Title
Interference in Human Predictive Learning when Associations Share a Common Element

Permalink
https://escholarship.org/uc/item/8pw3f0st

Journal
International Journal of Comparative Psychology, 13(1)

ISSN
0889-3675

Authors
Pineño, Oskar
Matute, Helena

Publication Date
2000

DOI
10.46867/C4BS3P

Copyright Information

Peer reviewed

Interference in Human Predictive Learning when Associations Share a Common Element

Oskar Pineño and Helena Matute
Universidad de Deusto, Spain

Interference between cues is generally regarded as different from interference between outcomes in the (animal and human) predictive learning literature. In consequence, current theories of learning generally attempt to explain one or the other effect but not both. In general, cues are assumed to interfere with each other only if they are trained in compound as predictors of the same outcome, whereas outcomes are assumed to interfere with each other only if they have been individually paired to the same cue. In two experiments with humans, we examined the parallel between interference between elementally-trained cues and between elementally-trained outcomes, using a behavioral preparation. Experiment 1 showed that both interference effects are similarly affected by identical contextual manipulations. Experiment 2 showed that the two effects take place when the interfering association shares an element with the target association: When the shared element is the outcome, interference between cues takes places; when the shared element is the cue, interference between outcomes occurs. These results add to the growing body of evidence that calls for the integrative study of interference between cues and between outcomes in predictive learning situations.

In the last two decades, there has been an increasing interest in the study of human causal and predictive learning in the same framework used in animal conditioning research (e.g., Allan, 1980; Dickinson, Shanks, & Evenden, 1984; Wasserman, Chatlosh, & Neunaber, 1983) as well as an interest in the study of animal conditioning in the same framework as that used to study human causal and predictive learning (e.g., Miller & Matute, 1996a, 1996b). Within the associative framework, the main difference between human and animal research on learning is that animal experiments normally use biologically significant stimuli (i.e., unconditioned stimuli), whereas human studies normally use biologically neutral stimuli (e.g., Miller & Matute, 1996a, 1996b). Thus, for example, an animal study can use flavors that become associated with illness, and the experimenter can assess the strength of the resultant conditioned response to infer the strength

of the association. An analogous human study can use, for example, fictitious meals that become associated with fictitious illness. The dependent variable that will be used in this case to infer the strength of the association can be, for example, the causal or predictive judgment, that is, the degree to which the subject expects the outcome to occur when the fictitious patient ingests the fictitious meal (e.g., Dickinson & Burke, 1996; Matute, Arcediano, & Miller, 1996; Wasserman, 1990). Another type of dependent variable that can also be used in human research is the number of instrumental responses (e.g., pressing the space-bar of a computer keyboard) that the subject gives in response to stimuli that predict reinforcement (e.g., Pineño, Ortega, & Matute, 2000). This later dependent variable is more similar to those used in animal conditioning research and is the one used in the present series of experiments.

Some of the effects that have been most prominently investigated with both humans and other animals in associative learning research are interference effects between cues and between outcomes. Many studies have shown that when two cues predict the same outcome, the two cues interfere with each other so that, if one of them is a strong predictor of the outcome, responding to the other is impaired. Quite symmetrically, other studies of predictive learning have shown that when two different outcomes are predicted by the same cue, and one of these outcomes is being strongly predicted at a given moment, this interferes with appropriate responding for the other outcome. Examples of interference between cues (most commonly called competition between cues) are blocking (Kamin, 1968), overshadowing (Pavlov, 1927), and the cue-validity effect (Wagner, Logan, Haberlandt, & Price, 1968). Examples of interference between outcomes are extinction and counterconditioning (Pavlov, 1927). Although these effects have traditionally been studied in the area of animal conditioning, they have been also reported in the area of human predictive learning (see, e.g., Dickinson et al., 1984; Matute et al., 1996; Wasserman, Kao, Van Hamme, Katagari, & Young, 1996 for interference between cues; Paredes-Olay & Rosas, 1999; Vila, 2000, for interference between outcomes). To our knowledge, however, none of the studies in the predictive learning tradition has compared interference between cues and between outcomes within a single experiment.

Indeed, current theories of (human and nonhuman) learning generally tend to deal with one or the other interference effect but not both. For example, the Rescorla-Wagner (1972) model is generally regarded as providing a compelling explanation of interference between cues, but it makes the wrong predictions with respect to interference between outcomes (see, e.g., Rescorla, 1996a, 1996b). Conversely, theories designed to explain interference between outcomes do not attempt to explain interference between cues (e.g., Bouton, 1993).

At first glance, interference between cues and between outcomes might seem to have little to do with each other. Interference between cues is generally shown in predictive learning when two cues (A and B) are trained in compound as predictors of the same outcome (overshadowing: AB-O_1), and interference between outcomes is generally shown when two or more outcomes (O_1 and O_2) are independently paired to the same cue in different phases of the study (counterconditioning: Training with A-O_1 during Phase 1 and with A-O_2 during Phase 2). However, both interference between elementally-trained cues and between elementally-trained outcomes was largely studied in the verbal learning tradition under the A-B C-B (elementally-trained cues) and A-B A-C (elementally-trained outcomes) paradigms (Underwood, 1966). Although the paradigm that is most frequently used in current studies of interference is the A-B A-C paradigm (e.g., Bäuml, 1996, 1998; Chandler, 1993; Chandler & Gargano, 1998), many studies show interference between elementally-trained cues (A-B C-B paradigm) in the literature of the 1960s and 1970s (Abra, 1967; Cheung & Goulet, 1968; Johnston, 1968; Keppel, Bonge, Strand, & Parker, 1971; Schwartz, 1968).

Thus, interference between elementally-trained outcomes in predictive learning (e.g., extinction and counterconditioning) can be regarded as an instance of what was known in the paired-associate tradition as the A-B A-C interference paradigm, and it can be readily integrated within the framework of a general theory of interference (see Bouton, 1993). By contrast, cues that are independently trained (A-B C-B paradigm in the paired associate tradition) are not supposed to interfere with each other in predictive learning. These discrepancies are the main focus of the present research.

It could be argued that the similarities observed in the verbal learning tradition concerning interference between cues and between outcomes might not be relevant to predictive learning because the paired associate literature was concerned with language-related issues rather than with predictiveness. However, recent research on predictive learning is showing that cues that are trained independently of each other as predictors of the same outcome can interfere with each other. That is, learning a cue, B, as a predictor of the same outcome that was previously predicted by another cue, A, can impair responding to A at testing (Matute & Pineño, 1998a, 1998b). Moreover, interference between cues and between outcomes seems to be affected in a similar way by similar manipulations. For example, it is well known that if a change of context or a retention interval occurs between training and testing in an extinction or a counterconditioning experiment with animals (which can be collectively represented as A-O_1 training during Phase 1 followed by A-O_2 training during Phase 2), renewal of the response or spontaneous recovery can respectively occur at test (Bouton, 1993; Pavlov, 1927).

Several experiments have provided data suggesting that similar effects can take place in interference between cues in both humans (Matute & Pineño, 1998a, 1998b; Ortega & Matute, 2000; Pineño et al., 2000) and rats (Escobar, Matute, and Miller, 2001). For example, Matute and Pineño (1998a) showed that the interference between cues that occurred when A was tested after A-O training in Phase 1 followed by B-O training in Phase 2, did not occur (i.e., the response to A was renewed) when testing occurred in a context that was different from that in which the interfering cue, B, had been trained. Additionally, Pineño et al. (2000) showed that the introduction of a retention interval before testing recovered responding to A. These authors also showed that responding to A at test could be recovered by inserting either a retrieval cue for the A-O trials of Phase 1 or a novel cue just before testing, whereas inserting a retrieval cue for the interfering B-O trials did not recover responding to A at test. These results on interference between cues are analogous to those of Bouton and his colleagues on interference between outcomes, with rats, which show recovery from interference between outcomes (e.g., recovery from extinction) due to the introduction of a cue that was presented during training of the target association (Bouton, 1993; Brooks, 2000; Brooks & Bouton, 1993; Brooks, Palmatier, Garcia, & Johnson, 1999). Nevertheless, because different dependent variables and animal species have generally been used in each of those two sets of studies (i.e., rat subjects are normally used in extinction and counterconditioning experiments and human subjects are most frequently used in the studies of interference between elementally-trained cues), it is difficult to make a direct comparison between them and to infer, for instance, the relative strength of each of the two interference effects.

The present experiments were designed to provide such direct comparisons. Two experiments with human subjects investigated the parallel between both types of interference. More specifically, we studied the parallel between interference among elementally-trained cues and experimental extinction. The first experiment studied whether interference between cues and between outcomes was affected in a similar way by contextual manipulations such as those that have been frequently used in the study of interference between outcomes in animal research (e.g., Bouton, 1993, 1994; Bouton & Bolles, 1979). That is, we used a renewal design to assess the parallel between these two effects. The second experiment aimed at showing that both types of interference take place only when the target and the interfering associations share a common element, being this the cue (in the case of interference between elementally-trained outcomes) or the outcome (in the case of interference between elementally-trained cues).

Experiment 1

Experiment 1 was designed to study whether contextual manipulations affect the two types of interference in a similar way. Specifically, we used a renewal design in this experiment. The renewal effect has been demonstrated in the framework of interference between outcomes with rats (e.g., Bouton, 1993, 1994) and humans (e.g., Paredes-Olay & Rosas, 1999), and consists of the strong responding that is observed when the target cue is tested outside of the context in which the interfering association was trained. More specifically, we used an XYX renewal design, and applied it to both interference between cues and between outcomes: We tested the $A\text{-}O_1$ association in the context in which $A\text{-}O_1$ had been trained (Context X), after the interfering training ($A\text{-}O_2$, in the case of interference between outcomes; or $B\text{-}O_1$, in the case of interference between cues) was given either in the same context (Context X) or in a different context (Context Y).

Method

Participants and Apparatus. Forty-eight undergraduate students from Deusto University volunteered for the study. Participants were randomly assigned to one of four groups ($n = 12$). The experiment was run using personal computers and participants were run in individual cubicles.

Procedure. Table 1 summarizes the design of this experiment. During Phase 1, all groups were exposed to identical treatment in Context X: Fifteen presentations of cue A, which was always followed by O_1 (i.e., $A\text{-}O_1$), interspersed with fifteen presentations of cue C, which was always followed by O_2 (i.e., $C\text{-}O_2$). (Presentations of $C\text{-}O_2$ trials were included in order to prevent stimulus generalization that would result in strong responding appropriate to O_1 to all cues.) In Phase 2, Groups Cues-Same and Cues-Diff were exposed to fifteen presentations of a different cue, B, predicting O_1 (i.e., $B\text{-}O_1$). In contrast, Groups Outcomes-Same and Outcomes-Diff were exposed to fifteen presentations of a different Outcome, O_3, predicted by A (i.e., $A\text{-}O_3$). Orthogonally, Groups Cues-Same and Outcomes-Same were exposed to Phase 2 in Context X and Groups Cues-Diff and Outcomes-Diff were exposed to Phase 2 in Context Y. Then, in the test phase, all groups were exposed to a single presentation of A in Context X. There were no interruptions between the different phases of the experiment.

Table 1
Design Summary of Experiment 1

| | Treatment | | |
Group	Phase 1	Phase 2	Test
Cues-Same	X: $A\text{-}O_1$, $C\text{-}O_2$	X: $B\text{-}O_1$	X: A
Cues-Diff	X: $A\text{-}O_1$, $C\text{-}O_2$	Y: $B\text{-}O_1$	X: A
Outcomes-Same	X: $A\text{-}O_1$, $C\text{-}O_2$	X: $A\text{-}O_3$	X: A
Outcomes-Diff	X: $A\text{-}O_1$, $C\text{-}O_2$	Y: $A\text{-}O_3$	X: A

The preparation used in this experiment is the same as that used in the experiments reported by Pineño et al. (2000), but different from that used in the experiments reported by Matute and Pineño (1998a). The reason to use different preparations is to ensure that the effects are not task specific. The preparation used by Matute and Pineño (1998a) consisted of a Martians video game in which the dependent variable was the degree to which responding was suppressed when the critical cues were presented (see Arcediano, Ortega, & Matute, 1996, for a detailed description of this preparation). In contrast, the present task required the participants to rescue a group of refugees by helping them to escape from a war zone in several trucks. The dependent variable was the mean number of responses given during cue presentations - that is, just the opposite to that used in Matute and Pineño. In addition, this new preparation seems to be more sensitive to a variety of effects. A translation from Spanish of the instructions that the participants received reads as follows:

Screen 1

Imagine that you are a soldier for the United Nations. Your mission consists of rescuing a group of refugees that are hidden in a ramshackle building. The enemy has detected them and has sent forces to destroy the building... But, fortunately, they rely on your cunning to escape the danger zone before that happens.

You have several trucks for rescuing the refugees, and you have to place them in those trucks. There are two ways of placing people in the trucks:
a) *Pressing the space bar repeatedly, so that one person per press is placed in a truck.*
b) *Maintaining the space bar pressed down. In such manner, you will be able to load people very rapidly.*

If you rescue a number of persons in a given trip, they will arrive at their destination alive, and you will be rewarded with a point for each person. You must gain as many points as possible!

Screen 2

But... your mission will not be as simple as it seems. The enemy knows of your movements and could have placed deadly mines on the road. If the truck hits a mine, it will explode, and the passengers will die. Each dead passenger will count as one negative point for you.

Fortunately, the colored lights on the SPY RADIO will indicate to you the state of the road. The lights can indicate that:
a) *The road will be free of mines. $\rightarrow$ The occupants of the truck will be liberated. $\rightarrow$ You will gain points.*
b) *The road will be mined. $\rightarrow$ The occupants of the truck will die. $\rightarrow$ You will lose points.*
c) *There are no mines, but the road is closed. $\rightarrow$ The occupants of the truck will neither die nor be liberated. $\rightarrow$ You will neither gain nor lose points: You will maintain your previous score.*

Screen 3

At first, you will not know what each colored light of the SPY RADIO means. However, as you gain experience with them, you will learn to interpret what they mean.

Thus, we recommend that you:
a) *Place more people in the truck the more certain you are that the road will be free of mines (keep the space bar continuously pressed down ONLY if you are completely sure that there are no mines, because in this way you will put a lot of people in the truck...).*
b) *Introduce fewer people in the truck the more certain you are that the road is mined.*

After these instructions, participants were shown a fourth screen that gave instructions about contextual changes. Although contextual changes are not being used in the second experiment, in order to avoid making more changes than necessary between different experiments conducted with the same preparation, we maintained the four instructional screens of this program in both experiments. A translation of the fourth screen reads as follows:

Screen 4

Finally, it is important to know that your mission may take place in several different towns. The colors on the SPY RADIO can mean the same or a very different thing depending on the town in which you are. Thus, it is important to pay attention to the message that indicates the place in which you are. If you travel to another town, the message indicating the name of the town will change. When a change of destination is occurring, you will read the message "Traveling to

another town," so you will be continuously informed about such changes. Nevertheless, sometimes you might end up returning to the same town even if you have seen the message that indicates that you are traveling. Do not worry if all this looks like very complex at this point. Before we start, you will have the opportunity to see the location of everything (radio, town name, messages, scores, etc.) on the screen, and to ask the experimenter about anything that is unclear.

The top of the screen showed a "spy radio" that consisted of a panel in which six colored lights could be presented. Cues A, B, and C were blue, red, and yellow lights in the Spy Radio, counterbalanced. In this experiment, each time that a cue was presented, all six panels were illuminated with the color of that cue. Cue duration was 3 s. During the intertrial intervals (ITIs), the lights were turned off (i.e., gray). The ITI duration was random with a range between 3 and 7 s, and a mean of 5 s. While the lights were on, each response (i.e., pressing the space bar once) placed one refugee in the truck. If the participant maintained the space bar pressed down while the lights were on, up to 10 refugees per second could be placed in the truck in this experiment. On each trial, the offset of the cue coincided with the onset of the outcome.

Outcome 1 (O_1) consisted of (a) the message "[N] refugees safe at home!!!" (with [N] being the number of refugees introduced in the truck during the cue presentation) and, (b) gaining one point for each refugee who was liberated. Outcome 2 (O_2) consisted of (a) the message "[N] refugees have been killed!!!" and, (b) losing one point for each refugee who died in the truck. Outcome 3 (O_3) consisted of (a) the message "Road closed" and, (b) maintaining previous score. Outcome messages were presented for 3 s. Outcomes were not counterbalanced because in order to test interference between outcomes (i.e., extinction or counterconditioning) in the two G Groups Outcomes, the outcomes used in Phase 1 and Phase 2 need to have opposite motivational value and be consistent for all participants. That is, participants should learn to respond to Cue A during Phase 1 (A-O_1), and should decrease their responding to Cue A during Phase 2 (A-O_3). It is by doing so that we can expect weak responding at test in those participants that received extinction treatment (or counterconditioning, if O_3 is interpreted as a different outcome rather than as the absence of O_1) in the same context, and stronger responding at test in those participants that received extinction training in a different context. Otherwise, strong or weak responding at test would depend on outcome counterbalancing rather than on whether the extinction treatment is given in the same or in a different context.

The number of refugees that participants risked taking in each truck was our dependent variable. Presumably, the more certain they were that the trip would be successful (i.e., O_1), the greater number of refugees they would take; whereas the more certain they were that the truck would explode, the smaller number of refugees they would take (i.e., O_2). Additionally, because introducing refugees did not have any effect on the score when O_3 followed the cue, we expected the participants to extinguish their responding when the cue was followed by O_3.

One score panel on the screen provided information during the experiment. The panel showed the number of people that the participant was introducing in the truck on each trial. Although bar presses that occurred while the outcome message was present had no consequences, this panel remained visible during the presentation of the outcome and showed the number of people that had been boarded while the cue was present. At the offset of the outcome, this panel was set to zero. Responses that occurred during the ITIs had no consequences and were not reflected in the panel. Thus, only responses that occurred while a cue was presented resulted in refugees traveling in the truck and participants gaining or losing points if O_1 or O_2 followed the cue.

Contexts X and Y consisted of two fictitious towns, Bow Town and Hal Land, counterbalanced. The name of the town was shown at all times in the middle of the screen, and these towns were given contextual status via instructions (see Screen 4 of instructions). When a context change was taking place, the message "Traveling to another town" was shown for 3 s in the middle of the screen. In order to control for the potential influence of the message of context change itself, this message was also presented to those groups for which there was no change of context (but recall that the instructions told participants that, despite seeing this message, they could end up returning to the same town). For better clarity concerning this procedure, a demonstration version of the programs here mentioned can be downloaded from http://sirio.deusto.es/matute/software.html

Preanalysis Treatment of the Data. We normally use a data selection criterion in order to ensure that participants are paying attention to the experiment and have acquired the

discrimination during Phase 1. According to this criterion, the number of responses given to A during the last trial in which it is presented during Phase 1 of training has to be higher than the number of responses given to the last trial of C. Following this criterion, no participant was eliminated from this experiment.

Results and Discussion

As expected, contextual manipulations affected interference between elementally-trained cues and between elementally-trained outcomes in a similar way. For both types of interference effects, responding to A at test was stronger when testing took place in the context in which the target association had been trained, as compared to those conditions in which testing took place in the context of Phase 2 (i.e., interfering) training.

The critical results of this experiment are depicted in Figure 1, which shows the mean number of responses to A at test. As can be appreciated in this figure, responding to A was stronger in Groups Cues-Diff and Outcomes-Diff, when compared to Groups Cues-Same and Outcomes-Same, respectively. A 2 x 2 (Type of Interference x Context of Phase 2) analysis of variance (ANOVA) on responding to A at test yielded a main effect of context of phase 2, $F(1, 44) = 22.80$, $p<.001$. The type of interference and the interaction were not significant, $ps>.5$. Planned comparisons showed that responding in Group Cues-Same was weaker than in Group Cues-Diff, $F(1, 44) = 11.25$, $p<.01$, and responding in Group Outcomes-Same was weaker than in Group Outcomes-Diff, $F(1, 44) = 11.55$, $p<.01$.
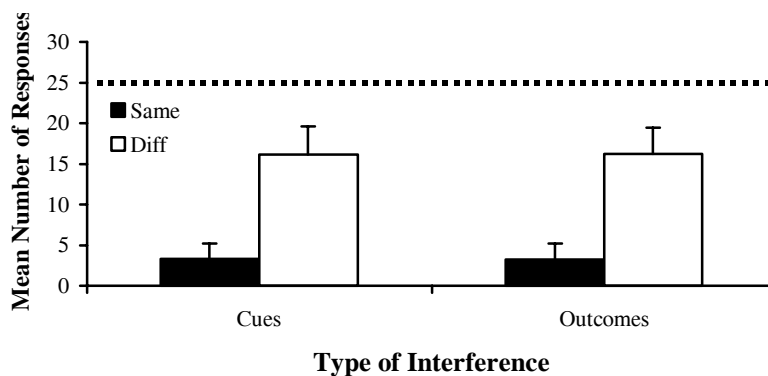

Figure 1. Experiment 1. Mean (+ SEM) number of responses to A during the test trial. For reference, the horizontal dotted line represents the mean number of responding given to A in the last 5-trial block of training during Phase 1. In this experiment, the maximum number of responses that could be recorded during the cue presentation was 30.

Importantly, the observed differences in responding to A at test can not be attributed to a different level of responding to A during training in the

different groups. The mean number of responses to A in the last 5-trial block of training was 25.83 ($\pm$0.31), 25.86 ($\pm$0.29), 24.93 ($\pm$0.54), and 25.25 ($\pm$0.77), for Groups Cues-Same, Cues-Diff, Outcomes-Same, and Outcomes-Diff, respectively. A 2 x 2 (Type of Interference x Context of Phase 2) ANOVA carried out on responding to A in the last 5-trial block of training yielded no main effects or interaction (all ps>.1). For reference, the dotted line in Figure 1 shows the mean number of responses given to A in the last 5-trial block of training, collapsed for all four groups.

Counterbalancing did not affect responding at test either, as shown by a one-way ANOVA on responding to A during testing, $F < 1$. Also, the mean number of responses during the 3 s that preceded testing was zero in all the groups. Therefore, differential baseline responding can also be discarded as a potential explanation of the observed results.

Thus, the results of this experiment show that contextual manipulations affected interference between cues and interference between outcomes almost identically. When testing of A was given in the same context in which training of the interfering association (either B-O_1 or A-O_3) took place, responding to A at test was weak. However, when the interfering association was trained in a different context, strong responding to A at test was observed. This was observed regardless of whether the competing association was B-O_1 (i.e., interference between elementally-trained cues) or A-O_3 (i.e., interference between elementally-trained outcomes). Thus, the contextual manipulations that are frequently reported in animal studies of interference between outcomes are also shown to affect human interference between cues and between outcomes.

Experiment 2

Experiment 1 showed that contextual manipulations affect analogously the effects of interference between cues and between outcomes. The second experiment was designed to study whether, as Matute and Pineño (1998b) suggested, these two types of interference take place only when the interfering association shares a common element with the target association. Following their rationale, when the shared element is the outcome, an effect of interference between elementally-trained cues occurs. On the other hand, an effect of interference between outcomes occurs when the shared element is the cue. Although both types of effects have been demonstrated in different studies, this parallelism remains to be demonstrated in a single experiment.

Method

Participants and Apparatus. Thirty-six undergraduate students from Deusto University volunteered for the study. Participants were randomly assigned to one of three groups (n = 12).

The apparatus was identical to that used in Experiment 1, except that the software was run under different computers and operating system, which speeded up the introduction of refugees in the truck (i.e., the maximum number of refugees that could be introduced in each trial was elevated from 10 refugees per second in Experiment 1 to about 30 refugees per second in this experiment).

Procedure. Table 2 summarizes the design for this experiment. During Phase 1, all groups were exposed to identical treatment: Fifteen presentations of cue A, which was always followed by O_1 (i.e., A-O_1), interspersed with fifteen presentations of cue C, which was always followed by O_2 (i.e., C-O_2). In Phase 2, Group Cues was exposed to fifteen presentations of cue B, which was always followed by O_1 (i.e., B-O_1). Group Outcomes was exposed to fifteen presentations of cue A, which was always followed by O_3 (i.e., A-O_3). Finally, Group Control was exposed to fifteen presentations of cue B, which was always followed by a different outcome, O_3 (i.e., B-O_3). Thus, in this control condition, the second-learned association did not share any element with the target, A-O_1, association. An alternative to this control group could have been to use a group receiving no training at all during Phase 1. However, in this experiment, we had two experimental groups that were to acquire during Phase 2 an association that shared either the cue or the outcome with the target association acquired during Phase 1. Thus, it seemed important to use as a control group a group that received the same number of interfering pairings during Phase 2, but to ensure that these interfering pairings shared neither the cue nor the outcome with the association acquired during Phase 1. This control is equivalent to the A-B C-D control condition frequently used in the paired-associate literature.

Then, in the test phase, all groups were exposed to a single presentation of A. There were no interruptions between the different phases of the experiment.

Preanalysis Treatment of the Data. In this experiment, the same data selection criterion as described in Experiment 1 was used. Following this criterion, no participant was eliminated from this experiment.

Table 2
Design Summary of Experiment 2

| | Treatment | | |
Group	Phase 1	Phase 2	Test
Cues	A-O_1, C-O_2	B-O_1	A
Outcomes	A-O_1, C-O_2	A-O_3	A
Control	A-O_1, C-O_2	B-O_3	A

Results and Discussion

The critical results are depicted in Figure 2, which shows the mean number of responses to A at test in each of the three groups. This figure suggests that Groups Cues and Outcomes, for which the interfering association shared one element with the target association, showed weak responding at test compared to Group Control, for which the two associations had no common elements. These impressions were confirmed by a one-way ANOVA on the test of A, which showed an overall Group effect, $F(2, 33) = 22.00$, $p<.01$. Planned comparisons showed that responding in Group Cues was weaker than responding in Group Control, $F(1, 33) = 17.77$, $p<.01$, and that responding was weaker in Group

Outcomes than in Group Control, $F(1, 33) = 42.81$, $p<.01$. Thus, interference both between elementally-trained cues and between elementally-trained outcomes were replicated in this experiment when the interfering association shared the outcome or the cue with the target association.

Importantly, the observed differences in responding to A during the test cannot be attributed to a different level of responding to A during training in the three groups. The mean number of responses to A in the last 5-trial block of training was 66.38 ($\pm$4.09), 63.10 ($\pm$1.72), and 67.08 ($\pm$1.84), for Groups Cues, Outcomes, and Control, respectively. A one-way ANOVA on responding to A in the last 5-trial block of training showed that these differences were not reliable, $F<1$. For reference, the dotted line in Figure 2 shows the mean number of responses to A in the last 5-trial block of training collapsed for all three groups.

Counterbalancing did not affect responding at test in this experiment either, as shown by a one-way ANOVA on responding to A during testing, $F<1$. Also, the mean number of responses during the 3 s that preceded testing was zero in all the groups. Therefore, differential baseline responding can also be discarded as a potential explanation of the observed results.
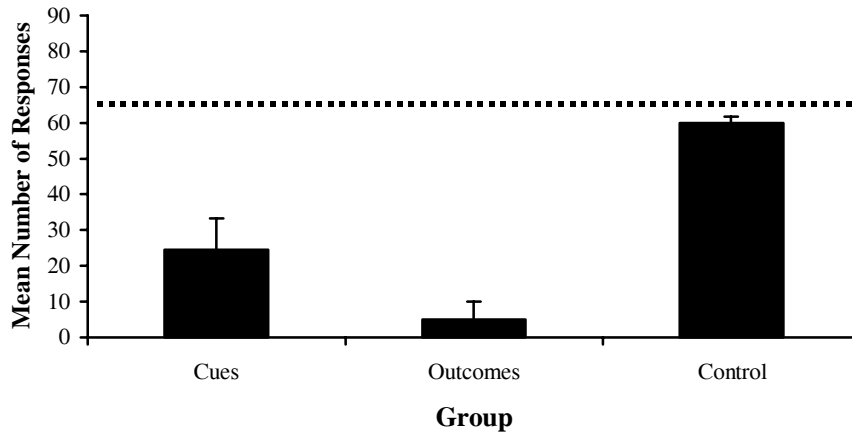

Figure 2. Experiment 2. Mean (+ SEM) number of responses to A during the test trial. Error bars represent standard error of means. For reference, the horizontal dotted line represents the mean number of responding given to A in the last 5-trial block of training during Phase 1. Because in this experiment the software was run under different computers and operating system than those used for Experiment 1, the maximum number of responses that could be recorded during the cue presentation was 90 rather than 30.

Overall, these results show that, as suggested by Matute and Pineño (1998b), interference takes place when the interfering association shares an element with the target association, be it the cue or the outcome.

General Discussion

The present experiments studied the parallel between the effects of interference between elementally-trained cues and between elementally-trained outcomes in human participants. Experiment 1 used contextual manipulations in order to study whether those manipulations affected both effects analogously. That experiment showed that a contextual switch before testing renewed responding to the target cue, A, in both effects. Experiment 2 showed that interference takes place when the interfering association shares a common element with the target association: When the shared element is the cue, interference between outcomes occurs; when the shared element is the outcome, interference between cues takes place.

The results of the Groups Outcomes replicate previous findings of interference between outcomes that had been studied primarily with non-human animals (e.g., Bouton, 1993); the results of the Groups Cues replicate previous findings of interference between cues, which had been studied primarily with humans (e.g., Matute & Pineño, 1998a). The present findings are also consistent with results reported in the human paired-associates literature (i.e., A-B A-C and A-B C-B paradigms; see Introduction), and show that those effects can also be found in predictive learning. The conjoint study of the two effects in the present experiments highlights the similarities between interference between cues and between outcomes in predictive learning as well.

It is not implied here that interference between elementally-trained cues can only be observed in humans. Indeed, Escobar et al. (2001) have recently replicated with rats the original findings with humans that cues trained apart can interfere with each other for behavioral control. Embedding the cue interference treatment into a sensory preconditioning procedure (e.g., Rizley & Rescorla, 1972), Escobar et al. have shown that rats are also subject to the effect of interference between elementally-trained cues. Nevertheless, it is important to note that the use of sensory preconditioning in rats was shown to be critical in order to obtain the effect, which was not observed using regular conditioning, rather than sensory preconditioning (Escobar et al.). This is consistent with recent work showing that cues that have biological significance (either inherent biological significance because they are unconditioned stimuli, or acquired biological significance because they have been paired to unconditioned stimuli) are resistant to cue interference effects such as forward blocking, backward blocking, cue-validity, and the contingency-degradation effect (Denniston, Miller, & Matute, 1996; Miller & Matute, 1996b; Oberling, Bristol, Matute, & Miller, 2000). The work of Escobar et al. shows that cues that are biologically significant (i.e., conditioned stimuli) are also resistant to the effect of interference between elementally-trained cues. However, cues that

are biologically irrelevant (such as those used in their sensory preconditioning procedure with rats, or the cues used in the present experiments with humans) can interfere with each other under certain circumstances, even if they are trained apart.

Given the observed similarity between interference between cues and between outcomes, it is appropriate to seek a unified explanation for both effects. Indeed, these two effects can be seen as symmetrical: Either two different cues predict the same outcome, or one cue predicts two different outcomes. Regardless of whether the common element is the cue or the outcome, as shown by Experiment 2, interference occurs when the interfering association has a common element with the target one. Moreover, Experiment 1 showed that, if the competing association is trained in a second context, interference is not observed in either case. This integrates previous findings showing that interference between elementally-trained cues (Matute & Pineño, 1998a, 1998b; Pineño et al., 2000) and between elementally-trained outcomes (e.g., Bouton, 1993) does not occur when testing occurs in a context different from that used for interfering training, or when a retention interval or a retrieval cue is inserted between interfering training and testing. That is, the two effects seem to take place during the expression of the association, if the interfering association is more strongly activated in memory at the time of testing than the target one. Moreover, the stronger activation of the interfering association at testing that seems to be necessary for interference to occur, can take place, not only through the use of contextual manipulations and retrieval cues, but also through greater recency: The effect does not occur if the two cues are trained within the same phase of the study or when the interfering association is trained prior to training with the target association (Matute & Pineño, 1998a; Pineño et al., 2000). Apparently, the stronger activation of the interfering association during testing interferes with the expression of the target association.

With respect to interference between outcomes, Bouton (1993) suggested that when two outcomes are predicted by the same cue, the stronger expression of one association interferes with the expression of the other association. On the other hand, Bouton also proposed that the context of testing (the physical context, the temporal context, or both) serves the purpose of modulating the expression of one or the other association. Following Bouton's theory, the second-learned association is context specific, whereas the first-learned association easily transfers to a new context (see, e.g., Bouton, 1997; Bouton & Nelson, 1998). Thus, interference between outcomes takes place only when the test context activates the expression of the interfering association. In other words, if testing takes place in a context that is different from that in which the second-learned association was trained, this association will no longer

interfere with the expression of the first-learned association, leading to a recovery of responding appropriate to the first-learned association. The results of our Groups Outcomes support of Bouton's theory, since we observed responding appropriate to the first-learned association when the second-learned association was trained in a second context and testing took place in the context in which the first-learned association had been trained. On the other hand, we observed responding appropriate to the interfering association when the context was not changed, and thus testing took place in the same context in which the more recent interfering association had been trained. Most importantly, this rationale can be extended to account for Groups Cues' result as well: When two different cues are associated with the same outcome representation in different contexts, the context of testing will activate the expression of the second-learned association (i.e., Cue_2-O) only when this context is the same as that in which the Cue_2-O association was trained. In addition, if the context of testing is activating the expression of the Cue_2-O association, the Cue_1-O association will not be expressed and hence, interference takes place. On the other hand, if the context of testing is different from the context in which the Cue_2-O association was trained, interference no longer takes place and recovery (in this case, renewal) of responding appropriate to the Cue_1-O association occurs.

Several recent revisions of the Rescorla and Wagner (1972) model could account for the results of Groups Cues in terms of acquisition rather than retrieval effects (Markman, 1989; Tassoni, 1995; Van Hamme & Wasserman, 1994; see also Dickinson & Burke, 1996, for an analogous revision of Wagner's, 1981 SOP model). However, like the original Rescorla-Wagner model, these revised theories are unable to explain available data on interference between outcomes, such as, for example, the effect of the passage of time on the spontaneous recovery of the response (Pavlov, 1927). Moreover, these revised theories do not offer an explanation for the available data on interference between elementally-trained cues. These theories would explain this effect by assuming that a stimulus that is expected to occur during Phase 2 (i.e., A), but does not occur, has a negative salience. Following Van Hamme and Wasserman (1994), this negative salience leads the absent stimulus, in the cases in which the outcome is presented, to lose some of the previously acquired associative strength. But the stimulus has to be expected by the subject in order to have a negative salience. Following Dickinson and Burke (1996; see also Wasserman & Berglan, 1998), a cue can only be expected if it has a within-compound association with the cue that is present or with the training context. Thus, these theories can only predict the effect of interference between elementally-trained cues when the context in which A and B are trained is the same. This could provide an explanation for the Groups Cues in the present experiments, but it cannot explain other available results on

interference between elementally-trained cues. For example, this cannot explain the effects of retention intervals and retrieval cues mentioned above (see Pineño et al., 2000), nor can this account provide an explanation for the finding that interference can occur between elementally-trained cues even when the two cues are trained in different contexts (but as long as the target cue is tested in the context in which the interfering cue is trained; see Matute and Pineño, 1998a, Experiment 3).

Therefore, negative salience models could account for the results of the present experiments, but their difficulties in explaining many other results on interference between elementally-trained cues suggest that this kind of interference (as well as interference between outcomes) is better accounted for as a retrieval, rather than as an acquisition effect.

The idea that interference is a retrieval rather than an acquisition effect is consistent with Bouton's (1993) explanation of interference between outcomes, as well as with Miller and Matzel's (1988) explanation of interference between compound-trained cues as retrieval effects. However, researchers in the predictive learning tradition have usually assumed that interference between cues was different from interference between outcomes, because cues were assumed to interfere with each other only if they were trained in compound and outcomes were assumed to interfere with each other only if they were trained elementally. The present experiments show that elementally-trained cues interfere with each other in a way that is similar to that in which elementally-trained outcomes interfere with each other. Moreover, several other experiments have shown that compound-trained outcomes can also interfere with each other in a way similar to that in which compound-trained cues interfere with each other (Esmorís-Arranz, Miller, & Matute, 1997; Miller & Matute, 1998; Rescorla, 1980).

Nevertheless, it is important to note that the present experiments were not aimed at demonstrating that a common mechanism underlies both types of interference effects, but that the similarities between the two effects warrant further attention.   Indeed, the possibility exists that the similar results that we have observed are being produced by different mechanisms. In addition, it should be noted that the effects of interference between outcomes (e.g., extinction, counterconditioning) are well established, and they are easily replicated in different species, and using different preparations, whereas interference between elementally-trained cues seems to be more parameter-dependent.   For example, whereas biological significance plays an important role on interference between cues (see above), effects of interference between outcomes, such as extinction, counterconditioning, or reversal learning, are not subject to this constraint. In addition, as shown elsewhere (Matute & Pineño, 1998b), elementally-trained cues can be resistant to interference under some conditions even when cues biologically-neutral cues are used, as in the case of human causal

judgment preparations.

The present experiments clearly show that interference between cues and between outcomes are more similar than is often assumed in the literature of predictive learning and that it is important to consider the possibility that they might be due to a common mechanism. If this were the case, the results of studies of interference in the predictive learning tradition and in the paired associate tradition could be integrated into a general framework of interference theories.

References

Abra, J. C. (1967). Time changes in the strength of forward and backward associations. *Journal of Verbal Learning and Verbal Behavior*, **6**, 640-645.

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

Arcediano, F., Ortega, N., & Matute, H. (1996). A behavioural preparation for the study of human Pavlovian conditioning. *Quarterly Journal of Experimental Psychology*, **49B**, 270-283.

Bäuml, K.-H. (1996). Revisiting an old issue: Retroactive interference as a function of the degree of original and interpolated learning. *Psychonomic Bulletin and Review*, **3**, 380-384.

Bäuml, K.-H. (1998). Strong items get suppressed, weak items do not: The role of item strength in output interference. *Psychonomic Bulletin and Review*, **5**, 459-463.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological Bulletin*, **114**, 80-99.

Bouton, M. E. (1994). Context, ambiguity, and classical conditioning. *Current Directions in Psychological Science*, **3**, 49-53.

Bouton, M. E. (1997). Signals for whether versus when an event will ocurr. In M. E. Bouton & M. S. Fanselow (Eds.), L*earning, motivation, and cognition. The functional behaviorism of Robert C. Bolles*. Washington: American Psychological Association.

Bouton, M. E., & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation*, **10**, 445-466.

Bouton, M. E., & Nelson, J. B. (1998). Mechanisms of feature-positive and feature-negative discrimination learning in an appetitive conditioning paradigm. In N. A. Schmajuk & P. C. Holland (Eds.): *Occasion setting. Associative learning and cognition in animals*. Washington: American Psychological Association.

Brooks, D. C. (2000). Recent and remote extinction cues reduce spontaneous recovery. *Quarterly Journal of Experimental Psychology*, **53B**, 25-58.

Brooks, D. C., & Bouton, M. E. (1993). A retrieval cue for extinction attenuates spontaneous recovery. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 77-89.

Brooks, D. C., Palmatier, M. I., Garcia, E. O., & Johnson, J. L. (1999). An extinction cue reduces spontaneous recovery of a conditioned taste aversion. *Animal Learning and Behavior*, **27**, 77-88.

Chandler, C. C. (1993). Accessing related events increases retroactive interference in a matching recognition test. *Journal of Experimental Psychology: Learning, Memory and Cognition*, **19**, 967-974.

Chandler, C. C., & Gargano, G. J. (1998). Retrieval processes that produce interference in modified forced-choice recognition tests. *Memory and Cognition*, **26**, 220-231.

Cheung, C. G., & Goulet, L. R. (1968). Retroactive inhibition of R-S associations in the A-B, B-C, C-B paradigms. *Journal of Experimental Psychology*, **76**, 327-328.

Denniston, J.C., Miller, R.R., & Matute, H. (1996). Biological significance as a determinant of cue competition. *Psychological Science*, **7**, 325-331.

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

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

Escobar, M., Matute, H., & Miller, R. R. (2001). Cues trained apart compete for behavioral control in rats: Convergence with the associative interference literature. *Journal of Experimental Psychology: General*, **130**, 97-115.

Esmorís-Arranz, F. J., Miller, R. R., & Matute, H. (1997). Blocking of antecedent and subsequent events. *Journal of Experimental Psychology: Animal Behavior Processes*, **23**, 145-156.

Johnston, W. A. (1968). Bidirectional interference in an A-B, C-B paradigm. *Journal of Verbal Learning and Verbal Behavior*, **7**, 305-311.

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

Keppel, G., Bonge, D., Strand, B. Z., & Parker, J. (1971). Direct and indirect interference in the recall of paired associates. *Journal of Experimental Psychology*, **88**, 414-422.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **99**, 22-44.

Markman, A. B. (1989). LMS rules and the inverse base-rate effect: Comment on Gluck and Bower (1988). *Journal of Experimental Psychology: General*, **118**, 417-421.

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.

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

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

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

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

Miller, R. R., & Matute, H. (1998). Competition between outcomes. *Psychological Science*, **9**, 146-149.

Miller, R. R., & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed.), *The psychology of learning and motivation, Vol. 22* (pp.51-92). San Diego, CA: Academic Press.

Oberling, P., Bristol, A. S., Matute, H., & Miller, R. R. (2000). Biological significance attenuates overshadowing, relative validity and degraded contingency effects. *Animal Learning and Behavior*, **28**, 172-186.

Ortega, N. & Matute, H. (2000). Interference between elementally trained stimuli can take place in one trial. *Learning and Motivation*, **31**, 323-344.

Paredes-Olay, M. C., & Rosas, J. M. (1999). Within-subjects extinction and renewal in predictive judgments. *Psicológica*, **20**, 195-210.

Pavlov, I. P. (1927). *Conditioned reflexes*. London: Clarendon Press.

Pineño, O., Ortega, N., & Matute, H. (2000). The relative activation of the associations modulates interference between elementally-trained cues. *Learning and Motivation*, **31**, 128-152.

Rescorla, R. A. (1980). *Pavlovian second-order conditioning*. Hillsdale, NJ: Erlbaum.

Rescorla, R. A. (1996a). Preservation of Pavlovian associations through extinction. *Quarterly Journal of Experimental Psychology*, **49B**, 245-258.

Rescorla, R. A. (1996b). Spontaneous recovery after training with multiple outcomes. *Animal Learning and Behavior*, **24**, 11-18.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations

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

Rizley, R.C., & Rescorla, R.A. (1972). Associations in higher order conditioning and sensory preconditioning. *Journal of Comparative and Physiological Psychology*, **81**, 1-11.

Schwartz, M. (1968). Effect of stimulus class on transfer and RI in the A-B, A-C paradigm. *Journal of Verbal Learning and Verbal Behavior*, **7**, 189-195.

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

Underwood, B. J. (1966). *Experimental psychology* (2nd edition). New York: Appleton-Century-Crofts.

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

Vila, N. J. (2000). Extinción e inhibición en juicios de causalidad. *Psicológica*, **21**, 257-273.

Wagner, A.R. (1981). SOP: A model of automatic memory processing in animal behavior. In N.E. Spear & R.R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5-47). Hillsdale, NJ: Erlbaum.

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

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

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

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

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