

# Asymmetries in the Bidirectional Associative Strengths Between Events in Cue Competition for Causes and Effects

Deanah Kim ([kmkim@usc.edu](mailto:kmkim@usc.edu))

Department of Psychology, University of Southern California  
Los Angeles, CA 90089-1061

Stephen J. Read ([read@usc.edu](mailto:read@usc.edu))

Department of Psychology, University of Southern California  
Los Angeles, CA 90089-1061

## Abstract

Two experiments using social stimuli tested a recurrent neural network model's predictions for cue competition for causes and effects. The delta-rule based model predicts the presence of cue competition for effects as well as for causes as a result of an asymmetry in the bidirectional associative strengths between the relevant cue-outcome pairs. This model can capture cue competition for effects when cues are encountered in the cause-effect direction, unlike associative and feed-forward models. Results support the model's prediction of cue competition for both effects and causes. The implications of these results for causal model theory and for various associative accounts of cue competition are discussed.

## Introduction

Since the advent of research on causal induction many researchers have focused on causal models that explain the competitive nature of learning cues that predict or indicate the occurrence of an event (e.g., Rescorla & Wagner, 1972; Gluck & Bower, 1988; Shanks, 1991; Waldmann & Holyoak, 1992). The first account of competitive learning of cues was the Rescorla-Wagner (1972) model. When multiple cues are present preceding an event, these cues compete with each other for predictive strength, resulting in the competitive learning of cues. A classic example of this is blocking, whereupon learning that stimulus A predicts outcome X during initial training, there is a deficit in learning that B also perfectly predicts X when AB are presented together preceding X in the second training phase.

Although cue competition of causes is well established in both the animal and human causal learning literature, the question of whether competition occurs among multiple effects of a common cause has produced somewhat inconsistent findings and has resulted in a heated debate (e.g., Shanks, 1991; Waldmann & Holyoak, 1992; Van Hamme & Wasserman, 1993; Price & Yates, 1995; Matute, Arcediano & Miller, 1996; Shanks, & Lopez, 1996). Cue competition for effects describes a two-stage conditioning phenomenon whereupon first learning that cause A perfectly predicts the occurrence of effect X during Phase 1 of training, there is a deficit in subsequently learning that cause A also perfectly predicts the occurrence of effect Y when the XY compound is presented together after the presentation of cause A during the Phase 2 training.

A number of researchers have provided evidence for cue competition for effects. Some researchers (e.g., Shanks, 1991; Shanks & Lopez, 1996; Price & Yates, 1995; Cobos,

Lopez, Cano, Almaraz & Shanks, 2002) interpret the findings as consistent with associative learning theories. Others (e.g., Matute et al., 1996; Miller & Matute, 1998) assert that the findings are more consistent with contiguity theory, which assumes that associations are learned noncompetitively and bi-directionally through simple contiguity, and that cue competition takes place at judgment.

In contrast, proponents of causal model theory (e.g., Waldmann & Holyoak 1992; Melz, Cheng, Holyoak & Waldmann, 1993; Waldmann, 2000) deny the evidence for cue competition for effects, suggesting that whereas causes compete, effects do not. They argue that multiple effects should not compete with each other because they provide new information about the effects of a common cause.

The goal of this paper is to contrast the predictions of a recurrent neural network, with associative learning theory, causal model theory and contiguity theory's predictions for cue competition for causes as well as for effects. Two experiments designed to test the different predictions for the occurrence of cue competition between effects will be presented. These experiments use various social behaviors as target stimuli for cues and outcomes in an attempt to extend the research in multiple cue contingency learning beyond the traditional settings of biological, physical or abstract events and their consequences.

In addition, these studies assess both directions of reasoning between causes and effects. Typical research in this domain only investigates one direction of reasoning and thus cannot say anything about the relative strength of the forward link from causes to effects and the backward link from effects to causes.

## Cue Competition in Associative Models

The Rescorla-Wagner (R-W) model (Rescorla & Wagner, 1972) formally describes the change in associative strength during learning by:  $\Delta V_{cs(n)} = \alpha_{cs} \alpha_{us(n)} (\alpha_{us(n)} - V_{total(n)})$ , where  $\Delta V_{cs(n)}$  is the change in the associative strength (V) of CS as a result of a pairing with US on trial n;  $\alpha_{cs}$  is the learning rate parameter of the CS;  $\alpha_{us(n)}$  is the learning rate parameter of the US on trial n;  $\alpha_{us(n)}$  is the asymptote of learning or the maximum associative strength that the US can support on trial n; and  $V_{total(n)}$  is the sum of associative strengths of all CSs present on trial n, or the extent to which the US is predicted on trial n. The basic principle behind the R-W model is that associative learning is determined by the extent to which an US is surprising, represented by the difference

between the US that is actually presented on trial  $n$  and the US that is expected on the basis of the summed predictive value of all the cues that are present on trial  $n$ .

With respect to blocking and cue competition, the R-W model predicts the consequences of presenting multiple causes. Cue competition for causes is observed in blocking experiments because by the end of Phase 1 training, the animal has learned that  $CS_1$  perfectly predicts the US. During Phase 2 training, when  $CS_1$  and  $CS_2$  are presented together with the US, no learning occurs for  $CS_1CS_2$  because changing the associative strength of the  $CS_2$  cannot improve the already perfect predictability of the US.

However, the R-W model is unable to naturally handle cue competition between multiple effects of a common cause, because it is a predictive model that assumes a cause-to-effect directionality in learning associations. In other words, the difference term ( $\Delta V_{total}$ ) only applies to the ability of cue (CS) to predict outcome (US), but not outcome to cue. However, several researchers have obtained cue competition for effects and have provided associative accounts for them (e.g., Shanks, 1991; Shanks & Lopez, 1996; Price & Yates, 1995; Cobos et al, 2002). In order to do so, these researchers have had to resort to the somewhat convoluted procedure of presenting participants with the effects *preceding* the cause (for instance, where symptoms predict the disease that caused them). In other words, the R-W model can accommodate cue competition for effects by using a diagnostic learning procedure, where the multiple effects can be presented as antecedent events, which are understood to occur after their cause, even though the effects are presented prior to the cause. Thus the effects (antecedents) compete with each other in predicting the cause. However, the R-W rule cannot handle the more typical situation in which causes occur before effects.

Gluck and Bower (1988) and Shanks (1991) demonstrated that a simple two layer feedforward network using the delta rule, which is closely related to the R-W rule, correctly predicted competition of cues (symptoms) for a common outcome (a rare disease). The delta rule is an error-correcting learning rule that says that the changes in weights,  $\Delta w_{ij}$ , from input node  $i$  to output node  $j$  is given by the following equation:  $\Delta w_{ij} = \partial (t_j - o_j) a_i$ , where  $a_i$  is the activation on input node  $i$ ;  $t_j$  is the target activation on output node  $j$ ;  $o_j$  is the observed or actual activation on output  $j$ ; and  $\partial$  is the learning rate (constant). As in the R-W rule, the change in weight between the input and output nodes,  $\Delta w$ , depends on the extent to which the target activation of the output differs from the observed activation of the output.

However, this network can only learn forward links from cues to outcomes. Thus, as with the R-W rule they could capture cue competition for effects only by assuming diagnostic learning where the effects precede the cause.

We will show that a recurrent network model with delta-rule learning does not have this limitation, but can handle cue competition for effects when causes precede effects.

### Cue Competition in Causal Model Theory

Causal model theory argues that people use meaningful world knowledge about the basic characteristics of causal relations in conjunction with contingency information to

build causal models of the relations between causes and effect (Waldmann & Holyoak, 1992). The causal model theory uses a contingency rule to deal with a multiple cue situation, where the contingency is the difference between the proportion of cases in which the effect and cause co-occur and the proportion of cases in which the effect occurs in the absence of the cause. When the causal model is predictive, cue competition between causes is expected in the classic blocking paradigm because during Phase 2 training, the new cue,  $Cue_2$ , always co-occurs with first cue,  $Cue_1$ . Waldmann and Holyoak (1992) argue that because "it is impossible to determine whether the observed unconditional contingency between  $Cue_2$  and the effect is genuine or spurious," this should lead to uncertainty, which should further lead to a lowered predictiveness of  $Cue_2$ , or partial blocking (p. 226). On the other hand, they assert that effects do not compete with each other, because each effect provides further information about the cause and there is no uncertainty (Waldmann & Holyoak, 1992; Waldmann, 2000).

Waldmann has done a number of studies that fail to find cue competition for effects (with the exception of Study 2 in Waldmann and Holyoak (1992)). However, he typically uses complicated learning tasks where effects temporally precede causes (diagnostic learning). This has apparently been motivated by the necessity of using diagnostic learning to compare the predictions of the R-W rule.

### Cue Competition in Contiguity Models

Some researchers propose that a noncompetitive, contiguity theory of learning may better accommodate cue competition for effects by asserting that cue competition does not arise during learning, but during later judgment (e.g., Matute et al., 1996). Further, Matute et al. (1996) found that the wording of test questions moderates the observance of cue competition for effects. They obtained cue competition for causes when they used test questions that implicitly probed the conditional probability of an effect given a cause compared to its probability given an alternative cause ( $p[E|C]$  with  $p[E|C']$ ), and they found cue competition for effects when they probed the conditional probability of a cause given an effect compared to its probability given an alternative effect ( $p[C|E]$  with  $p[C|E']$ ). Their work seems to suggest that the direction of the relationship queried may be related to whether evidence is found for cue competition.

### Cue competition in Recurrent Neural Networks

Recently, Read (2003) demonstrated that a recurrent network, based on McClelland and Rumelhart's (1988) auto-associator, with bidirectional links between the input and output nodes and using a modified version of the delta rule, can predict cue competition for both causes and effects. Unlike a feed-forward model, the recurrent model acquires bidirectional links or associations between the input and output nodes, and thus is able to accommodate cue competition for effects with predictive learning, where the cause *precedes* the effects.

One of the reasons this literature has become so confusing is that in order to use the R-W rule or a feed-forward network as a model of cue competition for effects, one must test cue competition for effects with diagnostic learning,

where effects are encountered before causes. However, with the recurrent network with delta-rule learning, this is not necessary. The current model allows one to test an associative model, closely related to R-W, with the more natural situation in which causes temporally precede effects.

In a recurrent neural network, the associative strengths of the bidirectional links between any two events may differ, and this possible asymmetry can be illustrated in the classic blocking paradigm. In the recurrent model, the observation of blocking depends on the direction of the association between the redundant cause B and the common effect X. If the association between cause A and X is trained in phase 1, then when A and B are subsequently paired preceding X, the link from B→X should exhibit competitive learning because, when B is activated, X is already activated due to the simultaneous presence of the previously trained cause A, which is a perfect predictor of X. Thus, there is little change in the link from B→X during Phase 2 training because B does not provide any new information about X. However, the link from X→B should not exhibit cue competition during Phase 2 training, as activating X does not predict B because B is not initially predicted by anything. Therefore, there is a great discrepancy between the target activation of B and the actual activation of B, which results in a greater weight change in the link from X→B. The result is that the link from X→B will be stronger than the link from B→X, suggesting that cue competition should only be observed in the link from B→X. As seen in Figure 1, Read demonstrated this asymmetry in weights using the recurrent model, with delta rule learning, with a learning rate of .15, with 10 passes through 10 learning instances with the same contingencies as in the current experimental stimuli.

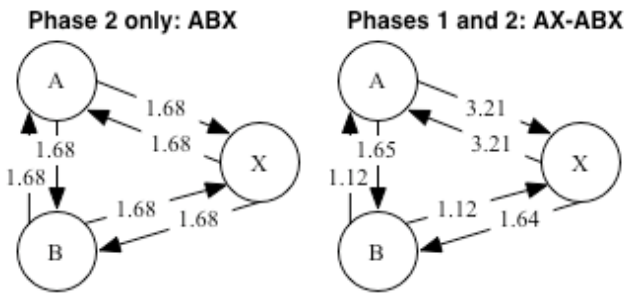


Figure 1: Weights demonstrating cue competition for causes after Phase 1 and 2 training (but not after Phase 2 only training). A and B are causes, X is the common effect.

Similarly, the recurrent model predicts cue competition for effects using the same rule. Again, the model predicts an asymmetry between the associative strengths of the links from A→Y and Y→A. The associative link from Y→A should exhibit competitive learning because when Y is activated during Phase 2 training, A is already highly activated from the simultaneous presence of X, which is a perfect predictor of A. Thus, there is very little weight change in the link from Y→A. However, the link from A→Y should not exhibit cue competition in learning because when A is activated during Phase 2 learning, Y is not initially predicted by anything. Therefore, the discrepancy between the target and the actual activations of Y are large, resulting in a bigger

weight change in the link from A→Y. Thus, the associative strength from A→Y should be stronger than the associative strength of Y→A. Read's simulation results in Figure 2, with the same parameters, reflect this as well.

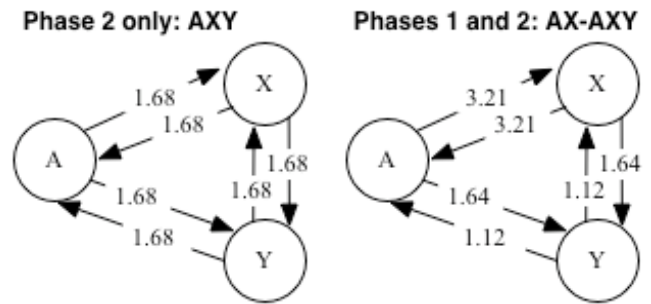


Figure 2: Weights demonstrating cue competition for effects after Phase 1 and 2 training (but not after Phase 2 only training). A is the cause and X and Y are the effects.

**Purpose.** Our purpose is threefold. First, we will provide further evidence that cue competition for effects occurs. Second, we will demonstrate that this effect can be obtained using social stimuli. Finally, we will test the predictions of the recurrent network model for cue competition and the asymmetry in the associative strengths of the bidirectional links between the cue and the target outcome. We will ask subjects to make judgments about the associative strengths of each of the two possible links between a cue and an outcome. Doing this in the same study has not previously been done. It is expected that cue competition for effects will be observed as well as the asymmetry in the associative strengths of the two possible links: the target outcome, Y, will exhibit cue competition with outcome, X, for the weight from Y→A, but not for the weight from A→Y.

We will get directional measures of strength by asking subjects to make conditional probability judgments between all pairs of nodes. Interpreting these judgments depends on the relationship between weights and conditional probabilities. As O'Reilly and Munakata (2000) show, in neural networks, the weight from an input  $i$  to an output  $o$  is a function of the conditional probability of the input given the output ( $p[i|o]$ ). The output node can be thought of as corresponding to a hypothesis and the input node to data concerning the hypothesis. Thus, the weight from input to output captures the conditional probability of the data given the hypothesis. The critical implication is that judgments of the conditional probability  $p[Y|A]$  will be sensitive to the strength of the weight from Y to A whereas judgments of  $p[A|Y]$  should be sensitive to the weight from A to Y.

### Study 1

Stimuli were four unrelated behaviors that had no known preexisting relationship with each other: breaking a glass (cause A); shaving one's head (effect X); lighting a tree on fire (effect Y); and meditating (filler effect Z). Information was presented simultaneously in list format (as Van Hamme et al. (1993) and Matute et al. (1996) did) with a dashed line break in between the antecedent and subsequent events to separate the common cause from the multiple effects.

Participants were instructed to learn the causal relationships between the antecedent and subsequent events. Finally, questions designed to assess the conditional probabilities and probe the associative strengths between all four behaviors were used ( $A \rightarrow X$ ;  $X \rightarrow A$ ;  $A \rightarrow Y$ ;  $Y \rightarrow A$ ;  $X \rightarrow Y$ ;  $Y \rightarrow X$ ;  $A \rightarrow Z$ ;  $Z \rightarrow A$ ).

## Method

**Participants.** 93 undergraduates from the University of Southern California volunteered for extra credit. The study was a between subjects design with repeated measures on judgments, where the experimental group received both Phase 1 and Phase 2 of training, and the control group only received the Phase 2 training.

**Materials and Procedure.** Subjects were randomly assigned to the experimental or control group and seated in front of a computer, on which the entire experiment was done. The cover story asked subjects to imagine that they were anthropologists in the distant future traveling to a long lost human colony on a faraway planet to study their culture and social customs. They were instructed that their goal was to learn the various behavioral patterns of the colonists by observing individual instances of sets of behaviors. They were instructed to learn the causal relationships among the behaviors. Before seeing the behaviors, all subjects made initial judgments about the extent to which the four stimulus behaviors were related to each other to establish that they had no known relationship. They were asked to rate the extent to which the occurrence of one behavior affected the likelihood of another behavior on a scale from -10 to 10, where -10 indicates “strongly inhibits,” 10 indicates “strongly promotes”, and 0 indicates “no relationship.” These questions and rating scale were also used for the testing after the training phase(s).

After the initial ratings, subjects in the experimental condition received Phase 1 training, where they saw 10 behavior sets exhibited by 10 individuals. Each set was presented individually on separate screens along with the name of the individual exhibiting the behaviors. Each set was displayed until the subject pressed the space bar. Eight of the 10 sets involved “breaking a glass” (cause A) followed by “shaving one’s head” (effect X); 2 of the 10 sets involved “breaking a glass” (cause A) followed by “meditating” (filler effect Z). The order of the 10 sets was randomized for each participant. After Phase 1, subjects started Phase 2 training, where they saw 10 more behavior sets exhibited by 10 new individuals. Eight of the 10 sets involved “breaking a glass” (cause A) followed by “shaving one’s head” (effect X) and “lighting a tree on fire” (effect Y). As in Phase 1, two of the 10 sets involved “breaking a glass (cause A) followed by “meditating” (filler effect Z).

Subjects in the control condition only received Phase 2 training. Immediately after training, they made judgments about the extent to which one behavior affects the likelihood of another behavior, for all four behaviors in both directions. Thus, subjects made judgments about the extent to which “breaking a glass” affects the likelihood of “shaving one’s head” ( $A \rightarrow X$ ); “shaving one’s head” affects the likelihood of “breaking a glass” ( $X \rightarrow A$ ); “breaking a glass” affects the

likelihood of “lighting a tree on fire” ( $A \rightarrow Y$ ); “lighting a tree on fire” affects the likelihood of “breaking a glass” ( $Y \rightarrow A$ ); “shaving one’s head” affects the likelihood of “lighting a tree on fire” ( $X \rightarrow Y$ ); “lighting a tree on fire” affects the likelihood of “shaving one’s head” ( $Y \rightarrow X$ ); “breaking a glass” affects the likelihood of “meditating” ( $A \rightarrow Z$ ); and “meditating” affects the likelihood of “breaking a glass” ( $Z \rightarrow A$ ).

## Results and Discussion

The mean of the eight initial ratings was -.62 for the experimental condition, and -.49 for the control, indicating that the four behaviors had no preexisting causal relationships. The mean of the eight final ratings was 5.77 for the experimental condition, and 6.19 for the control, indicating that subjects learned the causal contingencies.

A between-groups comparison for each of the eight final ratings found all of them to be non-significant with one exception. As predicted, the difference between experimental and control group judgments of the  $Y \rightarrow A$  rating ( $p[Y/A]$ ) was found to be highly significant,  $t(91)=3.02, p=.003$  (experimental  $M= 5.29$  vs. control  $M=7.21$ ) (See Figure 3). This provides evidence for cue competition between effects in the direction predicted by the recurrent network.

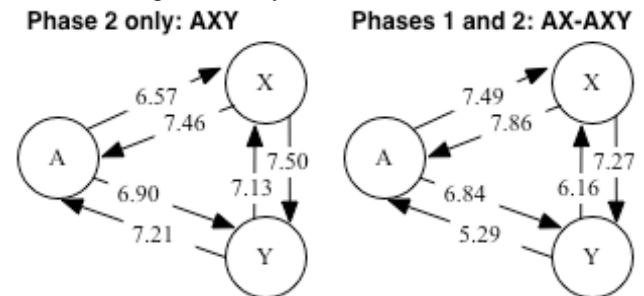


Figure 3: Mean ratings for cue competition for effects in the AXY (control) and AX-AXY (experimental) conditions. A is the cause, X and Y are the two competing effects.

Although the results of Study 1 were consistent with our predictions, we wondered whether the test questions did a good job of measuring conditional probabilities. After learning the behavioral contingencies, subjects were asked to make some “final estimates about the extent to which the occurrence of the first behavior affects the likelihood of the occurrence of the second behavior.” However, for the backward reasoning test questions (e.g.,  $X \rightarrow A$ ,  $Y \rightarrow A$ , and  $Z \rightarrow A$ ) the actual test questions asked subjects to indicate the likelihood that an individual doing X (or Y or Z) had previously done A. In other words, the task instructions implicitly asked subjects to make forward casual judgments from the first listed behavior to the second, while the actual test questions asked subjects to make backward causal judgments from the first listed behavior to the second (at least for the backward reasoning questions). Thus, it is unclear whether we were successful in measuring the bidirectional associative strengths among the four behaviors.

## Study 2

Study 2 involved several changes. First, the wording of the test questions was changed to more clearly measure

conditional probabilities. Test questions probing the associative strength of the link from  $A \rightarrow X$  were more clearly phrased as the probability of A given X ( $p[A|X]$ ),  $X \rightarrow A$  was more clearly phrased as the probability of X given A ( $p[X|A]$ ), and so forth (The phrasing of the new questions is presented later). Next, because the recurrent network model also predicts cue competition for causes as a result of asymmetrical associative strengths in the links between the redundant cue and outcome, the design from Study 1 was used to investigate cue competition for causes as well as for effects. For the Effects condition, the design and stimulus behaviors (Cause A, Effects X, Y and Z) were identical to those of Study 1. For the Causes condition, Cause A and Effect X remained the same with the addition of a new redundant Cause B, (“ringing a bell”) in the Phase 2 training portion, and changing the previous filler effect Z to filler cause Z. Finally, Study 2 was conducted on the web.

## Method

**Participants.** 168 adults ranging from ages 18 to 67 participated in this study on the Internet. Mean age was 39.28 (SD = 12.792). Participants were from previous on-line studies, unrelated to causal reasoning, who indicated that they were interested in future on-line studies. They were recruited by email and were residents of the US, with three exceptions. They were entered into a lottery for a \$50 cash prize, with the odds of winning at 1/50. The study was a between subjects design with repeated measures on judgments. As before, the experimental groups for causes and effects received both Phase 1 and Phase 2 training, and the control groups only received Phase 2 training.

**Materials and Procedure.** Subjects clicked a link in their email directing them to the study. Upon clicking the button that initiated the experiment, subjects were randomly assigned to one of four conditions: experimental and control conditions for Causes and experimental and control conditions for Effects. Subjects were presented with the same cover story as in Study 1. Because Study 1 showed that there were no preexisting relationships between the behaviors, the initial judgments were dropped for Study 2.

The procedures for the Effects conditions were identical to Study 1. Procedures for the Causes conditions were identical to those for Effects, with the exception of changes in the behaviors. In Phase 1 for the experimental group, 8 of the 10 sets involved “breaking a glass” (cause A) followed by “shaving one’s head” (effect X); 2 of the 10 sets involved “meditating” (filler cause Z) followed by “shaving one’s head” (effect X). In Phase 2, 8 of the 10 sets involved “breaking a glass” (cause A) and “ringing a bell” (redundant cause B) followed by “shaving one’s head” (effect X). As in Phase 1, 2 of the 10 sets involved “meditating” (filler cause z) followed by “shaving one’s head” (effect X). Control subjects only received Phase 2 training.

The order of the eight test questions was randomized for each subject. Immediately following training, subjects were asked to make final judgments about the extent to which the occurrence of one behavior affects the likelihood of the occurrence of another behavior in terms of conditional probabilities for all four stimulus behaviors both directions.

They made their judgments by clicking on a radio button on a rating scale from 0 to 10, where 0 indicates “No chance”, 5 indicates “50-50”, and 10 indicates “Certain”. The test questions were phrased: for  $p[A|X]$ , “Assuming that someone shaves their head, how likely is it that they had broken a glass,” and for  $p[X|A]$ , “Assuming that someone has broken a glass, how likely is it that they will shave their head?,” and so forth.

## Results and Discussion

As presented in Figure 4, the results for the Effects condition in study 2 replicate those of Study 1.

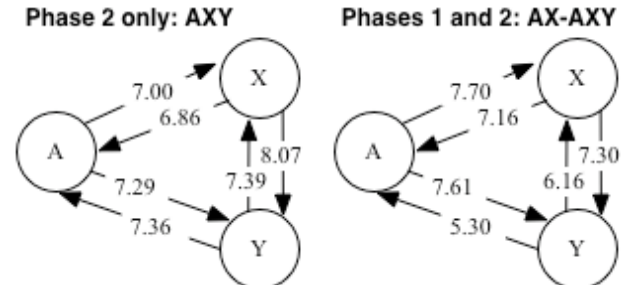


Figure 4: Mean ratings for cue competition for effects in the AXY and AX-AXY conditions in Study 2.

Between groups comparison of the final ratings for the experimental and control conditions showed a difference in the critical variable of  $Y \rightarrow A$  rating, ( $p[Y|A]$ )  $t(70)=3.75$ ,  $p=.00$  ( $M=5.30$  for experimental vs.  $M=7.36$  for control). No other comparisons were significant. Cue competition for effects was asymmetric in the predicted direction.

For the Causes condition, there was a significant difference between the experimental and control groups for the critical variable of  $B \rightarrow X$  rating ( $p[B|X]$ ),  $t(76) = 2.57$ ,  $p = .01$ . However, the difference for the non-critical variable of  $X \rightarrow B$  ( $p[X|B]$ ) was also significant,  $t(76) = 2.88$ ,  $p = .00$ , as presented in Figure 5. Cue competition for causes is obtained, but contrary to the network model’s predictions, there is no asymmetry in the associative strengths.

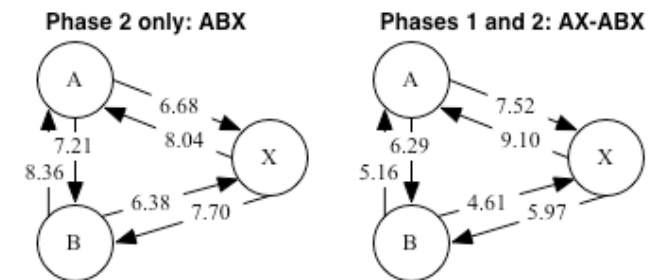


Figure 5: Mean final ratings for cue competition for causes in the ABX and AX-ABX conditions in Study 2. A and B are the competing causes, and X is the common effect.

## General Discussion

These results replicate the well-established phenomenon of competition between causes (e.g., Van Hamme et al., 1993; Waldman & Holyoak, 1992) as well as the more controversial competition between effects (Shanks, 1991; Price &

Yates, 1993; Matute et al., 1996). Further, these studies show that this effect can be obtained with social behaviors and is not limited to biological or physical events.

The present study is the first to study cue competition for causes and effects by systematically exploring all possible directional links between causes and effects. Studies 1 and 2 demonstrated that cue competition between effects occurs on the weight from the redundant effect Y to the cause A, rather than on the weight from the cause A to the redundant effect Y. However, Study 2 seemed to indicate that cue competition between causes occurs on both the weight from the redundant cause B to effect X as well as on the weight from effect X to the redundant cause B.

The results clearly contradict causal model theory, which states that effects do not compete. As for the associative learning and the neural network models, the results support their prediction of competitive learning and the presence of cue competition between effects.

One advantage of the recurrent model is that it provides an account of cue competition for effects without the necessity of requiring diagnostic learning (effects precede causes). Further, the recurrent model predicts and the current results confirm that the extent of cue competition depends on the direction of the weight or relationship between cue and effect. Previous models would have been unable to make such a prediction.

The results for cue competition between effects are consistent with the idea that the weights are sensitive to the conditional probabilities between causes and effects. These studies show that cue competition between effects occurs on the weights from Y to A and not from A to Y after AX-AXY training. This makes sense in terms of conditional probability, in that, taking both Phase 1 and 2 into account, every time Y was presented, it was always preceded by A, and thus  $p(A|Y)$  is 100%. (note the weight from A to Y should encode this conditional probability). However, whenever A was presented, Y followed A only half the time (during Phase 2), and thus  $p(Y|A)$  is 50% (the weight from Y to A should encode this conditional probability). Thus, the asymmetry in cue competition between effects is consistent with the conditional probabilities.

However, with regard to cue competition for causes the same does not seem to apply. A conditional probability analysis should predict a similar asymmetry. Instead, the results indicate no asymmetry in the bi-directional associative links between the redundant cause and the common effect; cue competition occurs when reasoning both from  $B \rightarrow X$  and from  $X \rightarrow B$ . It is unclear why we do not get weight asymmetry for cue competition for causes. However, in further research, using different social stimuli, we did find the predicted asymmetry, suggesting that the current results may be specific to the current stimuli.

We note one other caveat. Research in this area has not separated the effects of learning from the judgment process. Future studies in cue competition should be designed to examine the various types of processes that participants may use to arrive at their judgments of contingency.

## References

Cobos, P. L., Lopez, F. J., Cano, A., Almaraz, J. & Shanks,

- D. R. (2002). Mechanisms of predictive and diagnostic causal induction. *Journal of Experimental Psychology: Animal Behavior Processes*, 28, 331-346.
- Gluck, M. A. & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, 117, 225-244.
- Matute, H., Arcediano, F. & Miller, R. R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 182-196.
- McClelland, J. L. & Rumelhart, D. E. (1988). *Explorations in parallel distributed processing: A handbook of models, programs, and exercises*. Cambridge, MA: MIT Press/Bradford Books.
- Mels, E. R., Cheng, P. W., Holyoak, K. J., & Waldmann, M. R. (1993). Cue competition in human categorization: Contingency or the Rescorla-Wagner learning rule? Comment on Shanks (1991). *Journal of Experimental Psychology: Learning, Memory and Cognition*, 19, 1398-1410.
- Miller, R. R., Barnet, R. C. & Grahame, N. J. (1995). Assessment of the Rescorla-Wagner Model. *Psychological Bulletin*, 117, 363-386.
- Miller, R. R. & Matute, H. (1998). Competition between outcomes. *Psychological Science*, 9, 146-149.
- O'Reilly, R. C., & Munakata, Y. (2000). *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.
- Price, P. C. & Yates, J. F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 21, 1637-1655.
- Read, S. J. (2003). *An integrative model of causal learning and causal reasoning using a feedback neural network*. Unpublished manuscript, U. of Southern California.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64-99). New York: Appleton-Century-Crofts.
- Shanks, D. R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology*, 37B, 1-21.
- Shanks, D. R. & Lopez, F. J. (1996). Causal order does not affect cue selection in human associative learning. *Memory & Cognition*, 24, 511-522.
- Van Hamme, L. J. & Wasserman, E. A. (1993). Cue competition in causality judgments: The role of manner of information presentation. *Bulletin of the Psychonomic Society*, 31, 457-460.
- Waldmann, M. R. (2000). Competition among causes but no effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 26, 52-76.
- 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.