Judging the Contingency of a Constant Cue: Contrasting Predictions from an Associative and a Statistical Model

Frédéric Vallée-Tourangeau, **Robin A. Murphy**, and **A. G. Baker**

Department of Psychology
University of Hertfordshire
Hatfield, Hertfordshire
UNITED KINGDOM AL10 9AB
f.vallee-tourangeau@herts.ac.uk

Department of Psychology
McGill University
Montréal, Québec
CANADA H3A 1B1

Abstract

Two contingency judgment experiments are reported where one predictive cue was present on every trial of the task. This constant cue was paired with a second variable cue that was either positively correlated (Experiment 1) or negatively correlated with the outcome event (Experiment 2). Outcome base rate was independently varied in both experiments. Probabilistic contrasts could be calculated for the variable cue but not for the constant cue since the probability of the outcome occurring in the absence of the constant cue was undefined. Cheng & Holyoak's (1995) probabilistic contrast model therefore cannot uniquely specify the way in which the constant cue will be judged. In contrast, judgments of the constant cue were systematically influenced by the variable cue's contingency as well as by the outcome base rate. Specifically, judgments of the constant cue 1) were discounted when the variable cue was a positive predictor of the outcome but were enhanced when the variable cue was a negative predictor of the outcome, and 2) were proportional to the outcome base rate. These effects were anticipated by a connectionist network using the Rescorla-Wagner learning rule.

Introduction

In situations where little is known about the causal structure underlying the occurrence of a predictor event followed by the occurrence of an outcome event, their covariation serves as an important cue that informs a reasoner's judgment of their contingency. A measure of covariation is provided by the difference between the conditional probabilities of the occurrence of the outcome in the presence of the predictor, $p(O \mid P)$, and in its absence, $p(O \mid \text{no P})$, a measure termed ΔP.

Two classes of models, namely statistical and associative, have been developed to explain people's ability to judge inter-event contingency. The probabilistic contrast model (henceforth PCM) is the statistical model that has recently received the most attention (Cheng & Holyoak, 1995). In a judgment task that involves only one predictor event and one outcome, the PCM reduces to ΔP. That is, it postulates that reasoners explicitly consider the difference between the conditional probabilities $p(O \mid P)$ and $p(O \mid \text{no P})$. In situations where two predictors (A and B) signal a common outcome the PCM specifies that the derivation of the conditional probabilities for any given cue must itself be conditional on the presence or absence of another cue. Such a conditionalizing cue may be any cue that covaries with the outcome. Thus, if the outcome is judged to be contingent on B, B may be a conditionalizing cue for A. A's contingency with the outcome is then assessed by calculating a pair of contrasts, namely its contingency in the presence of B, or $\Delta P_{A|B}$ [i.e., $p(O \mid A \& B) - p(O \mid \text{no A} \& B)$] and its contingency in the absence of B or $\Delta P_{A|\text{no}B}$ [i.e., $p(O \mid A \& \text{no B}) - p(O \mid \text{no A} \& \text{no B})$]. If B is perfectly correlated with the outcome, i.e., the outcome always occurs in B's presence but never in its absence, each contrast for A will equal 0, namely $\{[p(O \mid A \& B) = 1] - [p(O \mid \text{no A} \& B) = 1] = 0\}$ and $\{[p(O \mid A \& \text{no B}) = 0] - [p(O \mid \text{no A} \& \text{no B}) = 0] = 0\}$. In this situation, the conditionalized probabilistic contrasts dictate that A should not be attributed causal importance; it is a redundant cause. And as many have reported (e.g., Baker, Mercier, Vallée-Tourangeau, Frank, & Pan, 1993; Price & Yates, 1993) in judgment tasks where A's contingency is moderately positive but B's contingency is perfectly positive, subjects discount the causal importance of A and rate its contingency near zero.

Associative models do not postulate that reasoners derive conditional probabilities and compute probabilistic contrasts in order to formulate a judgment of contingency. Rather they assume that a reasoner's contingency intuitions reflect the associative strength between a predictor and an outcome that develops on the basis of the *contiguity* between the two events. An associative model commonly discussed is the Rescorla and Wagner (1972; henceforth RW) model of learning which is a single layer localist connectionist network where the input nodes correspond to the predictor events and the output node corresponds to the outcome event. The weights between each predictor and the outcome reflect the strength of the hypothesized association. On any given learning trial, the weight connecting predictor j and the outcome is modified following a delta rule of the form,

$$\Delta w_j = \alpha_j \beta (\lambda - \Sigma w_k)$$

which is the weighted difference between the target activation value of the output node λ (which equals 1 when the outcome is present and 0 when it is absent) and the sum of the weights of the k predictors present on that trial (α_j and β are learning parameters coding for the associability of predictor j and the outcome respectively). This learning rule constrains the nature of the connection weights in two important ways: 1. The connection weight of a given predictor is influenced by the weights of the accompanying predictors and 2. their sum is bounded by λ since when $\Sigma w_k > \lambda$, $(\lambda - \Sigma w_k)$ is negative resulting in a negative adjustment of the weights. The predictions of the RW model are derived by training the network with event frequencies that correspond to the contingencies experienced by human

subjects in a given judgment task and comparing the magnitude, order, and polarity of the weights of each predictor with the magnitude, order, and polarity of the judgments of the contingency for these same predictors.

In many judgment tasks the predictions of the PCM and the RW model are identical (Baker, Murphy, & Vallée-Tourangeau, in press; Spellman, in press). For example, in the situation described above where A is a moderate predictor of the outcome but B is a perfect one, the RW model predicts discounting: with training the weight of the perfect predictor B approximates λ, and the moderately correlated cue A develops a connection weight that asymptotes at or near zero.

Judgment tasks that involve a constant cue offer an interesting forum to assess the merit of both models. When a cue is present on every trial of the task, and if the experimental trials make up the set of focal instances over which contrasts are calculated, then probabilistic contrasts cannot be calculated for that cue. There is no conditionalizing cue whose presence or absence identifies a focal set of trials where the conditional probability of the outcome occurring in the absence of a constant cue can be calculated. "Accordingly, subjects will have no positive evidence that any constant cue is causal" (Melz, Cheng, Holyoak, & Waldmann, 1993, p. 1404)[1]. Consequently, the PCM is unable to specify uniquely how people will judge the influence of a cue present on every trial of a judgment task. The power equations in Cheng, Park, Yarlas, and Holyoak (in press; e.g., Eq. 3) suffer the same fate since some of their terms are undefined. In turn, the RW model is able to formulate predictions about how people will judge relationships involving constant cues since the mechanism underlying the predictions is driven by the contiguity between the predictor and the outcome (and not their contingency) as well as by the magnitude of the weights of the accompanying predictors.

The judgment task designed for this study involved two predictor variables and one outcome variable. One of the two predictors (called X) was present on every trial whereas the second (A) was present on some trials and absent on others. In Experiment 1, predictor A was either positively correlated with the outcome, $p(O \mid A) > p(O \mid \text{no A})$, or was not correlated with the outcome, $p(O \mid A) = p(O \mid \text{no A})$. In Experiment 2, predictor A was either negatively correlated with the outcome, $p(O \mid A) < p(O \mid \text{no A})$ or was not correlated. The RW model predicts that judgments of predictor X, the constant cue, will be systematically influenced by the nature of A's contingency. Specifically, in Experiment 1 judgments of X should be lower when A's contingency is positive than when it is zero; in other words, X will have a weaker association with the outcome when A's contingency is positive. In Experiment 2 judgments of X should be greater when A's contingency is negative than when it is zero; that is, X will have a stronger association with the outcome when A's contingency is negative. These predictions hinge on the fact that the weight of predictor A is proportional to its contingency, and that the weight of the constant cue X is inversely proportional to the weight of A. When A's contingency is greater than zero, Σw_k will be larger than when A's contingency equals zero. Consequently the weight of X will be smaller when A's contingency is positive than when it is zero. In turn, when A's contingency is smaller than zero, Σw_k will be smaller than when A's contingency equals zero. Consequently the weight of X will be larger when A's contingency is negative than when it is zero. Thus while the RW model can formulate predictions about how judgments of X should be influenced by the presence of a variable predictor, the PCM is unable to formulate any prediction about the judgment of a constant cue since the probabilistic contrasts pertaining to X are undefined.

In both experiments, the outcome base rate, namely the proportion of trials where the outcome is present, was manipulated independently of the contingency of the variable predictor. Three different base rates were created: .25, .5, and .75. The RW model predicts that as the outcome base rate increases, judgment of the constant cue X should increase since X's contiguity with the outcome is directly proportional to the outcome base rate. Once again, the PCM is unable to advance predictions on the effect of outcome density on the judgments of a constant cue.

Method for Experiments 1 and 2

Task Scenario and Procedure

Subjects were asked to evaluate the relationships between each of two viruses with a certain disease in six samples of forty fictitious patients. Each sample showed new viruses and a new disease. For each sample the record of each patient was presented on a monitor one at a time informing the subjects of the presence or absence of the two viruses. Subjects were prompted for a diagnosis and then were told whether or not the disease was present. One of the two viruses (X) was present for all patients and the other (A) was sometimes present and sometimes absent. In each sample subjects were asked to rate the relationship between each virus and the disease, using a scale from -100 to 100, after 20 and 40 patients; the analyses reported below were conducted only on the terminal estimates. A virus could be negatively correlated with a disease since, as subject read in the instructions, "some viruses could afford immunity against a disease. The more negative the rating, the greater the immunity."

Design

Each sample of patients corresponded to one of six conditions derived from a 2 by 3 factorial design. The first independent variable was the contingency of the variable virus A and had two values, namely .5 and 0 in Experiment 1, and -.5 and 0 in Experiment 2. The second independent variable was the disease base rate in the sample which could take three values: .25, .5, and .75. The three conditions where A had a zero contingency were designed in both experiments: in the Low Density Zero condition $p(O \mid A) = p(O \mid \text{no A}) = .25$; in the Even Density Zero condition $p(O \mid$

[1] In fact, even if such a constant cue is part of a known physical mechanism involving the effect, it is understood to be an "enabling condition" and not a "cause" (Cheng & Novick, 1992).

A) = p(O| no A) = .5; and in the High Density Zero condition p(O | A) = p(O | no A) =.75.

In Experiment 1 the remaining three conditions were the three samples where virus A had a contingency of .5. In the Low Density .5 condition p(O | A) = .5 and p(O | no A) = .0; in the Even Density .5 condition p(O | A) = .75 and p(O | no A) = .25; and in the High Density .5 condition p(O | A) = 1 and p(O | no A) = .5. In Experiment 2, the remaining conditions were the negative image of these three conditions. Thus, in the Low Density -.5 condition p(O | A) = 0 and p(O | no A) = .5; in the Even Density -.5 condition p(O | A) = .25 and p(O | no A) = .75; and in the High Density -.5 condition p(O | A) = .5 and p(O | no A) = 1. The frequencies of the different kinds of trials in the six conditions of both experiments are shown in Table 1.

Experiment 1						
Trial Type	Low Density Zero	Even Density Zero	High Density Zero	Low Density Positive .5	Even Density Positive .5	High Density Positive .5
AX → O	5	10	15	10	15	20
AX → No O	15	10	5	10	5	0
X → O	5	10	15	0	5	10
X → No O	15	10	5	20	15	10
Experiment 2						
Trial Type	Low Density Zero	Even Density Zero	High Density Zero	Low Density Negative .5	Even Density Negative .5	High Density Negative .5
AX → O	5	10	15	0	5	10
AX → No O	15	10	5	20	15	10
X → O	5	10	15	10	15	20
X → No O	15	10	5	10	5	0
A = Variable Cue; X= Constant Cue; O = Outcome						

Table 1. Event frequencies in the six conditions of Experiments 1 and 2. Frequencies add up to 40 in each condition corresponding to the number of fictitious patients.

The order in which these conditions were presented to the subjects was randomized within each experiment. The labels assigned to the pairs of viruses and the six diseases were counterbalanced.

Subjects

Two different groups of 24 undergraduates from the University of Hertfordshire received course credits for their participation in Experiments 1 and 2. Data from one subject in each experiment were only partially recorded due to a computer malfunction. These subjects were not included in the analyses.

Experiment 1 Results

The mean terminal estimates of the variable predictor (A) and of the constant predictor (X) are plotted in the top left and top right quadrants of Figure 1. The effects of the two independent variables can be clearly observed in both panels. Starting with the judgments of the variable predictor, 1) judgments of A were greater when A's contingency was .5 than when it was zero, and 2) judgments in both contingency conditions were greater the higher the outcome base rate. A two-factor repeated measures analysis of variance (ANOVA) supported these observations (a .05 rejection criterion was used in all analyses). The main effect of contingency was reliable, $F(1, 22) = 12.6$, as was the main effect of outcome base rate, $F(2, 44) = 11.9$; the interaction was not reliable [$F < 1$].

The nature of the judgments of the constant predictor seemed clearly determined by the contingency of A as well as the outcome density. Thus, judgments of X were lower when A's contingency was .5 than when it was zero and judgments of X in all conditions were ordered as a function of the outcome base rate. A two factor repeated measures ANOVA confirmed these impressions. The main effect of A's contingency was reliable, $F(1, 22) = 11.9$, as was the main effect of base rate, $F(2, 44) = 39.2$; the interaction was not reliable [$F < 1$] .

Experiment 2 Results

The bottom two quadrants of Figure 1 show the mean terminal estimates for the variable predictor (left) and the constant predictor (right) in Experiment 2. Judgments of the variable predictor were again determined by the actual contingency and the outcome base rate. Judgments of A's contingency were more negative when the contingency was -.5 than when it was 0 and judgments in all conditions were greater the higher the base rate. A two-factor repeated measure ANOVA confirmed these impressions: the main effect of contingency was reliable, $F(1, 22) = 121$, as was the main effect of outcome base rate, $F(2, 44) = 13.4$; the interaction was not reliable [$F < 1$].

The judgments of the constant predictor in Experiment 2 were the mirror image of the judgments of the constant predictor in Experiment 1. That is, whereas A's positive contingency lowered the judgments of the constant cue in Experiment 1, A's negative contingency increased the judgments of the constant cue in Experiment 2. Again, in all conditions judgments were greater the higher the outcome

base rate. Statistical analyses again yielded reliable main effects of A's contingency, $F(1, 22) = 25.5$, and of base rate, $F(2, 44) = 64.1$; the interaction was not reliable [$F(2, 44) = 1.86$].

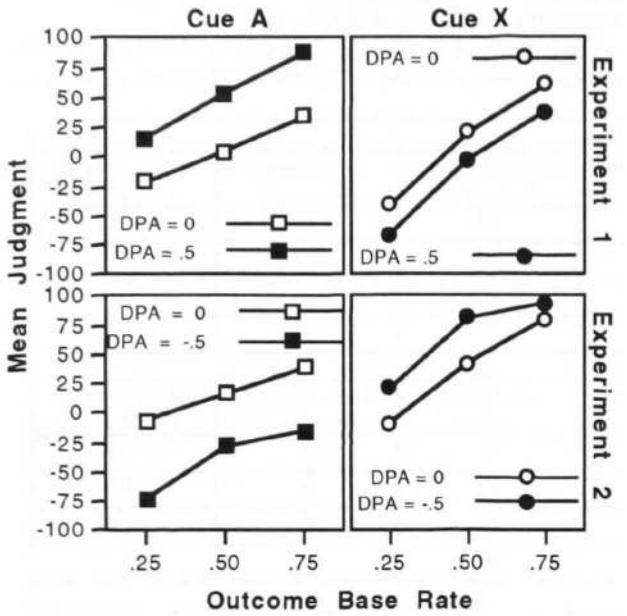

Figure 1. Mean contingency judgments for the variable cue A and the constant cue X in Experiment 1 (top left and top right) and in Experiment 2 (bottom left and bottom right). DPA = Delta P of cue A, or p(O | A) - p(O | no A).

Discussion

In two contingency judgment experiments, a predictor variable was present on every trial. In the first experiment, the constant predictor was paired with a second predictor that had either a positive or a zero correlation with the outcome, and in the second experiment it was paired with a predictor that had either a negative or a zero correlation with the outcome. In spite of the constant cue's identical probability of being paired with the outcome in each outcome base rate condition, judgments of the constant cue were systematically influenced by the nature of the correlation between the second predictor and the outcome. Specifically, judgments of the constant cue were attenuated when the second predictor was positively correlated with the outcome but were enhanced when the second predictor was negatively correlated with the outcome. This is a novel finding and establishes that judgments of constant cues can be discounted or potentiated in the same way as the judgments of variable cues (e.g., Baker et al., 1993). Subjects could have ignored the constant cue and rated its relationship with the outcome as zero, but they did not. Furthermore, in both experiments, judgments of the constant cue were determined by the base rate of the outcome: the more prevalent the outcome, the more positive the judgments of the constant cue. Thus, the participants in these experiments had no difficulty evaluating the nature of a relationship between a constant cue and an outcome even if probabilistic contrasts could not be computed. Importantly, the dual influences of the contingency of the accompanying cue and of the outcome base rate were anticipated by the RW model.

To account for the judgments of the constant cue the PCM must postulate a focal set of instances whose nature is not constrained by the trials of the task, thereby enabling the derivation of the probability of the outcome in the absence of the constant cue, $p(O \mid \text{no X})$. The difficulty, however, is that due to the nature of the judgment task and the fictitious diseases, it is hard to conceive the kinds of life experiences subjects may recruit to define the probability of any of the fictitious diseases in the absence of the constant cues. For example, to use one set of virus-disease labels employed here, what experiences outside the laboratory could subjects use to define the probability of the occurrence of the disease *Ork's Complex* in the absence of *Threbbagia*? Let's assume that for some subjects *Ork's Complex* reminded them of a real world disease (and that these subjects might have assumed also that in those real world cases *Threbbagia* was absent), the probability (*Ork's Complex* | No *Threbbagia*) is no longer undefined. The disease labels used in these experiments (*Ork's Complex*, *Nachmose A*, *Grympox*, *Melastraz*, *Trachtosis*, *Voldusis*) may have reminded different subjects to different degrees of real world diseases, thereby producing focal sets, for some diseases, which defined the probability of the disease in the absence of the constant cue. However, one might predict that subjects would produce highly variable judgments of the constant cue given their variable backgrounds; yet systematic patterns were observed.

More generally, it might be argued that people use an abstract reasoning schema which by default sets the probability of a disease in the absence of a virus to zero. This would correspond to the common understanding of the pathogenic quality of viruses. But subjects would have been ill-served by this reasoning schema since in these experiments some viruses could grant immunity and indeed in the High Density -.5 condition of Experiment 2, the probability of the disease in the absence of virus A equalled 1! Evidently subjects were aware of the different kinds of viruses in this judgment task since they experienced no difficulty rating some virus-disease relationships negatively. It is thus unlikely that they assumed by default that the probability of the disease in the absence of a virus was zero.

Baker, Murphy, and Vallée-Tourangeau (in press) have pointed out that the PCM's difficulties with undefined contrasts may be alleviated by including the inter-trial intervals (ITI) in the calculation of conditional probabilities such as $p(O \mid \text{no X})$. For example, the time separating the presentation of each patient's record could be segmented in discrete time intervals where nothing is happening, that is where none of the viruses are present and where the disease is absent as well. When such ITI segments are included in the calculations, the probability of the outcome in the absence of the constant cue is no longer undefined and equals zero. Following this strategy, the probabilistic contrasts for the constant cue in the six conditions of both experiments account partly for the judgment of the constant cue. This can be assessed in Table 2 where the ordering of the mean judgments of X parallels loosely the ordering of

$\Delta P_{X|no\ A}$. This auxiliary assumption, however, is not without problems. One may question of course the plausibility of arguing that subjects consciously considered the inter-trial interval when evaluating the constant cue's effectiveness. More importantly, the medical context in which the task is couched means that the probability of a disease can only be defined with respect to patients that either have it or not, and such patients were absent during the ITI.

While it can be argued that the RW model better accounts for the patterns of judgments of the constant cue, neither model fares well in explaining the strong effect of outcome base rate on judgments of the variable predictor (see the left half of Fig. 1). Probabilistic contrasts are impervious to differences in outcome densities if these densities do not affect the overall contingencies. Again, the PCM can resort to including the ITI in calculating A's contingencies (the conditionalizing cue for A is no longer the constant cue X): counting time segments where nothing happens increases frequency of "no A" observations. The greater the number of ITI observations included in the calculations of $p(O \mid no\ A)$, the smaller $p(O \mid no\ A)$, and the more proportional to $p(O \mid A)$ A's contingency becomes. In this way, the higher $p(O \mid A)$, the higher the judgments. And this is certainly what was observed (assuming that the same number of ITI observations were included in the devaluation of $p(O \mid no\ A)$ in the positive, negative and zero conditions). With this auxiliary assumption however, the PCM is committed to predict an overall positive bias in the estimates of A across both experiments, namely the .5 contingencies in Experiment 1 should be judged more positive than the -.5 contingencies of Experiment 2 should be judged negative. Such an asymmetry was observed: The overall judgment means of the variable cue in the three .5 conditions of Experiment 1 and the three -.5 conditions of Experiment 2 were 53.8 and -38.3 respectively. However, the asymmetry was not statistically significant: The absolute magnitude of the overall means did not differ reliably ($t(136) = 1.73$). Furthermore, symmetric judgments of positive and negative contingencies are routinely observed in similar tasks (e.g., Vallée-Tourangeau, Baker, & Mercier, 1994).

Associative models, in turn, usually predict some effect of outcome density on learning as this changes the contiguity between the predictor and the outcome as well as the associative strength of the context in which learning takes place. Specifically, the RW model predicts that for a positively correlated predictor, lower base rates yield larger positive connection weights, and for a negatively correlated predictor, lower base rates yield more negative connection weights (Wasserman , Elek, Chatlosh, & Baker, 1993) ; for a non-correlated predictor, the weights, at asymptote, should equal zero regardless of the base rates. Ostensibly, the ratings of the varying cue exhibited none of these predicted effects. However, judgments of the varying cue might have been influenced by the associative strength of the constant cue. In an animal conditioning preparation, learning supported by a conditioned stimulus may be better determined on test trials conducted in a test context that is different from the training context. Analogously, subtracting the constant cue ratings from the ratings of the varying cue would yield estimates of the varying cue "freed" of the influence of the constant cue. These adjusted ratings of the varying cue in the two contingency conditions of Experiment 1 and Experiment 2 are shown in Figure 2 (left and right panel respectively). The adjusted ratings of the varying cue in the positive contingency conditions of Experiment 1 show the predicted effect of outcome base rates: they are more positive with smaller base rates.

Experiment 1				
Conditions	$p(O \mid X \& no\ A)$	$p(O \mid no\ X \& no\ A)$	$\Delta P_{X\|no\ A}$	Mean Judgments
HD 0	0.75	0.00	0.75	60.2
HD .5	0.50	0.00	0.50	36.0
ED 0	0.50	0.00	0.50	21.1
ED .5	0.25	0.00	0.25	-3.3
LD 0	0.25	0.00	0.25	-39.7
LD .5	0.00	0.00	0.00	-64.8
Experiment 2				
Conditions	$p(O \mid X \& no\ A)$	$p(O \mid no\ X \& no\ A)$	$\Delta P_{X\|no\ A}$	Mean Judgments
HD -.5	1.00	0.00	1.00	92.4
ED -.5	0.75	0.00	0.75	79.2
HD 0	0.75	0.00	0.75	78.7
ED 0	0.50	0.00	0.50	40.7
LD -.5	0.50	0.00	0.50	19.9
LD 0	0.25	0.00	0.25	-10.4
HD = High Density; ED = Even Density; LD = Low Density				

Table 2. Comparisons of the predictions of the PCM given by including the inter-trial interval in the focal set for the constant cue X and the ordering of the mean judgments for X (Experiment 1, top half; Experiment 2 bottom half).

However, this pattern holds for the two zero contingency conditions as well as for the negative contingency condition of Experiment 2. The effects in the noncontingent conditions are in fact preasymptotic predictions of the model, but not in the negative contingency condition. Thus, while the RW model can formulate predictions about the influence of base rates on the judgments of the varying cue, and that these predictions have often been confirmed (e.g., Wasserman et al., 1993, Fig. 5), they were only partially observed in these experiments. Base rate effects on contingency judgments have important implications and future research should aim to elucidate the conditions under which they are and are not observed.

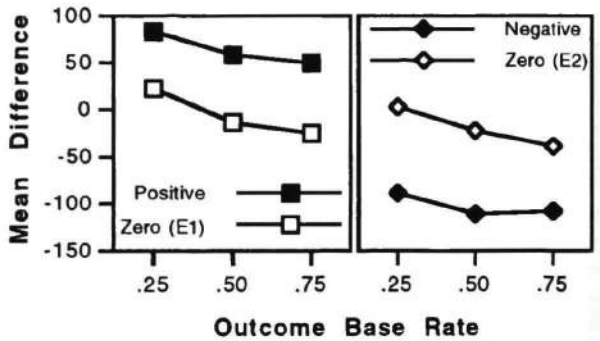

Figure 2. Mean difference between ratings of the varying cue A and the constant cue X for the two contingency conditions in Experiment 1 (left panel) and in Experiment 2 (right panel) as a function of outcome base rate.

In summary, the two experiments reported in this paper showed that discounting and enhancing effects found with variable cues can also occur with constant cues. These experiments raised an important concern about the PCM, namely whether any conceptually acceptable set of focal instances could be derived a priori for constant cues that could predict the dual effect of outcome base rate and the variable cue's contingency on the estimates of the constant cues. Judgments of the constant cue in both experiments were better explained by a mechanism operating on the basis of the contiguity between the constant cue and the outcome, a mechanism that is also constrained by the contiguity of other cues present.

Acknowledgments

This work was supported by grant R000221514 from the Economic & Social Research Council (UK) awarded to FVT. We wish to thank Susan Drew for recruiting and running the participants, and one anonymous reviewer for thoughtful comments.

References

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

Baker, A. G., Murphy, R. A., & Vallée-Tourangeau, F. (In press). Associative and normative models of causal induction: Reacting to versus understanding cause. To appear in D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The Psychology of Learning and Motivation*, Vol. 34. San Diego: Academic Press.

Cheng, P. W., & Holyoak, K. J. (1995). Complex adaptive systems as intuitive statisticians: Causality, contingency, and Prediction. In J.-A. Meyer, & H. Roitblat (Eds.), *Comparative approaches to cognitive science* (pp. 271-302). Cambridge, MA: MIT Press.

Cheng, P. W., & Novick, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, 99, 365-382.

Cheng, P. W., Park, J., Yarlas, A. S., & Holyoak, K. J. (In press). A causal-power theory of focal sets. To appear in D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The Psychology of Learning and Motivation*, Vol. 34. San Diego: Academic Press.

Melz, E. R., Cheng, P. W., Holyoak, K. J., & Waldmann, M. R. (1993). Cue competition in human categorization: Contingency or the Rescorla-Wagner learning rule? Comment on Shanks (1991). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1398-1410.

Price, P. C., & Yates, F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory & Cognition*, 21, 561-572.

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

Spellman, B. A. (In press). Conditional contingency as an explanation of cue-interaction effects. To appear in D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The Psychology of Learning and Motivation*, Vol. 34. San Diego: Academic Press.

Vallée-Tourangeau, F., Baker, A. G., & Mercier, P. (1994). Discounting in causality and covariation judgments. *Quarterly Journal of Experimental Psychology*, 47B, 151-171.

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