant, with the activation of the indicator lights (cues) being caused by the problems in the plant (outcomes). Indicator lights were labeled with letters (A to N) and problems in the plant were labeled with numbers (1 to 6). After the training phase, subjects were given a questionnaire on which they rated on a scale from 0 to 100 how strongly each of cues A, C, E, G, I, and L was associated with the six different outcomes.

#### Results and Discussion

The subjects were able to detect the relationships between the cues and the different outcomes by the end of the training phase. Table 2 shows that the percentages of correct responses on the last two trials of each trial type ranged from 84.4% to 94.5%.

As in Experiment 1, A and E ratings (for the correct outcome) were collapsed into a noncontingent rating and C and G ratings into a contingent rating. In accordance with the results of Experiment 1, contingent ratings ( $M = 74.8$ ,  $SEM = 5.31$ ) were higher than noncontingent ones

( $M = 56.6$ ,  $SEM = 6.25$ ) [ $t(31) = 3.04$ ,  $p < .01$ ]. Hence the results show a cue competition effect in this new abstract/EC task. In a postexperimental debriefing, all the subjects indicated that they construed the causal relationship in the appropriate way, namely that problems in the plant caused the warning lights to be illuminated.

Although the task used in this experiment is novel, we have replicated the key result obtained in our Experiment 1. Cue competition in EC tasks seems to be a quite general phenomenon.

### EXPERIMENT 3

Our first two experiments suggest that Waldmann and Holyoak (1992) were mistaken when they claimed that results from so-called diagnostic tasks are problematic for associationist theories. In fact, cue selection does occur in such tasks and therefore associationist theories are untroubled. However, the results we have obtained so far do not discriminate between associationist and causal model theories, since both are able to account for the results. Although cue selection in the EC groups in Waldmann and Holyoak's (1992) experiments would have been at variance with their theory, the changes in design we have had to adopt in Experiments 1 and 2 mean that the causal model theory does predict an apparent "selection" effect.

To see how causal model theory can explain the findings, consider the design of Experiment 1, shown in Table 1. For the abstract/EC condition of this experiment (and of Experiment 2), the cues need to be interpreted as effects and the outcomes as causes. This means that the  $AB \rightarrow 1$  and  $B \rightarrow 1$  trials are interpreted by subjects as, respectively,  $E_1 E_2 \leftarrow \text{cause}_1$  and  $E_2 \leftarrow \text{cause}_1$  trials. Recall that the causal model theory assumes that subjects compute the degree of contingency ( $\Delta P$ ) between the cause and effect via Equation 1. Applying the  $\Delta P$  equation to the trial types yields a  $\text{cause}_1 \rightarrow E_1$  contingency of 0.5 for the noncontingent  $AB \rightarrow 1$  and  $B \rightarrow 1$  trials:

$$\Delta P = P(E/C) - P(E/-C) = 0.5 - 0.0 = 0.5.$$

Now consider the contingent  $CD \rightarrow 2$  and  $D \rightarrow 0$  trials, and let us denote the causal model interpretation of these events as  $E_3 E_4 \leftarrow \text{cause}_2$  and  $E_4 \leftarrow \text{no-cause}$ . Computation of the  $\text{cause}_2 \rightarrow E_3$  contingency yields a value of 1.0:

$$\Delta P = P(E/C) - P(E/-C) = 1.0 - 0.0 = 1.0,$$

and hence greater judgments for the contingent than for the noncontingent cue are predicted. Thus like the associative theory, causal model theory is perfectly capable of explaining the contingency effects obtained in the abstract/EC conditions of Experiments 1 and 2.<sup>2</sup>

How can we distinguish between these accounts? A simple modification to our experimental design allows this to be achieved (Table 3). In this new design, the basic arrangement for the target contingent and noncontingent cues is maintained but the frequencies of occurrence of the outcomes associated with these cues are equated by introducing additional trial types. Specifically, the trial types for the noncontingent Cue A become  $AB \rightarrow 1$ ,  $B \rightarrow 1$ ,



**Table 3**  
**Trial Types and Mean Percentages of Correct Responses**  
**Across the Last Two Trials for Each Trial Type**  
**for Subjects in Experiment 3**

Condition	Cues→Outcomes	% Correct
Noncontingent	AB→1/GH→3	79.3
	B→1/H→3	74.0
	C→0/I→0	89.4
Contingent	DE→2/JK→4	65.4
	E→0/K→0	74.0
	F→2/L→4	67.8

Note—A–L are the cues and 1–4 are the outcomes. 0, no outcome.

C→0, and the trial types for the contingent Cue D become DE→2, E→0, F→2. Except for the inclusion of the C and F trials, the arrangement is as before; however, these additional trials alter the obtained values of  $\Delta P$ . Applying the causal model contingency formula to the target Cues A and D yields a value of 0.5 in each case:

$$\Delta P = P(E/C) - P(E/-C) = 0.5 - 0.0 = 0.5.$$

For instance, the AB→1, B→1, and C→0 trials are now interpreted as  $E_1E_2 \leftarrow \text{cause}_1$ ,  $E_2 \leftarrow \text{cause}_1$ , and  $E_3 \leftarrow \text{no-cause}_1$  trials, from which the crucial  $\text{cause}_1 \rightarrow E_1$  contingency is readily seen to be  $\Delta P = 0.5$ . The calculation is the same for the contingent Cue D.

Thus Waldmann and Holyoak's (1992) causal model theory predicts the absence of any difference between ratings for the contingent and noncontingent cues in this design. What about associationist theories? The addition of the extra trial types (C→0 and F→2) to equate the outcome frequencies does not affect the predictions of the Rescorla–Wagner (1972) theory, since the contingent cue remains a better predictor of the outcome than the noncontingent cue, and a selection effect is again predicted (see Shanks, 1991, p. 438, for a simulation of this trial design). The inclusion of the C→0 trials does not alter the fact that Cue B is able to “block” Cue A and prevent it (at asymptote) from acquiring any associative strength for Outcome 1. Similarly, the inclusion of the F→2 trials does not impair Cue D's ability to obtain an asymptotic associative strength for Outcome 2. Experiment 3 therefore provides a crucial test between the two theories. Note that Shanks (1991, Experiment 2) conducted an experiment with this design that yielded a significant difference in judgments for the target contingent and noncontingent cues. But although that study used an EC medical diagnosis task, the cues (symptoms) were concretely specified, and hence (although we have disputed the viability of this process) the result is open to the objection that subjects may have recruited real-world knowledge of extra-experimental causal factors. The present experiment therefore uses abstract cues.

### Method

**Subjects.** Fifty-two undergraduate students of both sexes from the University of Malaga took part in this experiment.

**Procedure.** The procedure was identical to that used for the abstract/EC group of Experiment 1 except that the trial types were modified as in Table 3. Also, there were 120 trials, with 10 of each trial type rather than the 12 of the previous experiments. The symp-

toms were labeled with letters (A to L) and diseases with numbers (1 to 4), and the specific letters were initially assigned at random to the trial types given in Table 3. Across subjects, symptoms and diseases were counterbalanced in the same way as in Experiment 1.

### Results and Discussion

Subjects were able to detect the relationships between the cues and the different outcomes by the end of the training phase. Table 3 shows that the percentages of correct responses on the last two trials of each trial type ranged from 65.4% to 89.4%. Chance performance is 20% correct.

Ratings for Cues A and G (for the correct outcome) were collapsed into a noncontingent rating and ratings for D and J into a contingent rating. In accordance with the results of Experiments 1 and 2, contingent ratings ( $M = 66.6$ ,  $SEM = 4.03$ ) were reliably higher than were noncontingent ones ( $M = 58.2$ ,  $SEM = 4.66$ ) [ $t(51) = 3.40$ ,  $p < .01$ ]. Although the difference between these means is somewhat smaller than that in the previous experiments, the results plainly show that a cue selection effect occurs with this modified design in which the outcome frequencies are equated. The reduced size of the effect is probably attributable to the reduced numbers of trials of each type and the increased overall difficulty of the task brought about by having more trial types.

In this experiment we have again replicated the key result of Experiment 1, namely cue competition in an abstract/EC task. The present results, though, are at variance with the idea that subjects compute contingency in the manner prescribed by Waldmann and Holyoak's (1992) causal model theory.

### GENERAL DISCUSSION

The studies reported here confirm previous results demonstrating cue selection in animal (e.g., Kamin, 1968; Wagner, 1969) and human (e.g., Chapman, 1991; Chapman & Robbins, 1990; Shanks, 1991) learning. The extent to which a cue comes to be associated with an outcome is a function not simply of repeated pairing, but rather of the extent to which the cue conveys nonredundant information about the outcome. The differences we have obtained between contingent and noncontingent cues illustrate that if a target cue is accompanied by another cue that, on its own, is a reliable predictor of the outcome, the target cue will tend to be discounted. As explained earlier, this sort of cue selection effect is straightforwardly accounted for by associative models that use competitive learning algorithms such as the delta rule.

As Melz, Cheng, Holyoak, and Waldmann (1993) have pointed out, the Rescorla–Wagner model in fact predicts rather more extreme ratings than those observed in the present experiments. At asymptote, the associative strength of a noncontingent cue should be zero and that of a contingent cue should be at ceiling, yet judgments for these cues were much less extreme than zero and 100. Two obvious reasons for this discrepancy (see Shanks, 1993) are that (1) learning may not have reached asymptote by the end of the experiment, in which case intermediate judgments

would be expected, and (2) the relationship between internal association strengths and judgments may not be an identity mapping: Because of the rating scale used, associative strengths of zero, for instance, may map onto significantly positive judgments. The present experiments do not allow these possibilities to be tested (and in any case Shanks, 1993, has suggested that they may in general be very difficult to distinguish).

While the present results are perfectly compatible with associative accounts of learning, our findings contradict those reported by Waldmann and Holyoak (1992) and also contradict their causal model theory. They claimed that cue selection does not occur in EC tasks, except when concrete cues are used, and only then because extra-experimental knowledge leads to the introduction of implicit competing causes. If the cues are abstract, the absence of cue selection would be at variance with an associative theory, but would be consistent with a causal model theory in which subjects use prior knowledge about causal order to interpret the events.

We have found, though, that cue selection is just as evident in abstract EC tasks as in CE tasks (Experiment 1), and in Experiment 2 we confirmed that abstract EC tasks can yield a contingent/noncontingent difference. Finally, in Experiment 3 we obtained a difference between ratings of A and D in an  $AB \rightarrow 1, B \rightarrow 1, C \rightarrow 0, DE \rightarrow 2, E \rightarrow 0, F \rightarrow 2$  design. Waldmann and Holyoak's (1992) causal model theory cannot explain this result because calculating the degree of contingency  $\Delta P$  between the relevant cause and effect yields the same result in each case.

There are several reasons to question Waldmann and Holyoak's (1992) apparent failure to obtain cue selection in the abstract EC situation, and instead to give more weight to our contrary findings. First, the main result Waldmann and Holyoak obtained that is problematic for associative theory is the null result of their Experiment 3, but that null result is based on the data from only 12 subjects, each of whom made one rating for Cue P and one for Cue R. In our third experiment, a much larger sample was used (52 subjects), each making two contingent cue ratings and two noncontingent cue ratings. We suggest that lack of power in their experiment may be one reason why Waldmann and Holyoak failed to observe a difference in ratings between the cues.

Second, as noted, Waldmann and Holyoak (1992) adopted a design in which cue selection was confounded with the amount of exposure to each cue, such that none of their experiments can be said with confidence to speak to the issue of cue selection at all. The appropriate comparison to establish cue competition is between two cues each of which has been paired with another cue, where one of those other cues is a good predictor and the other is a poor predictor (see, e.g., Kamin, 1968). Waldmann and Holyoak merely compared Cues P and R after subjects had been presented with  $P \rightarrow$ outcome and  $PR \rightarrow$ outcome training trials, a comparison that fails to meet the standard design requirements. Any difference between ratings of P and R could be attributed to the fact that more trials with Cue P have been observed, rather than to competition between the cues.

Finally, Waldmann and Holyoak (1992) used a blocked-trials design in which subjects saw the  $P \rightarrow$ outcome and  $PR \rightarrow$ outcome trials separately. There is evidence to suggest that such a design is less than optimal for obtaining cue selection effects (see Lovibond et al., 1988) for reasons having little to do with selection per se. We used an intermixed-trials design to avoid this problem.

### Waldmann and Holyoak's Experiment 1

In the discussion to this point, we have said nothing about Waldmann and Holyoak's (1992) first experiment. The design and procedure were similar to those of their Experiment 2. In the CE context, subjects had to learn relationships between certain features of a person's appearance (perspiration, skin, posture, weight) and the outcome (an emotional response in observers of that person), while in the EC context, the same cues were interpreted as symptoms (i.e., effects) of a virus. In the CE version, subjects rated the causal predictiveness of the cues as in Experiment 2, and Waldmann and Holyoak observed significant cue competition with Cue R being rated lower than Cue P.

In the EC task, by contrast, subjects made *effect* rather than causal predictiveness ratings. Specifically, they were told to rate each cue individually as to whether it was an effect of the disease. For these ratings, Waldmann and Holyoak (1992) obtained no significant difference between Cues P and R (although R was rated about 30% lower than was P), and concluded—in accordance with their causal model theory but contradicting the associative account—that cue competition had not occurred in this EC task.

Quite apart from design problems in this experiment similar to those discussed above for the other experiments in Waldmann and Holyoak (1992), and quite apart from the fact that (to our knowledge) no one has ever proposed associationist theories to account for how subjects make "effect" ratings, we argue that the result from the EC group is not as problematic for associationist learning models as Waldmann and Holyoak suggested. Remember that in this condition the symptoms appeared first and were followed by immediate feedback. Subjects must have reasoned from symptoms to disease in order to make their prediction responses on each trial, but the crucial point is that the reverse form of reasoning may also have occurred. Waldmann and Holyoak (1992, p. 229) reported that subjects were told at the outset that they would have to make "effect" ratings: Specifically, they would have to rate each symptom periodically as to whether it was affected by the disease. Given these instructions, and given a lifetime's history of reasoning from causes to effects, it is not unlikely that subjects were motivated to reverse in working memory the effective ordering of the events at the end of each trial so that the disease was literally processed before the symptoms. This may have been the only course of action open to the subjects if they were to be successful at making later "effect" ratings.

But if this happened, then the EC condition would have been subjectively transformed into a condition involving a single cue (the disease) and multiple outcomes (the symptoms), a case for which associative theories such as

the Rescorla–Wagner model predict no competition. Since it is the outcomes and not the cues that restrict the total amount of associative strength that can be shared, multiple outcomes connected to a single cue can acquire independent strengths without competing with one another (and indeed, there is excellent evidence from animal instrumental learning studies that outcomes do not compete; see Rescorla, 1991). Although speculative, the upshot of this analysis of Waldmann and Holyoak's (1992) first experiment is that the results may not after all be inconsistent with associative accounts (see Matute, Arcediano, & Miller, 1996, and Van Hamme, Kao, & Wasserman, 1993, for similar observations).

### Role of Prior Knowledge

The results reported here show no effect of subjects' extra-experimental knowledge concerning causal order on cue competition. We have obtained similar degrees of selection in CE and EC versions of our tasks. However, we do not wish to imply that subjects' prior beliefs never exert any influence on judgments. Indeed, there is good evidence that prior knowledge and expectancies can affect the course of learning, as the extensive review by Alloy and Tabachnik (1984) illustrates. In fact, the observation of cue selection (see, e.g., Shanks, 1985) when trials are presented in blocks (B→outcome, AB→outcome) could be interpreted as evidence that prior knowledge (from the first block) affects learning in the second block. But as Alloy and Tabachnik (pp. 138-139) explained, there is no reason why effects of prior knowledge should necessarily be incompatible with associationist theories. All that is required is to say that prior knowledge is represented in terms of preexisting weights on certain cue–outcome connections, with these weights influencing the later course of learning.

Rather than challenging the idea that prior knowledge may affect associative learning, our results, in contrast to those of Waldmann and Holyoak (1992), support the much more limited conclusion that interpreting cues and outcomes in terms of their real-world causal order does not affect the course of learning.

### REFERENCES

- 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.
- BUSEMEYER, J. R., MYUNG, I. J., & MCDANIEL, M. A. (1993a). Cue competition effects: Empirical tests of adaptive network learning models. *Psychological Science*, *4*, 190-195.
- BUSEMEYER, J. R., MYUNG, I. J., & MCDANIEL, M. A. (1993b). Cue competition effects: Theoretical implications for adaptive network learning models. *Psychological Science*, *4*, 196-202.
- 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.
- CHENG, P. W. (1993). Separating causal laws from causal facts: Pressing the limits of statistical relevance. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 215-264). San Diego: Academic Press.
- CHENG, P. W., & HOLYOAK, K. J. (1995). Complex adaptive systems as intuitive statisticians: Causality, contingency, and prediction. In H. L. Roitblat & J.-A. Meyer (Eds.), *Comparative approaches to cognitive science* (pp. 271-302). Cambridge, MA: MIT Press.
- CHENG, P. W., & NOVICK, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, *99*, 365-382.
- COHEN, J. D., DUNBAR, K., & MCCLELLAND, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*, 332-361.
- DAVEY, G. C. L., & SINGH, J. (1988). The Kamin "blocking" effect and electrodermal conditioning in humans. *Journal of Psychophysiology*, *2*, 17-25.
- DICKINSON, A. (1980). *Contemporary animal learning theory*. Cambridge: Cambridge University Press.
- GLUCK, M. A., & BOWER, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, *117*, 225-244.
- HALL, G. (1991). *Perceptual and associative learning*. Oxford: Oxford University Press, Clarendon Press.
- HOLYOAK, K. J., KOH, K., & NISBETT, R. E. (1989). A theory of conditioning: Inductive learning within rule-based default hierarchies. *Psychological Review*, *96*, 315-340.
- KAMIN, L. J. (1968). "Attention-like" processes in classical conditioning. In M. R. Jones (Ed.), *Miami Symposium on the Prediction of Behavior, 1967: Aversive stimulation* (pp. 9-31). Coral Gables, FL: University of Miami Press.
- LOVIBOND, P. F., SIDDLER, D. A. T., & BOND, N. (1988). Insensitivity to stimulus validity in human Pavlovian conditioning. *Quarterly Journal of Experimental Psychology*, *40B*, 377-410.
- MACKINTOSH, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276-298.
- MACKINTOSH, N. J. (1983). *Conditioning and associative learning*. Oxford: Oxford University Press, Clarendon Press.
- MARTIN, I., & LEVEY, A. B. (1991). Blocking observed in human eyelid conditioning. *Quarterly Journal of Experimental Psychology*, *43B*, 233-256.
- MATUTE, H., ARCEDIANO, F., & MILLER, R. R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *22*, 182-196.
- MCCLELLAND, J. L., & RUMELHART, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General*, *114*, 159-188.
- MELZ, E. R., CHENG, P. W., HOLYOAK, K. J., & WALDMANN, M. R. (1993). Cue competition in human categorization: Contingency or the Rescorla–Wagner learning rule? Comments on Shanks (1991). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *19*, 1398-1410.
- PEARCE, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, *94*, 61-73.
- PEARCE, J. M., & HALL, G. (1980). A model for Pavlovian conditioning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, *87*, 532-552.
- RESCORLA, R. A. (1983). Effect of separate presentation of the elements on within-compound learning in autoshaping. *Animal Learning & Behavior*, *11*, 439-446.
- RESCORLA, R. A. (1991). Associations of multiple outcomes with an instrumental response. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*, 465-474.
- RESCORLA, R. A., & WAGNER, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64-99). New York: Appleton-Century-Crofts.
- SEIDENBERG, M. S., & MCCLELLAND, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523-568.
- SHANKS, D. R. (1985). Forward and backward blocking in human contingency judgement. *Quarterly Journal of Experimental Psychology*, *37B*, 1-21.
- SHANKS, D. R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *17*, 433-443.

- SHANKS, D. R. (1993). Associative versus contingency accounts of category learning: Reply to Melz, Cheng, Holyoak, and Waldmann (1993). *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 1411-1423.
- SHANKS, D. R. (1994). Human associative learning. In N. J. Mackintosh (Ed.), *Animal learning and cognition* (pp. 335-374). San Diego: Academic Press.
- SHANKS, D. R. (1995). Is human learning rational? *Quarterly Journal of Experimental Psychology*, **48A**, 257-279.
- VAN HAMME, L. J., KAO, S.-F., & WASSERMAN, E. A. (1993). Judging inter-event relations: From cause to effect and from effect to cause. *Memory & Cognition*, **21**, 802-808.
- WAGNER, A. R. (1969). Stimulus selection and a "modified continuity theory." In G. H. Bower & J. T. Spence (Eds.), *The psychology of learning and motivation* (Vol. 3, pp. 1-41). New York: Academic Press.
- WALDMANN, M. R., & HOLYOAK, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General*, **121**, 222-236.
- WASSERMAN, E. A., ELEK, S. M., CHATLOSH, D. L., & BAKER, A. G. (1993). Rating causal relations: The role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **19**, 174-188.

#### NOTES

1. Moreover, Waldmann and Holyoak (1992) did not counterbalance Cues P and R. Thus in their Experiment 2, for example, Cue P was skin condition and Cue R was weight for all subjects. Any difference in ratings is therefore confounded with possible preexisting differences or biases concerning these cues, such as the extent to which subjects will learn new associations involving them.

2. We thank Michael Waldmann and Keith Holyoak for suggesting this contingency analysis.

#### APPENDIX

##### Instructions Used in Experiment 1

###### CE Groups:

This experiment examines how difficult it is for people to learn relationships between foods and different types of allergic reactions. You will be presented with the case histories of 120 people. You will be shown the foods [abstract group only: which

are labeled by the letters A to N] that each person ate, and then asked to say which allergy you think the person develops. Some of these people develop Type 1 allergy, some Type 2, some Type 3, some Type 4, some Type 5, some Type 6, and some develop no allergy.

When you have made your prediction, you will be told what the correct allergy is for that person. All you have to do is to try to make as many correct predictions as possible. You will have as much time as you like to make your prediction.

After you have seen all the people, you will be asked to rate how strongly you think each of the foods is related to each of the allergies.

###### EC Groups:

This experiment examines how difficult it is for people to learn to make medical diagnoses. You will be presented with the case histories of 120 people [concrete group only: Some of these patients have Marshall-Isaacs disease, some have Phipp's syndrome, some have Coralgia, some have Dempes disorder, some have Burlosis, some have Terrigitis, and some have no illness].

You will be shown the symptoms [abstract group only: which are labeled by the letters A to N] that each patient has, and then asked to say which illness you think the person is suffering from.

[Abstract group only: Some of these people have Disease 1, some Disease 2, some Disease 3, some Disease 4, some Disease 5, some Disease 6, and some have no illness.] When you have made your diagnosis, you will be told what the correct diagnosis is for that patient. All you have to do is to try to make as many correct diagnoses as possible.

You will have as much time as you like to make your prediction. You do not need to memorize the disease names, since they will be shown on the screen when you make your diagnosis. After you have seen all the patients, you will be asked to rate how strongly you think each of the symptoms is related to each of the illnesses.

(Manuscript received February 14, 1995;  
accepted for publication February 20, 1995.)