THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2000, *53B* (3), 193–207

Does the type of judgement required modulate cue competition?

Pedro L. Cobos, Antonio Caño, Francisco J. López, Juan L. Luque, and Julián Almaraz
Universidad de Málaga, Málaga, Spain

According to the comparator process hypothesis (Matute, Arcediano, & Miller, 1996), cue competition in the learning of between-events relationships arises if the judgement required involves a comparison between the probability of the outcome given the target cue and the probability of the outcome given the competing cue. Alternatively, other associative accounts (the Rescorla–Wagner model: Rescorla & Wagner, 1972) conceive cue competition as a learning deficit affecting the target cue–outcome association. Consequently, the comparator process hypothesis predicts that cue competition occurs in inference judgements but not in contiguity ones, for only the first type of judgement implicitly involves such a comparison. On the other hand, the Rescorla–Wagner model predicts cue competition in both inference and contiguity judgements, because it establishes no relevant role for the type of judgement in producing cue competition. In Experiments 1 and 2 we manipulated the relative validity of cues and the type of question (inference vs. contiguity) in a predictive learning task. In both experiments we found a cue competition effect, but no interaction between the relative validity of cues and the type of question, suggesting that the Rescorla–Wagner theory suffices to explain cue competition.

Learning to predict is an essential requirement for surviving in our environment. Forecasting that it will rain from the sight of clouds, predicting that we will get burned if we have too much exposure to the sun, foreseeing an impending danger from somebody else's scream, predicting a gas leak from the identification of a certain smell, and so forth, are some of the numerous and important tasks we can solve because of our ability to infer predictive relationships between antecedent events (the clouds, the sun, the scream, the smell) and consequent events (the rain, the burns, the impending danger, the gas leak) by means of our experience. The aim of the study of this type of learning is to find out how this knowledge about predictive relationships is acquired from our experience with the events.

Requests for reprints should be sent to Pedro L. Cobos, Departamento de Psicología Básica, Psicobiología y Metodología de las Ciencias del Comportamiento, Facultad de Psicología, Universidad de Málaga, Campus de Teatinos, 29071, Málaga, Spain. Email: p_cobos@uma.es

This work is part of a research project supported by Junta de Andalucia (HUM0105). We thank David R. Shanks for his valuable comments on previous drafts of this article. We would also like to thank the reviewers of the previous version of this work, and especially Helena Matute, for helping us to improve this article.

http://www.tandf.co.uk/journals/pp/02724995.html

A common experimental procedure used in the study of predictive learning involves the presentation of a series of cases (trials) in which participants are given information about the antecedent events (cues) and, later, about the consequent events that have taken place (outcomes). This procedure is intended to simulate the natural conditions in which learning through direct experience with the events usually occurs. Individuals' performance in this type of task can be evaluated in different ways—for example, through the predictions they make on a trial-by-trial basis concerning the occurrence or not of the outcome given a certain cue, or through participants'judgements about the relationship between cues and outcomes. These judgements may be required during the learning process itself or in a later test phase. In the present article we specifically focused on people's judgements in predictive learning situations.

One of the most important phenomena in the development of predictive learning theories and models is that of cue competition. This takes place when several cues occur together preceding a particular outcome. In this case, the detection of the relationship between a target cue and the outcome depends on the relationship between other cues and this outcome. Let us take Experiment 3 by Shanks and López (1996) as an illustration of cue competition. In a medical diagnostic task, participants learnt predictive relationships between different symptoms (cues) and diseases (outcomes). On each trial, people were required to predict what the right outcome would be for a certain configuration of cues. Table 1 shows the different predictive relationships that were used. In a low relative validity condition, participants could see that Symptoms A and B indicated Disease 1 ($AB \rightarrow O_1$), that B on its own indicated Disease 1 ($B \rightarrow O_1$), and that patients suffering from Symptom C did not develop any disease ($C \rightarrow$ no O). In a high relative validity condition, participants could see across a series of trials that patients suffering from Symptoms D and E developed Disease 2 ($DE \rightarrow O_2$), those suffering from Symptom E did not develop any disease ($E \rightarrow$ no O), and those suffering from Sympton F also developed Disease 2 ($F \rightarrow O_2$). As can be seen, the number of pairings between Symptom A and Disease 1 and that between Symptom D and Disease 2 is the same for both conditions. Despite this, people tended to give higher ratings for the D–O_2 relationship than for the A–O_1 relationship, which has been said to demonstrate the importance of relative validity for learning about predictive relationships. Note that in the high-relative-validity condition Symptom D was paired with another symptom, E, that did not predict Disease 2 on its own, whereas in the low-relative-validity condition Symptom A was paired with Symptom B, which predicted Disease 1 on its own. Thus, D is a relevant predictor of Disease 2, and A becomes a redundant predictor of Disease 1 in the sense that it does not convey any additional information. In other words, Cue D is a better relative predictor of O_2 than Cue A of O_1. Cue competition, which was first identified and investigated in animals using classical conditioning (Kamin, 1968; Wagner, Logan, Haberlandt, & Price, 1968), has been obtained in humans in a variety of experimental situations (e.g., Baker, Mercier, Vallée-Tourangeau, Frank, & Pan, 1993; Chapman & Robbins, 1990; Dickinson & Shanks, 1985; Gluck & Bower, 1988; Shanks, 1991a).

Traditionally, this phenomenon has been explained by associationist theories. Within this framework we may distinguish between different kinds of explanations of cue competition. On the one hand, there is a group of theories that considers that relative validity has an effect on the establishment of associations between cues and outcomes.

According to these theories, the asymptotic associative strength between a cue and an outcome reflects the relative validity of the cue as a predictor for the outcome. Rescorla and Wagner's theory (1972) (RW hereafter) may be regarded as the prototype of this type of theory. According to this theory, the magnitude and sign of the change in the associative strength, in a given trial, between a certain cue and a certain outcome is directly proportional to the discrepancy between what is expected from all the cues present and what actually happens. All in all, the learning mechanism tries to minimize overall prediction errors through the entire set of training trials. If this principle is applied to the low-relative-validity condition of Shanks and López's (1996) Experiment 3, it is easy to see that in order to minimize errors, the associative strength of B (the competing cue) should be increased and that of A (the target cue) should be lowered to zero. With an associative strength of sufficient magnitude for Cue B, O_1 could be expected on $B \rightarrow O_1$ trials, and thus Cue A would not need any associative strength for subjects to anticipate O_1 accurately on $AB \rightarrow O_1$ trials. Moreover, if the target cue had some positive associative strength with O_1, the outcome would be overexpected in $AB \rightarrow O_1$ trials. In the high-relative-validity condition, in order to minimize errors, the associative strength of D (the target cue) needs to be increased and that of E needs to be lowered to zero. Cues C and F are not relevant for the present analysis, as they always occur in isolation. Therefore, RW predicts a stronger association between D and O_2 than between A and O_1.

This theory was originally used to explain the results of classical conditioning studies with animals, but it has also been used to explain predictive learning in humans (Dickinson, Shanks, & Evenden, 1984). An important assumption of this type of theory is that conditioned responses are the result of some monotonically increasing function of the associative strengths between cues and outcomes. This assumption has also been extended to relate judgement measures with associative strength in human learning. Therefore, the cue competition observed in participants' ratings comes from the different associative strengths of the various cues about which participants are asked.

Cue competition has also been explained by another kind of theory, which assumes that this effect is not due to a deficit in learning but to the mechanism responsible for the translation of what has been learned into a response. A noteworthy example of this point of view is the Comparator Theory (Miller & Matzel, 1988; Miller & Schachtman, 1985). According to this theory, the learning process entails the acquisition of associations reflecting contiguity rather than any high-level computation of relationships between events (i.e., computations of causality or predictiveness). Cue competition is the result of a comparator mechanism that takes place when the target cue is presented on its own. Such a mechanism draws a comparison between the associative strength of the target cue and that of the competing cue. When the competing cue is followed by the outcome more often than is the target cue, the former is supposed to be more strongly associated with the outcome than the latter. Thus, the target cue would elicit a weak response as a result of the comparison because it is more weakly associated with the outcome than is the competing cue.

Though Comparator Theory was developed in the animal learning field, it inspired Matute, Arcediano, and Miller (1996) to explain cue competition effects obtained with judgement measures in human learning. According to them, associations between cues and outcomes reflect contiguity or, specifically, the probability of the outcome conditionalized to the presence of the cue. However, when people have to make causal or inferential judgements (predictive or diagnostic), they draw a comparison between the probability of the outcome given the presence of the cue $[P(O|C)]$ and the probability of the outcome given the absence of the cue $[/P(O|C)/)]$—that is, in the presence of the competing cue. Cue competition is observed as a consequence of this comparison. This comparator process is supposedly activated because causality and inference questions implicitly require this comparison. Returning to the cue competition design mentioned above, the comparator process would yield different results for the target cues in the low- and the high-relative-validity conditions. In the low-relative-validity condition the probability $P(O_1|A)$ would have to be contrasted with the probability $P(O_1|B)$, whereas in the high-relative-validity condition the probability $P(O_2|D)$ would have to be contrasted with the probability $P(O_2|E)$. Thus, Target Cue D should obtain higher ratings than Target Cue A because the former is being compared with a lower conditional probability than is the latter.

Interestingly, according to Matute et al. (1996) different results should be expected if conditional probability (contiguity) judgements were requested. Conditional probability questions do not implicitly require any comparison between two conditional probabilities. Instead, they ask only for the $P(O|C)$. Thus, cue competition should not be observed with conditional probability questions because the comparator process is not activated by such questions.

In summary, the comparator process hypothesis (CPH hereafter) predicts that cue competition should only occur in inference or causal judgements but not in conditional probability judgements. Alternatively, RW predicts cue competition in inference or causal judgements as well as in conditional probability judgements. Thus, if it could be shown that cue competition is observed in inference but not in conditional probability judgements, CPH would be favoured over RW because only the former could explain such a result.

Matute et al. (1996) claimed to have found evidence supporting CPH. In the experiments by Matute et al., participants had to make inferences as well as conditional probability judgements about a given Target Cue X. During the training phase, half of the participants were exposed to a high-relative-validity condition, whereas the other half were exposed to a low-relative-validity condition. Following the relative validity design of Wagner et al. (1968), trial types for the high-relative-validity and for the low-relative-validity conditions were, respectively, as follows: (1) $AX \rightarrow O_1$, $AX \rightarrow$ no O, $BX \rightarrow O_1$, $BX \rightarrow$ no O; (2) $AX \rightarrow O_1$, $BX \rightarrow$ no O. Note that Target Cue X is always compounded either with Cue A or with Cue B, and that the probability of the outcome given Cue X is the same in both conditions. However, Cues A and B perfectly predict the outcome in the low-relative-validity condition, whereas such cues do not predict the outcome better than X in the high-relative-validity condition. Participants were given a summary list in which all the events occurring in every trial were specified. Thus, participants learned in a described rather than in an experienced[1] situation. In the instructions, participants

[1] We have preferred the term "experienced' rather than "trial-by-trial" situation because the latter is not informative enough regarding the way in which the information is presented. Note that even contingency tables may be used in trial-by-trial preparations. Alternatively, the term "described situations" refers to those situations in which information about the presence or absence of cues and outcomes is given in summary lists or tables (Shanks, 1991b).

were told that they had to learn the relationships between certain substances and certain allergic reactions that people could develop in response to the substances. Matute et al. (1996) found an interaction between the type of judgement required (inference vs. conditional probability) and the relative validity (high vs. low) of the cues. Specifically, relative validity effects, cue competition, occurred with inference judgements, but not with conditional probability judgements.

Matute et al.'s results are relevant because RW-like accounts of cue competition have been largely accepted in human as well as in animal learning. Thus, it is interesting to know whether such results generalize to other commonly used procedures in human learning (e.g., Dickinson & Burke, 1996; Shanks & López, 1996). Specifically, the two experiments reported here examined if the findings by Matute et al. (1996) could also be obtained with the relative validity design and the experienced situation used by Shanks & López (1996) in their Experiment 3.

This procedure and this design are interesting for several reasons. First, there are grounds for expecting that associative learning mechanisms are more likely involved in experienced than in described situations (see Shanks, 1991b). Only the former allows participants to experience the temporal distribution of the relevant pairings between cues and outcomes. Additionally, experienced situations bear more resemblance to the temporal distribution of cues and outcomes in real-world situations. Secondly, Shanks and López showed that this procedure and design were sensitive to the relative validity manipulation under different judgement conditions (i.e., predictive and diagnostic inferences). Thus, it would be interesting to test whether this sensitivity will generalise to contiguity judgements. Moreover, contrasting the relative validity design used by Matute et al. (1996) and that used by Shanks and López, it is only for the latter that predictions from RW are independent of the learning rates adopted (see Shanks, 1991a).

Thus, Experiment 1 used the same within subjects relative validity manipulation as in Experiment 3 by Shanks and López (see Table 1). As explained above, a cue competition effect would involve higher D–O_2 than A–O_1 ratings. Additionally, the type of judgement required was manipulated. Thus, two different groups were compared: an inference group and a contiguity judgement group, in which participants were required to make inference and contiguity judgements, respectively. According to CPH, cue competition should only arise in the inference judgement group, whereas RW predicts similar cue competition effects in both judgement groups.

EXPERIMENT 1

In Experiment 1 we evaluated whether the type of question people have to answer affects the occurrence of cue competition in a predictive learning task in which participants are trained in an experienced learning situation.

Method

Participants

The participants were 57 psychology undergraduate students from the University of Málaga who volunteered to take part in the experiment as part of their course requirements.

Apparatus and stimuli

The experiment was conducted in 12 cubicles, each of which contained a single computer. During the training phase, stimuli were presented to participants on the computer monitor. The cubicles were separated by screens that prevented participants from seeing what was taking place in adjoining cubicles. For the test phase, a paper-and-pencil questionnaire was provided to register participants' ratings.

Experiments 1 and 2 were conducted in Spanish. During the training phase, the following sentence appeared at the top of the screen on each trial: "*In example number X, the following indicator lights have come on.*" Below this sentence, those indicator lights that had come on in a given trial were listed. In the centre of the screen, 1.5 s later, under the sentence: "*Which substance has leaked*", the four substances that could leak from the plant were listed vertically and in alphabetical order. Under these substances, the option "*no substance*" appeared, because, as it can be seen in the design, no substance leaked in some trials. At the bottom of the screen, the sentence: "*Press ENTER after making your prediction*" appeared. The participants had to give the answer required either by typing the initial letter of the selected substance or by typing letter "*n*" for the "*no substance*" option. Once the "enter" key had been pressed, the display on the screen was replaced with the sentence "*The correct answer is*" at the bottom of the screen, followed a couple of lines further down by the name of the substance that had leaked or, alternatively, *no substance*. If the answer was wrong, a short high-pitched beep was produced by the computer. A second later, the sentence: "*Press the space-bar to continue*" appeared at the bottom of the screen. Participants could take as much time as they considered necessary before resuming with the task. Once the space-bar had been pressed, the program immediately proceeded to the following trial.

Procedure

Participants read the instructions that explained that the experimental task consisted of two phases: (1) that they had to learn the predictive relationships arranged between the events; and (2) that they had to make inference or contiguity judgements on the relationships detected between such events (see the Appendix for the instructions used in Experiment 1). They were told to imagine that they were operators at a chemical plant trying to predict the leakage of several substances from a series of indicator lights in a control room. In order to achieve this, they had to learn the relationships arranged between the indicators and the substances. Across a series of trials, participants had to predict which substance was leaking on the basis of the illuminated indicator light. Once the prediction had been made on each trial, participants were given corrective feedback. The training trials were based on the design in Table 1, which shows half of the trial types that occurred during the experiment. The other half was identical, except that different indicators (*G* to *L*) and different substances (3 and 4) were used. The relationships between these new indicators and substances mimicked those described in Table 1. Overall, the task comprised 12 indicators (from 1 to 12), 4 substances that might leak from the chemical plant, and 12 different trial types. The substance names were pseudowords (*betacina*, *dioxina*, *fenolina*, and *licaina*). For half of the participants, Target Cues *A* and *D* (see Table 1) were assigned to Indicators 2 and 9, respectively (Cues *G* and *J* were assigned

TABLE 1
Design and results of Experiments 1 and 2

			Experiment 1		*Experiment 2*	
Relative validity	*Training phase*	*Test*	*Inference*	*Contiguity*	*Inference*	*Contiguity*
Low	$AB \rightarrow O_1$ $B \rightarrow O_1$ $C \rightarrow$ no O	$A \rightarrow O_1$	66.6	70	47.3	58.8
High	$DE \rightarrow O_2$ $E \rightarrow$ no O $F \rightarrow O_2$	$D \rightarrow O_2$	86.1	85	66.7	70

to Indicators 8 and 7, respectively). For the remaining participants, the assignments for Cues *A* and *D* as well as for cues *G* and *J* were reversed.

Orthogonally, for half of the participants, Outcomes 1 and 2 were assigned to the substances *betacina* and *dioxina* (Outcomes 3 and 4 were assigned to *fenolina* and *licaina*). For the other half, the assignment of substances to Outcomes 1 and 2 was interchanged, as was the assignment to Outcomes 3 and 4. Trials were randomly presented. The order could differ from one participant to another, provided that each trial type appeared 10 times. Thus, the task consisted of 120 trials.

Once the training phase had been completed, participants had to make either inference or contiguity judgements through a paper-and-pencil questionnaire. In the inference group, the question: "*To what extent does indicator light X predict the leak of the different substances?*" was asked for each cue. In the contiguity condition, the question was worded in the following way: "*When indicator light X comes on, to what extent would you say the substance leaks?*" (see the Appendix for the Spanish wording of each question). Under these sentences, there was a list of the four substances, in alphabetical order. On the right hand side of each substance there was a scale from 0 to 100, calibrated in units of ten. To make their judgements, participants had to mark with an *X* the box of the scale they thought was correct. For all participants, judgements on Target Cues *A* and *D* opened the questionnaire, whereas those on Target Cues *G* and *J* ended it. For half of the participants, Cue *A* appeared before Cue *D*, and Cue *G* appeared before Cue *J*. For the other half this order was simply reversed. The questionnaires included specific instructions to ensure an adequate understanding of the questions and a correct use of the scale (see the Appendix for such instructions).

Results and discussion

Before carrying out an analysis of the results, we selected participants on the basis of their performance during the training phase, for two basic reasons: First, according to RW, cue competition effects shoud be most evident at asymptote. Second, according to preliminary work with this procedure, those participants who do not reach a high percentage of correct answers during the last 24 training trials usually show poor transfer when answering the questionnaire. Hence, we decided to select those participants who had made 2 mistakes or fewer during the last 24 training trials. As a result of this, the selection included 24 participants altogether—9 in the inference condition and 15 in the contiguity condition.

In the experiments reported, all the analyses adopted an α of 0.05. The results obtained are presented in Table 1. A2 (relative validity: high vs. low) $\times$ 2 (type of judgement: inference vs. contiguity) analysis of variance was carried out, these factors being within-subjects and between-subjects, respectively. The main effect of the relative validity factor was reliable, $F(1, 22) = 6.84$, $MSE = 488$—that is, ratings to the high-relative-validity cues were higher than those to the low-relative-validity ones. However, neither the main effect of the type of judgement, $F < 1$, nor the Relative Validity $\times$ Type of Judgement interaction, $F < 1$, was reliable.

To sum up, in this experiment we obtained cue competition, but we did not find evidence that this effect is modulated by the type of judgement required from participants. These results are not consistent with the CPH explanation of cue competition. On the other hand, they seem to fit RW's explanation insofar as the cue competition effect was not statistically different in the inference and contiguity conditions.

The absence of an interaction might, however, lie in the participants' difficulty with discriminating between the two types of judgements. In this sense, a between-subjects manipulation of the type of judgement required might have not compelled participants to treat differentially the inference and contiguity questions. Alternatively, a within-subjects manipulation may induce participants to look for the distinctive elements of the questions that would serve to emphasize their difference. In view of this possibility, Experiment 2 was based on the design of Experiment 1, but it used a within-group rather than a between-group methodology. There were, in addition, a number of minor changes to the design that was used for Experiment 1.

EXPERIMENT 2

Method

Participants

Participants were 59 psychology undergraduate students from University of Málaga, who volunteered to take part in this experiment as part of their course requirements. However, due to a technical problem with data registration, the initial sample of 59 participants was reduced to 55.

Experimental task, apparatus and stimuli

In this experiment, the experimental task, apparatus, and stimuli used were as in Experiment 1.

Procedure

First, participants went through a pretraining task, which was intended both to improve participants' performance in the training phase and, more importantly, to ensure that participants appreciated the difference between the two types of judgements required in the test phase of the experiment. The pretraining task included a training as well as a test phase. In the training phase, participants were exposed to 15 trials in order to learn the relationships between three symptoms (cues) and two diseases (outcomes). The symptoms were "*swollen gums*", "*excessive perspiration*", and "*loss of vision*", and the diseases were given fictitious names: *hocitosis* and *beralgia*. The design was very simple, as it only included 3 different trial types: $AB \rightarrow O_1$, $B \rightarrow$ no O, $C \rightarrow O_2$. In the test phase all participants had to make both inference and contiguity judgements for each cue–outcome relationship regarding all possible outcomes. Once all participants had finished the pretraining task, they did not continue with the experiment task itself until the experimenter had resolved their problems, if any, during the pretraining task (e.g., the differentiation between both types of judgements or any other aspect related to this task).

After this pretraining task, participants read the instructions of the experimental task (see the Appendix for the specific instructions used in Experiment 2). The training phase differed from that used in Experiment 1 (see Table 1 for the design) in the following aspects: Modifications were introduced to improve participants' performance during the training phase. Instead of limiting the number of trials to 120, the training phase ended once participants had answered correctly on 24 consecutive trials. If this condition was not fulfilled, the task ended after 240 trials. At the same time, to guarantee that, as in the previous experiment, all of the trial types occurred with the same frequency, trials were presented in sets of 12. Each set contained each of the 12 possible trial types. Within each set, trials occurred randomly. Another modification in the procedure was the appearance of a coloured rectangle at the top right corner of the screen, which indicated their performance level to the participants. The extent to which this rectangle was filled depended on the percentage of correct responses during the last 24 trials.

During the test phase, as in Experiment 1, participants had to estimate each target cue–outcome relationship. However, in contrast to that experiment, every participant had to answer inference and contiguity questions. In order to avoid a too lengthy test phase, each type of question was assigned to just half of the target cues. For this, the questionnaire was divided in two halves: one for inference judgements and another for contiguity ones. In each part, participants had to estimate a high-relative-validity cue and a low-relative-validity cue. To control the effect of the order in which each type of question occurred, half of the participants answered contiguity questions first, and the other half answered inference questions first. The rest of procedural details were as in Experiment 1, including the specific instructions concerning the questionnaire (see the Appendix).

Results and discussion

As in Experiment 1, only the data of those participants who made fewer than three mistakes during the last 24 trials were included in the statistical analysis. The procedural modifications introduced in this experiment proved to be effective, as only two of the participants did not satisfy the selection criterion. On the other hand, all participants claimed that they understood both questionnaires after the pretraining task.

Table 1 shows the results obtained. A within-subjects 2 (relative validity: high vs. low) $\times$ 2 (type of judgement: inference vs. contiguity) analysis of variance was carried out. The relative validity main effect was significant, $F(1, 52) = 11.6$, $MSE = 1{,}064$. However, neither the main effect of type of judgement, $F(1, 52) = 2.57$, $MSE = 1{,}115$, nor the Relative Validity $\times$ Type of Judgement interaction, $F(1, 52) = 1.12$, $MSE = 816$, was significant.

The results of this experiment replicated the findings from Experiment 1—that is, the main effect of relative validity was obtained, but neither the type of judgement main effect nor the interaction was significant. In contrast to Experiment 1, the type of judgement was a within-subjects manipulation. As mentioned above, this was used in order to encourage participants to discriminate between inference and contiguity questions.

However, the within-subjects manipulation of the type of judgement proved to be insufficient to obtain a significant Relative Validity × Type of Judgement interaction.

Regarding the theories considered, the results obtained in Experiments 1 and 2 are difficult to explain by CPH, since, according to this theory, cue competition depends on the type of question that has to be answered. Specifically, CPH predicts that cue competition should be exclusively obtained in inference judgements because such judgements implicitly require a comparison between $P(O|C)$ and $P(O|\bar{C})$, whereas conditional probability judgements explicitly ask only for $P(O|C)$. Thus, CPH has difficulties with explaining the cue competition we obtained in Experiments 1 and 2 in the contiguity condition. Alternatively, the results pose no problem for RW, since this theory assigns no relevant role to the type of judgement in the production of cue competition.

GENERAL DISCUSSION

In Experiments 1 and 2 we tried to find evidence showing that the type of judgement required (inference vs. contiguity) modulates cue competition in order to assess the relative merits of the RW and CPH associative learning theories. In terms of RW, the type of judgement should not be relevant for determining whether or not cue competition effects will be found. For CPH, the occurrence of cue competition effects is predicted to depend critically upon the nature of the judgement about the relationships between cues and outcomes. Specifically, cue competition should be found if the judgement required implicitly involves a comparison between the conditional probabilities $P(O|C)$ and $P(O|\bar{C})$, as is the case for inference but not for contiguity judgements. From this, we should expect an interaction between the type of judgement required and the relative validity of cues, whereas RW predicts cue competition in both inference and contiguity judgements.

In Experiment 1, although the relative validity main effect was significant, we obtained no evidence of an interaction between the type of judgement and the relative validity factors. However, the results of that experiment might have been due to participants treating the inference and contiguity questions in the same way. Accordingly, in Experiment 2, the type of judgement received a within-subjects manipulation to facilitate participants' discrimination between inference and contiguity questions. Despite this change in the procedure, we were unable to obtain any interaction between the relative validity and the type of judgement required. Thus, the misinterpretation hypothesis as an explanation for the results of Experiment 1 appears unlikely.

Regarding the theoretical implications of the results reported, they are consistent with RW and not with CPH. However it is noteworthy that CPH makes two different assumptions. On the one hand, it assumes that there exists a comparator mechanism responsible for cue competition. On the other hand, it establishes the conditions under which such a mechanism operates (i.e., when inference but not when contiguity judgements are required). Thus, it may be said that our results are inconsistent with the second assumption, but not necessarily with the first one.

An important issue to be dealt with in future investigations has to do with the origin of the differences between our results and those of Matute et al. (1996). As we said in the introduction, Matute et al. found that cue competition occurred in inference but not in contiguity judgements. However, there are so many differences between their experiments and ours that it is not possible at the moment to point out which are the critical factors responsible for the different results obtained. Among others, there are four important differences: (1) There is a difference concerning the way in which the information was provided during the training phase. As described in the introduction, the events for each trial were presented simultaneously in a list in Matute et al.'s experiments whereas the events we presented one at a time in our experiements. Moreover, participants were forced to guess what the right outcome would be on each trial of our task, but not in Matute et al.'s task. Such differences could be important if it could be shown that these two modes of information presentation engage different cognitive processes, as suggested by Shanks (1991b). For example, it is reasonable to argue that Matute et al.'s learning task facilitates the use of calculation strategies to a greater extent than does ours, as all the information concerning the occurrence of the events is accurately and explicitly available in their procedure.

(2) We used a different relative validity design and a greater number of events and trial types than did Matute et al. (see introduction). In their experiments, half of the participants were exposed to two trial types and five events, whereas the other half were also exposed to four trial types and five events. This contrasts with the 12 trial types and 17 events used in our experiments. Therefore, it is not unreasonable to suggest that our participants faced a more difficult learning task in which working memory is quickly overloaded from the outset of training. Both features might have had an effect on the way, as well as the extent to which, the information from the different trials was integrated. Thus, differences in difficulty and memory load may have had an effect on what was learned in both experimental situations.

(3) Matute et al.'s learning task was about substances and allergic reactions rather than substance leaks and indicator lights in a chemical plant. Though we do not know why this factor would influence performance, content has proved to affect people's performance in many cognitive tasks, such as deductive and inductive reasoning tasks (Evans, Barston, & Pollard, 1983; Evans, Over, & Manktelow, 1993; Evans & Pollard, 1990; Gigerenzer & Hug, 1992; Griggs & Cox, 1982; Manktelow & Over, 1991).

(4) There are some differences concerning the questionnaire administered in the test phase. In the questionnaire used by Matute et al., inference and contiguity questions were presented next to one another for each target cue. In our Experiment 2, half of the cue–outcome relationships were tested using inference ratings in one block, and the other half were tested with contiguity ratings in another block. It could be argued that Matute et al.'s procedure makes it more likely that participants will discriminate between the two types of questions. However, participants in our Experiment 2 went through a pretraining phase in order to familiarize them with the two experimental phases and to improve their discrimination between inference and contiguity judgements. Note that in that phase participants had to make both types of judgements for every cue–outcome relationship. Additionally, participants received instructions on how to interpret the scales for the different questions. Thus, the difference in the way in which the questionnaire was administered may well be of minor importance. Finally, the actual wording of the questions posed in our experiment slightly differed from that used by Matute et al., though both sets of questions were unequivocally inference and contiguity questions.

In summary, there are many procedural variations between the experiments reported by Matute et al. (1996) and those reported here. These procedural differences may account for the different results obtained. However, despite such differences, it should be borne in mind that Matute et al.'s experiments and ours deal with the same theoretical relevant phenomenon, namely cue competition in the learning of predictive relationships, and therefore both results are relevant for a proper theoretical characterization of the phenomenon.

To conclude, our results question the degree of generalization of Matute et al.'s results. Their results showed that the type of judgement required is a crucial factor to obtain cue competition. Thus, these results may be viewed as challenging RW-like accounts of the learning of predictive relationships, as these models assume that cue competition mirrors the cue–outcome association learnt. On the contrary, such results support those models for which cue competition is produced at the moment of information retrieval, like CPH. Therefore, Matute et al.'s results are important because they question the merits of an account largely accepted in the field of predictive learning, though their generality is called into question. The results reported here, under very different procedural circumstances, clearly contrast with their results and favour RW-like accounts of cue competition.

Future investigation, then, should shed light on the importance of these procedural differences to account for the contrasting evidence reported here and by Matute et al. Interestingly, a complete account for the origin of the differences will surely involve a greater commitment to the cognitive processes underlying individuals' performance in predictive learning tasks. In any case, more work is clearly needed to elucidate the explanatory power of RW-like accounts and comparator theory-like accounts in this type of learning task.

REFERENCES

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

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

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

Dickinson, A., & Shanks, D.R. (1985). Animal conditioning and human causality judgement. In L.G. Nilsson & T. Archer (Eds.), *Perspectives on learning and memory* (pp. 167–191). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Dickinson, A., Shanks, D.R., & Evenden, J.L. (1984). Judgement of act–outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, *45B*, 241–258.

Evans, J.St.B.T., Barston, J.L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory and Cognition*, *11*, 295–306.

Evans, J.St.B.T., Over, D.E., & Manktelow, K.I. (1993). Reasoning, decision making and rationality. *Cognition*, *49*, 165–187.

Evans, J.St.B.T., & Pollard, P. (1990). Belief bias and problem complexity in deductive reasoning. In J.P. Caverni, J.M. Fabre, & M. Gonzales (Eds.), *Cognitive biases*. North-Holland: Amsterdam.

Gigerenzer, G., & Hug, K. (1992). Domain-specific reasoning: Social contracts, cheating, and perspective change. *Cognition*, *43*, 127–171.

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

Griggs, R.A., & Cox, J.R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, *73*, 407–420.
Kamin, L.J., (1968) "Attention-like" processes in classical conditioning. In M.R. Jones (Ed.), *Miami symposium on the prediction of behaviour: Aversive stimulation* (pp. 9–33). Miami: University of Miami Press.
Manktelow, K.I., & Over, D.E. (1991). Social roles and utilities in reasoning with deontic conditionals. *Cognition*, *43*, 183–188.
Matute, H., Arcediano, F., & Miller, R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning Memory and Cognition*, *22*, 182–196.
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.
Miller R.R., & Schachtman, T.R. (1985). Conditioning context as an associative baseline: Implications for response generation and the nature of conditioned inhibition. In R.R. Miller & N.E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 51–88). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
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.K. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
Shanks, D.R. (1991a). Categorisation by a connectionist network. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *17*, 433–443.
Shanks, D.R. (1991b). On similarities between causal judgements in experienced and described situations. *Psychological Science*, *2*, 341–350.
Shanks, D.R., & López, F.J. (1996). Causal order does not affect cue selection in human associative learning. *Memory and Cognition*, *24*, 511–522.
Wagner, A.R., Logan, F.A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, *76*, 171–180.

Manuscript received 14 October 1998
Accepted revision received 23 February 2000

APPENDIX

Instructions for Experiment 1

This experiment examines how difficult it is for people to learn causal relationships. Specifically, you have to learn the causal relationships between the illumination of certain indicator lights (effects) in a control room of a chemical plant and the leak of certain substances (causes) in such chemical plant.

You will be shown 120 examples. On each example, you will know which indicator or indicators (numbered from 1 through 12) have come on, and you should say which substance has leaked. The system is still under testing so illumination of an indicator light may occur without the leak of any substance.

Your aim is to learn which substance has leaked whent he various indicator lights come on. Try to make as many correct responses as possible. To achieve this, you can take all the time you need.

Once you have examined all the examples, you will have to evaluate through a questionnaire to what extent each of the different indicator lights predict the leak of the different substances [to what extent the different substances leak when each of the different indicator lights comes on]. Note that substances are indexed by different letters, so you do not need to remember their names.

Press the space bar to start . . .

[1] The text between square brackets represents the variations for the contiguity group.

Instructions for Experiment 2

Imagine that you work in a chemical plant. Sometimes the leak of various substances occurs due to accidents in the plant. These substances cause the illumination of different indicator lights that warn you about the leaks. However, you have not been told which substance has leaked when the varius indicator lights come on. It is important for the safety of the plant that you learn the causal relationships between the illumination of the different indicator lights (effects) and the leak of the substances (causes). The system is still under testing, so the illumination of an indicator light may occur without the leak of any substance.

Your aim is to learn which substance has leaked when the different indicator lights come on. For this, you will be shown a series of examples. On each example, you will know which indicator or indicators have come on, and you should say which substance has leaked. If you make a mistake, the computer will beep. Moreover, after making your response, you will be told what the right answer was. Use this information to learn the relationship between the different indicator lights and the substances.

Press the space bar to continue reading the instructions . . .

On the screen you will see your percentage of correct responses. Try to make as many correct responses as possible. The task ends when you get a certain number of correct examples.

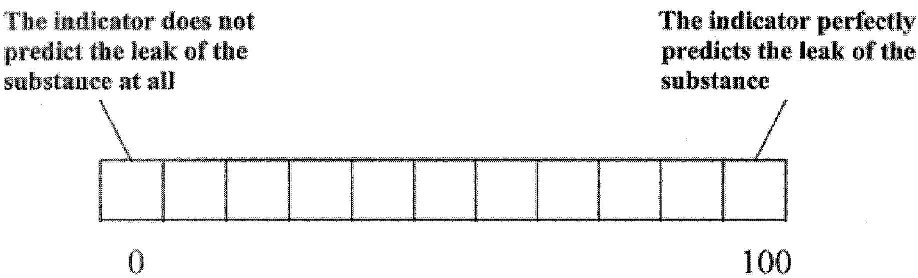
To ease the task, indicators are numbered from 1 to 12. Additionally, substances are indexed by different letters, so you do not need to remember their names.

Once you have examined all the examples, you will have to evaluate through a questionnaire to what extent each of the indicator lights predicts the leak of the different substances. Also, you will have to evaluate to what extent the different substances leak when each of the different indicators comes on.

Press the space bar to start . . .

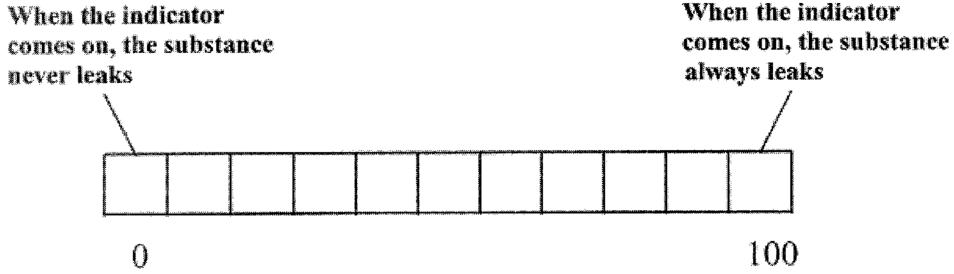
Instructions for the inference judgements

In what follows, you will find a questionnaire in which you will have to judge to what extent each of the different indicator lights predicts the leak of the different substances. For each question, you will have to mark, with an *X*, the following 0–100 rating scale, according to the meanings of the scale ends shown in Figure 1.



Instructions for the contiguity questionnaire

In what follows you will find a questionnaire in which you will have to judge to what extent the different substances leak when each of the indicator lights comes on. For each question, you will have to mark an *X* on the following 0–100 rating scale according to the meanings of the scale ends shown in Figure 2.



Spanish wording used for each type of judgement

a. Inference judgements: ¿En qué medida el indicador *X* predice el escape de las diferentes sustancias?
b. Contiguity judgements: Cuando se enciende el indicador *X*, ¿en qué medida afirmarías que se escapa la sustancia?