

The effectiveness of inhibitors in human predictive judgments depends on the strength of the positive predictor

DANIELLE M. KARAZINOV and ROBERT A. BOAKES
University of Sydney, Sydney, New South Wales, Australia

We tested whether the development of inhibitory strength, as measured by a summation test, is proportional to the strength of the positive cue (P) against which the inhibitory cue (I) is trained. P predicted the outcome, whereas the co-occurrence of P with I (PI) predicted no outcome. In Experiments 1, 2, and 3, we compared the latter design against a version in which P was overshadowed by another cue (X). In this design, the compound P^oX predicted the outcome, but P^oI^o predicted no outcome. In all three experiments, overshadowed cue I^o was less inhibitory than I. In Experiment 4, a P produced by fewer training trials also supported weaker inhibitory learning. Overall, the results were consistent with associative learning theories, especially Pearce's (1994) configural model. Contingency models need to make additional assumptions to accommodate this property of inhibitory learning.

Within the context of learning theory, an inhibitor generally is assumed to be a stimulus that signals the nonoccurrence of an event that would otherwise occur. In coming to learn that a stimulus is an inhibitor, an organism needs to expect that the event will occur before it is able to attribute its absence to the inhibitor. This situation is easily arranged in a predictive judgment experiment. In experiments of the kind reported here, participants are given a series of training trials in which certain cues, such as different foods, predict the presence or absence of a certain outcome, such as illness in a hypothetical person. They are then given a test in which they are asked to rate the likelihood of the outcome when particular cues are presented. What is referred to here as the *Pavlovian conditioned inhibition design* includes two types of training trial: one positive cue (P) is always followed by an outcome (P+ trials), but when P occurs with an inhibitory cue (I) the outcome does not follow (PI- trials). Typically, this training leads to I's being judged as having inhibitory properties (see, e.g., Chapman, 1991; Chapman & Robbins, 1990; Le Pelley & McLaren, 2001; Williams, 1996). Furthermore, the addition of I-alone trials followed by no outcome (I- trials) can enhance the inhibitory properties of the I element (Williams, 1995).

Notice that in the design described above the presence of P provides the basis for expecting the outcome when the PI compound occurs. In experiments using animals, pro-

cedures that do not involve P's occurrence with I can give rise to inhibitory learning also (see LoLordo & Fairless, 1985, for a review). These include explicitly unpaired (e.g., +, I-) and differential (P+, I-) training. The effectiveness of these procedures relies on the experimental context's giving rise to an expectation of the outcome during presentation of I (see, e.g., Miller, Hallam, Hong, & Dufore, 1991). Recently, we tested whether these procedures can produce inhibition in human predictive judgments, but we did not find evidence for such effects (Karazinov & Boakes, 2004). This may be because the experimental context is less well defined in human predictive judgment experiments and may not support the same level or type of learning that it does in animal experiments. Thus, we regard the explicit presence of P to be the most effective, and possibly the only, method of providing an expectation of the outcome during inhibitory learning in this experimental paradigm.

Like all other theories of inhibitory learning, recent contingency-based models (Cheng, 1997; Cheng & Novick, 1992; Cheng, Park, Yarlus, & Holyoak, 1996) assume that a positive cue must be present for inhibitory learning to take place. These models are developments of the original normative contingency model (Ward & Jenkins, 1965) with its central claim that participants base their judgments on the contingency between a particular cue and an outcome. Participants are assumed to estimate the probability of the outcome's occurring with the cue minus the probability of its occurring without the cue. The amount of inhibition that develops is determined by the extent to which there is a negative contingency between the cue and the outcome. In the probabilistic contrast model (Cheng & Novick, 1992), this account is modified to claim that humans estimate contingencies between target cues and outcomes only over a selected set of events

We thank Evan Livesey for programming the computer tasks for the experiments discussed in this article. This research was supported by an Australian Postgraduate Award and a Postgraduate Research Grant from the School of Psychology, University of Sydney, to D.M.K. and a grant from the Australian Research Council to R.A.B. Correspondence concerning this article should be addressed to D. M. Karazinov, School of Psychology, University of Sydney, Sydney, NSW 2006, Australia (e-mail: dannyk@psych.usyd.edu.au).

normally referred to as the *focal set*. Inhibitors are judged according to the negative contingency between the cue and the outcome over trials in which some positive cause is present. If, however, participants cannot identify an appropriate focal set, then judgment will be withheld.

According to this model, participants given training on a Pavlovian conditioned inhibition procedure (P+ vs. PI-) first determine whether there is a positive causal relation between P and the outcome and then learn about the inhibitory properties of I from its negative contingency with the outcome (-1) over trials in which P is present. In other words, the focal set for I is restricted to PI- and P+ trials. The important point here is that, although the probabilistic contrast model assumes that the *presence* of P is a necessary criterion for inhibition, unlike alternative accounts it does not assume that the strength of an inhibitor, I, depends on the *strength* of P, with which it has been paired. Rather, the strength of I, it claims, depends entirely on the value of the negative contingency between I and the outcome across the focal set defined by P.

This is quite different from the analysis of inhibition provided by associative learning theories (e.g., Rescorla & Wagner, 1972; also see Dickinson, 2001). Elemental versions of such theories propose that during training with the Pavlovian conditioned inhibition design inhibitory value accrues to I on PI- trials in order to reduce the discrepancy between expectancy (or V value) and outcome. Consequently, the greater the expectancy encouraged by the presence of P, the stronger the inhibition needed to counteract it. Configural versions of associative learning theories treat inhibitory learning in a different way. Since these assume that compound cues, such as PI, function as single units in acquiring predictive strength, generalization rules are needed to account for inhibitory learning. Assuming a particular set of rules, Pearce (1994) proposed that, once the configural unit for P has gained positive strength, this will generalize to the configural unit for PI. Thus, on early PI- trials there is a discrepancy between expectation (positive) and outcome (absence) that leads PI to acquire increasing inhibitory strength to counteract generalization of positive strength from P. The configural unit for PI becomes an effective inhibitor only because of generalization from P. The point here is that, just as with elemental associative theories, the effectiveness of PI as an inhibitor depends on the strength of P.

The different assumptions made by contingency and associative learning theories about the variable that drives inhibitory learning allow a test of the class of theory that best accounts for inhibitory learning in human predictive judgments. Accordingly, the main aim of the present study was to test whether inhibitory learning depends on the strength of P. In addition, some aspects of the design made it possible to assess whether elemental or configural versions of associative theory provided a better account of what our participants learned.

The claim that inhibition is proportional to the positive predictor was first made by Pavlov (1927): “[T]he

intensity of the primary inhibition is found to be exactly proportional to the intensity of the excitatory process on which it is based” (p. 100). He did not report the evidence on which his claim was based, and little relevant evidence has since been reported from either animal or human experiments. One animal conditioning experiment using the Pavlovian conditioned inhibition procedure has provided direct support for the claim (Wagner & Saavedra, briefly reported in Wagner, 1971), whereas a handful of animal experiments using the explicitly unpaired procedure have provided some indirect evidence that the positive strength of the experimental context determined the amount of inhibitory learning (Hearst, Bottjer, & Walker, 1980; Hearst & Franklin, 1977; Rescorla, 1969; Witcher & Ayres, 1980).

Although no studies using human participants have contrasted P with a low positive strength against another with a high positive strength during inhibitory training, in one study (Williams & Docking, 1995) the value of P was reduced *after* the PI- trials. In their Experiments 2 and 3, Williams and Docking found no reduction in inhibition after the positive value of P was reduced. However, in their fourth and final experiment, they found that a large number of P- trials relative to P+ trials did produce a significant reduction in inhibitory strength on I. This conclusion was based on a comparison with the results of a control group that received a higher overall number of P+ trials relative to the experimental group. Thus, this finding may have been due to further inhibitory learning in the control group rather than to a reduction in inhibition in the experimental group. Therefore, this series of experiments suggested that after inhibitory learning has taken place, it is relatively immune to a change in P's strength. Nonetheless, whether the amount of positive strength on P determines the amount of inhibition *during* learning remains an open question.

In the experiments reported here, we used a computer-based predictive learning task in which a hypothetical person ate different food items (cues) and suffered a migraine headache (outcome) after some of these foods but not others. Thus, participants could learn to predict the occurrence of the migraine outcome according to which food was consumed. In Experiments 1, 2, and 3, we used an overshadowing procedure to produce a cue with a smaller positive value. In Experiment 4, a positive cue that had been trained with a small number of trials was compared with one trained with a large number of trials. It was predicted that both manipulations would support less inhibitory learning than is supported by a cue with full positive value.

EXPERIMENT 1

In order to test whether overshadowing a positive predictor reduces inhibitory learning, we used a within-subjects comparison of two potential inhibitors. One inhibitor was trained with a cue that was always followed by the outcome alone. The other inhibitor was trained with a cue that, when followed by the outcome, always

Table 1
Design and Test Ratings for Experiment 1

	Test Cues	Ratings
Training cues: P+ (8), PI- (8), I- (8), P ^o X+ (8), P ^o I ^o - (8), I ^o - (8), FG+ (8), T+ (8), N- (16)	TI TI ^o TN TC P P ^o T	42.6 50.3 52.4 62.6 89.6 28.0 91.6

Note—P, positive cue; I, inhibitory cue; P^o, overshadowed positive cue; X, extra cue used for overshadowing; I^o, overshadowed inhibitory cue; F and G, filler cues; T, positive transfer cue; N, negative contingency control cue; C, novel control cue. Food names were allocated randomly to cue function for each participant. Presentation of the “migraine” outcome is denoted by “+,” and presentation of the “no migraine” outcome is denoted by “-.” The total number of presentations of each cue is given in parentheses. Ratings on an ungraduated analog scale were converted to ratings on a scale of 0 (“Migraine certain not to occur”) to 100 (“Migraine certain to occur”). The right-hand column shows the mean ratings over three trials for each type of test trial.

occurred in compound with another cue. Table 1 shows the full design for Experiment 1. We refer to the set of cues with a single positive predictor, P+, PI-, and I-, as the *standard inhibition design*, and to the set of cues with a compound positive predictor, P^oX+, P^oI^o-, and I^o- (where P^o and I^o are overshadowed positive and inhibitory cues, respectively), as the *overshadowed inhibition design*. A filler compound cue, FG+, was added as a further instance of a positive compound cue.

The most appropriate single test of inhibitory effects is a summation test (Rescorla, 1969; Williams & Docking, 1995). This requires training another positive cue that is always followed by the outcome (e.g., T+ trials). During test, I and I^o were combined with T. To ensure that any reduction in ratings produced by combining a cue with T was not due merely to the simple no-outcome training received by the cue, another no-outcome trained cue was needed as a control for inhibitory learning. Throughout the following experiments, we used the negative control cue N, which occurred with the same frequency as the potential inhibitory cues and was always followed by no outcome. Because I and I^o occurred both alone and in compound, N received double training trials. Thus, inhibitory learning about I would be revealed if TI is considered less likely than TN to be followed by the outcome. Furthermore, if I^o acquired less inhibition than I, TI would be considered less likely than TI^o to be followed by the outcome. For purposes of comparison with previous summation tests, we tested a novel cue, C, as a summation test compound cue, TC. In addition, we tested single cues P, P^o, and T to compare their positive strengths. (See Table 1.)

Method

Participants. Thirty-three first-year psychology students from the University of Sydney received course credit for taking part in the experiment.

Procedure. One to 12 participants at a time were tested in a single session of about 30 min. When the participants arrived, they were told that the experiment involved a computer-based task concerning a person who suffers from migraine headaches after eating certain foods. They then read the following instructions on the screen at their own pace:

We wish you to imagine that you are a doctor and that you want to discover what is causing a particular patient to have migraine attacks. You believe that some of the foods the patient is eating may be related to the occurrence of migraine attacks. In order to try and find out which foods are influencing migraines, you instruct the patient to consume particular foods on specific days and to record whether a migraine occurs or not. Your task is to look at these results and then try to determine whether any foods alter the likelihood of a migraine occurring. When you begin the simulation, you will be shown which foods have been consumed on a given day. A slider bar will be presented on the screen. You must make a prediction about whether you expect a migraine to occur by moving the slider bar. When you have made your prediction you will be shown the information indicating whether a migraine occurred or not. Numerous days will be presented and the foods consumed will sometimes be repeated, so that your ability to predict whether a migraine will occur should improve. Please use the slider scale carefully; only use the extreme ends of the scale if you are absolutely certain that a migraine will or will not occur.

After the experimenter had checked that the participants understood the task they were to perform, they were told to begin the training phase. On each training trial, the name of one (for single cues) or two (for compound cues) foods appeared in the center of the screen, and below this a horizontal scale appeared with a pointer set at the midpoint, which was marked “Don’t know.” The left end of the scale was marked “Migraine certain not to occur,” and the right end, “Migraine certain to occur.” The participant was then able to move the pointer to a location that represented how likely he or she expected the migraine to be. Immediately after the participant made a rating, a button labeled “Done” appeared on the bottom of screen beneath the scale bar. The cue or cues and the scale bar stayed on the screen until the “Done” button was pressed, and immediately afterward the outcome of that trial was revealed. This was presented as either the message, “A migraine DID occur” or the message “A migraine DID NOT occur.” The outcome for that trial remained alone on the screen for 2 sec before a “Display next” button appeared at the bottom of the screen. The next trial began immediately after the “Display next” button was pressed.

In each of the four 20-trial training blocks, each type of cue/compound cue listed in Table 1 occurred on two trials, except for N, which occurred on four trials. The trial sequence in each block was randomized for each participant. In the case of compound trials, on one trial within a block the first cue was presented to the left of the second (e.g., “bananas and melon”), and on the other the word order was reversed (e.g., “melon and bananas”). For each individual participant, 10 food names from the following list of 12 were allocated randomly by the program to each individual cue used in the overall design (i.e., P, I, P^o, X, I^o, T, N, F, and G, plus C used only in the test): *bananas, melon, potatoes, rice, beef, chicken, carrots, tomatoes, lettuce, biscuits, grapes, and apples*.

After the participants had completed the 80th training trial, the following message appeared on the screen:

Now that you have observed the pattern of migraines following different foods, it may be possible for you to anticipate which food or foods influence the occurrence of headaches. In the next phase of the simulation, you will be asked to rate the likelihood of a migraine when different foods are consumed. Please make this judgment on the same slider bar as before, but this time you will not be given feedback as to the accuracy of each prediction.

After the participants had read these instructions, they began the test phase. This consisted of three randomized blocks, each containing one of each of the seven types of test trials shown in Table 1.

Instead of reversing the left–right sequence of foods on compound trials (e.g., “potatoes and carrots” vs. “carrots and potatoes”) within participants as in training, the participants were divided into two groups for the testing phase so as to counterbalance cue sequence across participants. Thus, half of the participants were given the compound sequences TI, TI^o, TN, and TC (where C was novel), and the remaining half were given the reverse sequences—that is, IT, I^oT, NT, and CT. The only other change in procedure and in the appearance of the screen from training to test trials was that no feedback was given on any test trial.

Data and Analysis

In all the experiments, the position at which the pointer had been placed on each trial was converted by the program to a value on a scale ranging from 0 (“Migraine certain not to occur”) to +100 (“Migraine certain to occur”). The mean score over the three test trials for each cue was calculated for the analysis of test data. The test compound cue order (e.g., TI vs. IT) was treated as a separate groups factor in a preliminary analysis. No interaction involving this factor was found in any of the experiments, and the only main effect for this factor was found in Experiment 2; therefore, effects of compound cue order are reported only for that experiment. The main test analysis consisted of planned comparisons that excluded the groups factor of compound cue order during the test. A statistical significance level of .05 was used throughout all analyses.

Results

Performance during the training phase was analyzed in terms of three kinds of discrimination: The standard inhibition discrimination (P+ vs. PI–), the overshadowed inhibition discrimination (P^oX+ vs. P^oI^o–), and the simple discrimination (T+ vs. N–). Because N received twice as many training trials, only the first half of the N trials were used to calculate the simple discrimination scores reported here. (The same pattern of results was found when the analysis was conducted on the entire set of N trials rather than on only the first half of the N trials, as in the remaining experiments). For each participant, a discrimination score was obtained from each pair of cues (T+ vs. N–, P+ vs. PI–, and P^oX+ vs. P^oI^o–) by calculating the difference between the mean ratings of the two cues across the eight training trials in which the cues were presented. A larger discrimination score (maximum 100) over all blocks should indicate faster learning.

These scores indicated that both the standard and overshadowed inhibition discriminations were learned more slowly than the simple discrimination. The standard inhibition discrimination score (41.8) was significantly lower than the simple discrimination score [60.3; $F(1,32) = 11.66$]. The overshadowed inhibition discrimination score (46.0) also was significantly lower than the simple discrimination score [$F(1,32) = 9.41$]. The standard and overshadowed inhibition discrimination scores did not differ significantly [$F(1,32) < 1$].

In addition, we assessed the level of discrimination at the end of training by analyzing discrimination scores based on the last trial for each cue. For the simple, the standard, and the overshadowed discriminations, the scores were 93.7, 65.9, and 76.1, respectively. As with the scores based on all training trials, there were signif-

icant differences between the simple and standard inhibition discriminations [$F(1,32) = 10.91$] and between the simple and overshadowed inhibition discriminations [$F(1,32) = 5.51$], but not between the standard and overshadowed inhibition discriminations [$F(1,32) = 1.84$].

The most important results were the test ratings, shown in Table 1. Cue I trained in the standard inhibition design was found to have an inhibitory effect on T, in that TI was rated significantly lower than TN in the summation test [$F(1,32) = 5.78$]. The critical finding was that the standard inhibitor I was more inhibitory than the overshadowed inhibitor I^o, in that TI was rated significantly lower than TI^o [$F(1,32) = 6.01$]. However, the overshadowed inhibitor I^o was not reliably inhibitory, since TI^o was not rated significantly lower than TN [$F(1,32) < 1$]. Finally, P^o was compared to P to assess whether the presence of X in the P^oX compound produced an overshadowing effect on P^o. This comparison found that P^o was rated significantly lower than P [$F(1,32) = 55.75$].

Discussion

These results show that, whereas standard inhibition training produced a cue with inhibitory strength (I), training with an overshadowed positive predictor did not produce the same effect on a similar cue (I^o). Since the analysis of the training data failed to show a difference in the degree to which the two inhibitory discriminations were learned, it is unlikely that the greater level of inhibition on I resulted from better learning of the discrimination involving this cue.

As was intended, test ratings of the overshadowed positive cue (P^o) were lower than those for the single trained positive cue (P). Interestingly, the mean rating given to P^o was 28.0, a relatively low rating in comparison with other cues that had been paired with the outcome. An elemental associative model would have no trouble explaining such an effect at asymptote by assuming that cue X had acquired most of the positive strength available from the P^oX+ trials. This pattern is predicted because X always gains an increment in positive strength during the P^oX+ trials, whereas P^o gains increments on the P^oX+ trials but undergoes decrements on the P^oI^o– trials. Thus, P^o would be given low ratings, whereas X would be given a high rating. The fact that P^o had such low positive strength also would account for why no inhibitory effect was detected for I^o.

Higher ratings for X than for P^o can be predicted by configural theories also. Thus, following the stimulus generalization rules proposed by Pearce (1994), both X and P^o receive generalized positive strength from the unit P^oX+, but P^o also receives generalization from the slightly inhibitory unit P^oI^o–. However, configural theories would not predict as strong a difference between X and P^o as that anticipated by elemental theories, since only the latter assume competition between elements for the associative strength available on the P^oX+ trials such that the associative strength of P^o decreases as that of X increases. Unfortunately, we could not compare the strengths

of P^o and X in Experiment 1, since X was not included in the test. One way in which Experiment 2 differed from Experiment 1 was that it included X in the test.

EXPERIMENT 2

Using a design for the training phase similar to that of the previous experiment, the main aim of Experiment 2 was to provide further evidence that a stimulus combined with an overshadowed positive cue becomes less inhibitory than one combined with a simple positive cue. An additional aim was to test the predictions of both elemental and configural theories: first, that X should be given a higher rating than a simple overshadowed cue O, trained in OS+; and, second, that P^o should be given a lower rating than O. As can be seen from the overall design shown in Table 2, the compound cue OS+ was included in the training phase (equivalent to FG+ in Experiment 1; see Table 1) and the single cues O and X were included in the test.

Other changes from the previous experiment were the inclusion in the test of novel cue C and the two potential inhibitors, I and I^o. The latter two cues were added simply to ensure that some no-outcome cues occurred in the test. The addition of C allowed a comparison between the simple overshadowed cue O and a novel cue.

Finally, in order to distinguish between the two kinds of associative theories, we introduced a between-subjects factor to assess the effect of X- trials in the P^oX+, P^oI^o-, I^o- design. Group X- received intermixed X- trials throughout training, whereas Group Control received a control cue K- in their place. According to elemental theories, in the design P^oX+, P^oI^o-, I^o-, X-, the elements P^o and X will both acquire positive strength on the P^oX+ trials and will lose positive strength on the P^oI^o- and

X- trials, respectively. However, as I^o starts to acquire inhibitory properties, P^o will retain some positive strength on P^oI^o- trials, whereas the X- trials ensure that X continues to lose any positive strength it has gained. On subsequent P^oX+ trials, P^o will continue to gain positive strength as I^o gains inhibitory strength. The overall result is that the addition of X- trials will shift the positive strength acquired on P^oX+ trials from X to P^o and allow I^o to become as inhibitory as the standard inhibitor.

On the other hand, configural theories treat X- as a separate configural unit. Indeed, because the X configural unit draws generalized positive strength from the P^oX unit, it will become inhibitory. Generalization in the opposite direction will in turn produce an increase in positive strength held by the P^oX unit. Interestingly, the higher value now held by P^oX will allow the P^oI^o unit to gain a small amount of additional inhibitory strength. Because this gain is produced purely by generalization, it will remain low and well below the level of inhibition acquired by the standard inhibitor.

In summary, if participants encode compound cues as combinations of separate elements, associative theories predict that in Group X- the summation test will reveal I^o to be as effective an inhibitor as the standard inhibitor, I. In contrast, if participants encode compounds as configurations, then in Group X-I^o will be less effective than I.

Method

Participants. Sixty-seven first-year psychology students from the University of Sydney received course credit for taking part in the experiment. They were allocated to two groups according to the time when they arrived for participation ($n = 35$ for Group Control, $n = 32$ for Group X-).

Procedure. Experiment 2 was conducted in the same manner as Experiment 1, except for the following. In each of the four randomized 22-trial training blocks, each of the 11 types of trial listed in Table 2 occurred twice, except N, which occurred four times. When each participant logged on to a computer, the following 11 food names were randomly allocated to the 11 individual cues for Group Control, and 10 of the names were randomly allocated to the 10 individual cues for Group X-: *potatoes, rice, beef, chicken, carrots, melon, tomatoes, lettuce, biscuits, grapes, and apples*.

The test phase consisted of three randomized blocks, each containing 1 each of the 12 types of trial shown in Table 2. As in Experiment 1, the participants were divided into two additional conditions for the testing phase so as to counterbalance cue sequence across participants. Thus, 18 participants in Group Control and 16 participants in Group X- were given the compound sequences TI, TI^o, TN, and TC, whereas the remaining participants were given the reverse sequences—that is, IT, I^oT, NT, and CT.

Results

Performance during the training phase was analyzed in terms of three discrimination scores, exactly as in Experiment 1. When ratings were averaged over all eight training trials with each cue, the mean scores for the two groups were as follows. The simple discrimination score (T+ vs. N-) for Group Control was 66.4 and for Group X- it was 66.3. The standard inhibition score (P+ vs. PI-) was 43.0 for Group Control and 41.7 for Group X-. The mean overshadowed discrimination score (P^oX+ vs.

Table 2
Design and Test Ratings for Experiment 2

Test Cues	Ratings		
	Group Control	Group X-	
Training cues for Group Control/Group X-:			
TI	50.8	54.6	
TI ^o	61.1	50.3	
P+ (8), PI- (8), I- (8),	TN	60.6	49.8
P ^o X+ (8), P ^o I ^o - (8),	TC	66.2	65.3
I ^o - (8), T+ (8),	P	91.3	79.7
N- (16), OS+ (8)	P ^o	30.0	34.2
	T	85.9	88.1
Training cue for Group Control only: K- (8)	O	68.9	66.5
	X	68.3	10.2
	I	6.8	11.6
Training cue for Group X- only: X- (8)	I ^o	7.3	9.6
	C	43.9	46.5

Note—P, positive cue; I, inhibitory cue; P^o, overshadowed positive cue; X, extra cue used for overshadowing; I^o, overshadowed inhibitory cue; T, positive transfer cue; N, negative contingency control cue; K, control cue; O and S, overshadowing control cues; C, novel control cue. Presentation of the “migraine” outcome is denoted by “+,” and presentation of the “no migraine” outcome is denoted by “-.” The total number of presentations of each trial type is given in parentheses. Ratings were converted to a scale of 0 (“Migraine certain to occur”) to 100 (“Migraine certain to occur”) and averaged over three test trials.

P⁰I⁰-) was 50.6 for Group Control and 32.0 for Group X-; this was the only difference between groups detected during training [$F(1,65) = 7.95$]. Thus, the addition of X- in Group X- produced slower learning of the P⁰X+ versus P⁰I⁰- discrimination, but had no detectable effect on learning of the other two discriminations.

Within Group Control, the standard and overshadowed inhibition discrimination scores did not differ [$F(1,34) = 2.15$]. However, each was lower than the simple discrimination score [smallest $F(1,34) = 17.86$]. Within Group X-, the difference between the standard and the overshadowed inhibition discrimination scores approached statistical significance [$F(1,34) = 3.32, p = .08$]. As for Group Control, each of these scores was lower than the simple discrimination score [smallest $F(1,34) = 28.08$]. Thus, in both groups the simple discrimination between T+ and N- was learned at a faster rate than either inhibition discrimination.

Again as in Experiment 1, performance at the end of training was examined by basing discrimination scores on the last trial for each pair of cues. In Group Control, the simple, standard, and overshadowed discrimination scores were 84.5, 81.7, and 72.2, respectively, with no significant differences between them [largest $F(1,34) = 2.6$]. In Group X-, the simple, standard, and overshadowed discrimination scores were 87.1, 75.8, and 49.3, respectively; the differences between the simple and overshadowed discriminations [$F(1,31) = 12.6$] and between the standard and overshadowed discriminations [$F(1,31) = 5.8$] were both significant. Thus, the participants in this group were still performing poorly on the P⁰X+ versus P⁰I⁰- discrimination at the end of training.

Test ratings are shown in Table 2. A preliminary 2×2 analysis of variance with test compound cue order (e.g., TI vs. IT) and group (Group X- vs. Group Control) as factors was conducted across all test cues. It revealed a significant main effect of compound cue order whereby the group that was presented with the positively trained cue (T) to the left of its summation test compound partner (I, I⁰, N, or C) gave overall higher ratings across all test cues than did the group that received T as the second cue in the compound [$F(1,65) = 5.52$]. This effect—one we had not detected before—suggests that reading T first can bias ratings in a positive direction. Since this effect did not interact significantly with any of the within- or between-subjects comparisons, the main analysis did not include this factor.

The critical comparisons on the summation test cues were conducted across the groups factor in a mixed analysis and revealed a significant interaction between group and the TI versus TN difference [$F(1,65) = 7.51$], but no other significant interactions. These comparisons were then conducted within subjects separately for each group. In tests for standard inhibitory effects, ratings of TI were significantly lower than ratings of TN in Group Control [$F(1,34) = 6.83$] but not in Group X- [$F(1,31) = 1.63$]. Thus, standard inhibition training gave rise to inhibitory learning in Group Control only. In comparisons between the standard and the overshadowed inhibitor, TI was rated

significantly lower than TI⁰ in Group Control [$F(1,34) = 6.41$] but not in Group X- [$F(1,31) < 1$]. Neither group showed a significant difference between TI⁰ and TN (both F s < 1). In summary, the inhibitory effects found in Experiment 1 were replicated in Group Control but not in Group X-.

As in Experiment 1, ratings of P⁰ were low. The overshadowing effect on P⁰ was replicated in that P⁰ was rated significantly lower than P in a within-subjects comparison across both groups [$F(1,65) > 100$]. This result did not interact with group [$F(1,65) = 2.47$]. As was expected, a between-groups comparison showed that cue X was rated significantly lower by Group X- than by Group Control [$F(1,65) > 100$]. X and P⁰ were both compared to the simple overshadowed cue O in each group. In Group X-, both X [$F(1,31) > 100$] and P⁰ [$F(1,31) = 25.90$] were rated lower than O. In Group Control, P⁰ was rated significantly lower than O [$F(1,34) = 32.29$], but X was not significantly different from O [$F(1,34) < 1$]. As was noted above, it was expected that the inclusion of X- would increase ratings of P⁰. However, no group difference in ratings of P⁰ was found [$F(1,65) < 1$]. Finally, O was rated significantly higher than C, the novel test cue [$F(1,65) = 24.42$].

Discussion

As can be seen in a comparison of the ratings in Tables 1 and 2, the results for Group Control were very similar to those from Experiment 1. In particular, overshadowing the positive cue used in the training of an inhibitor again reduced the effectiveness of the inhibitor. As previously, this result did not appear to be compromised by differences in levels of learning about the two inhibitors in that, by the end of training, the overshadowed inhibition discrimination was well learned and not statistically weaker than the standard inhibition discrimination. In both experiments, it was also found that the overshadowed positive cue P⁰ was given low ratings, a result that is consistent with the finding that the overshadowed inhibitor, I⁰, did not produce a detectable inhibitory effect. The intention of adding X- to the overshadowed inhibition design was to test if this extra information would increase the positive strength of test cue P⁰ and, in turn, increase the inhibitory strength of the I⁰ element. However, Group X- did not confirm either prediction. The participants in Group X- gave P⁰ ratings as low as those given by Group Control. Furthermore, Group X- showed no evidence of inhibitory learning, even on the standard inhibitor (I).

Group X- showed relatively poor learning of the overshadowed inhibition discrimination even at the end of training. Although this might explain the ineffectiveness of the overshadowed inhibitor in the summation test, it does not explain why I was equally ineffective in this group. Group X- learned the standard inhibition discrimination well, performing at a level not significantly different from that of Group Control. An alternative explanation for the lack of inhibitory effects in Group X- is suggested by the finding that participants given predic-

tive tasks involving complex compound cues can adopt either an elemental or a configural strategy (Shanks, Charles, Darby, & Azmi, 1998). It may be that the extra complication resulting from the introduction of X- induced the participants in Group X- to use configural encoding to learn about PI- trials and then to show little or no generalization to the TI cue presented on test. In any case, it is clear that including X- in the overshadowed inhibition design did not lead to the increase in inhibitory learning predicted by elemental theory.

A further aim of Experiment 2 was to compare test ratings of P^o and X in Group Control. We expected that the participants would attribute more than an even share of positive strength to X and less than an even share to P^o. Our expectations were partly confirmed in that P^o was rated lower than the simple overshadowed cue O. On the other hand, ratings of X were no different than ratings of O. The lack of a difference between X and O ratings may reflect a measurement problem. Note that the novel cue C was rated close to the midpoint (marked "Don't know"), as is appropriate given its unknown value. But what rating should be made by a participant who assigns a 50% likelihood to a cue such as O? Experiment 3 addressed this issue by using a scale that did not include a "Don't know" label.

EXPERIMENT 3

The scale used in the previous experiments may have encouraged the participants to rate cues along two subjective dimensions: from "Don't know" to "No outcome" and from "Don't know" to "Outcome." If they judge a simple overshadowed cue to be about 50% predictive of the outcome, they may rate this cue halfway between "Don't know" and "Outcome" (i.e., a rating of 75) instead of halfway between "No outcome" and "Outcome" (i.e., a rating of 50). In other words, the simple overshadowed cue may have been given high ratings in Experiment 2 because the participants were reluctant to rate a cue at "Don't know" when they did know something about that cue's predictive strength.

With conditions similar to those of Group Control in the previous experiment, the aim of Experiment 3 was to compare the ratings of cues X, P^o, and O on a scale without a midpoint label. As is shown in Table 3, the training phase remained the same, but in the test a single cue S was added to provide an additional measure of a simple overshadowed cue. The single test cues I and I^o were omitted to reduce the length of the test.

Method

Participants. Forty-two first-year psychology students from the University of Sydney received course credit for taking part in the experiment.

Procedure. The procedure used in Experiment 3 was essentially the same as that used in the previous experiments, except for minor changes in instructions, in test cues, and in the rating scale. The following paragraph contains the only changes made to the instructions given at the beginning of training:

Table 3
Design and Test Ratings for Experiment 3

	Test Cues	Ratings
Training cues: P+ (8), PI- (8), I- (8), P ^o X+ (8), P ^o I ^o - (8), I ^o - (8), T+ (8), N- (16), OS+ (8)	TI TI ^o TN TC P P ^o T O S X C	50.1 59.4 57.1 65.7 82.7 31.4 95.8 68.0 70.4 79.5 48.3

Note—Cue functions are denoted by letters as in Table 2, except that control cue K was not used. The total number of presentations of each training cue is given in parentheses. Ratings were converted to a scale of 0 ("Migraine certain not to occur") to 100 ("Migraine certain to occur") and averaged over three test trials.

When you begin the simulation, you will be shown which foods have been consumed on a given day. A rating line will be presented on the screen. The rating line represents a scale from "Migraine certain NOT to occur" to "Migraine certain to occur." The midpoint represents a 50/50 chance that the outcome will/won't occur. You must make a prediction about whether you expect a migraine to occur by clicking on the line. An arrow will then mark the point that you have chosen. When you have made your prediction, you will be shown the information indicating whether a migraine occurred or not.

The rating scale was the same length and in the same position on the screen as in the previous experiments. The left and right end labels remained the same, as did a small vertical line marking the midpoint. Initially, no pointer appeared on the scale. Instead, when a participant used the computer mouse to click on the scale, a red arrow appeared at that position. The red arrow could be repositioned by clicking on another point on the scale.

Exactly as for the control group in Experiment 2, in each of the four 20-trial training blocks each type of trial listed in Table 3 occurred twice, except N, which occurred four times. All training blocks were randomized for each participant. Each participant received the following 10 food names randomly allocated to the 10 individual cues: *potatoes, rice, beef, chicken, carrots, bananas, melon, lettuce, biscuits, and tomatoes*. There were three randomized blocks in the test, each block containing 1 of each of the 11 types of test trial. As previously, the participants were divided into two groups for the test so as to counterbalance the cue sequence (e.g., TI vs. IT).

Results

Training performance was analyzed in terms of the three discrimination scores used in the previous experiments. As before, over the whole training phase the simple discrimination score, T+ versus the first eight trials of N (64.6), was greater than both the standard inhibition score, P+ versus PI- [41.6; $F(1,41) = 28.83$] and the overshadowed inhibition score, P^oX+ versus P^oI^o- [56.5; $F(1,41) = 4.31$]. Furthermore, for the first time in this series of experiments the overshadowed inhibition score was reliably greater than the standard inhibition score [$F(1,41) = 10.57$].

When based only on the last trial for each cue, the simple discrimination score (92.9) was significantly higher than both the standard (74.6) and the overshadowed (76.7) inhibition scores [$F_s(1,41) = 13.04$ and 5.46, respec-

tively]. However, there was no longer any difference between the standard and overshadowed discriminations [$F(1,41) < 1$].

The ratings given to each test cue are shown in Table 3. The basic overshadowing effect of X on P^o was indicated by a lower rating of P^o than of P [$F(1,41) > 50$]. Because for the first time both O and S elements from the trained OS+ compound were included in the test, the average rating of O and S was used for the simple overshadowed control measure, referred to as O/S. P^o was rated significantly lower than O/S [$F(1,41) = 43.83$]. More interesting was the finding that X was rated higher than O/S [$F(1,41) = 5.98$]. The fact that X was not rated as high as T [$F(1,41) = 4.27$] indicates that X was not at full strength as a positive cue. The comparison between O/S and C showed that O/S was given significantly higher ratings than C [$F(1,41) = 34.48$].

As for inhibitory effects, TI received significantly lower ratings than the control TN [$F(1,41) = 4.99$], indicating a reliable standard inhibition effect. The overshadowed inhibitor compound TI^o was rated significantly higher than TI [$F(1,41) = 8.71$], but not significantly different from TN [$F(1,41) < 1$].

Discussion

As in the first two experiments, the overshadowed inhibition design failed to produce any detectable inhibition, whereas the standard inhibition design produced a reliable effect. Furthermore, P^o gained less positive value than did a simple overshadowed control. Also, in the present experiment X gained more positive value than did the overshadowed control, whereas no such difference was found in the control condition of Experiment 2. It is likely that the main factor contributing to the successful detection of this predicted difference was the modification of the rating scale introduced in Experiment 3—namely, removal of the midpoint label “Don’t know.”

An interesting aspect of training was that the overshadowed inhibition discrimination was learned more quickly than the standard inhibition discrimination. The possible implication of this finding is discussed later. However, even though these discriminations were learned at different rates, they reach similar levels by the end of training.

EXPERIMENT 4

As an alternative to overshadowing, reducing the number of training trials on which the positive cue occurs can weaken that cue’s positive strength. This suggests a possibly simpler way of addressing the main issue of the present study. In an eye-blink conditioning experiment in which rabbits served as subjects, Wagner and Saavedra (briefly reported in Wagner, 1971) tested for inhibitory learning supported by either Stimulus A, which had been given extensive excitatory training, or Stimulus B, which had been given only brief training. A subsequent summation test revealed a greater inhibitory effect of a stimulus trained against A than of one trained

against B. In the present experiment, a similar within-subjects design was used. The first cue, I, was trained against P in the standard inhibition design. The second (reduced inhibitory) cue, I^r, received similar training (P^{r+}, P^rI^{r-}, I^{r-}) except that reduced positive cue P^{r+} was given half the number of positive training trials given to P.

Experiment 4 also included an additional control cue, M, for inhibitory learning. This control was trained in the combination LM⁻, M⁻, and the compound TM was added to the summation test. This additional control was intended to follow the pattern of training given to I (and I^r) and not simply to provide the same number of training trials, as was the case for the control cue N⁻ throughout this set of experiments. The complete design is shown in Table 4.

Method

Participants. Forty-five first-year psychology students from the University of Sydney participated in this experiment and received course credit for taking part.

Procedure. The basic procedure of Experiment 4 was the same as that of Experiment 3, including the use of the new scale. In each of the four 22-trial training blocks, each type of trial listed in Table 4 occurred twice, except N, which occurred four times, and P^r, which occurred only once. The names of the foods were the same as in Experiment 3. There were three randomized blocks in the test, each containing 1 of each of the 10 types of test trial. As in the previous experiments, the participants were divided into two groups for the testing phase so as to counterbalance cue sequence across participants.

Results

Training performance was analyzed in terms of four discrimination scores. Two simple discriminations, T+ versus the first eight trials of N⁻ and T+ versus LM⁻, were each compared with the standard inhibition discrimination, P+ versus PI⁻, and the reduced inhibition discrimination, P^{r+} versus P^rI^{r-}. The standard inhibition score (51.7) was again significantly lower than both the simple discrimination scores [64.4 and 63.2, respec-

Table 4
Design and Test Ratings for Experiment 4

Training cues	Test Cues	Ratings
P+ (8), PI ⁻ (8), I ⁻ (8), P ^{r+} (4), P ^r I ^{r-} (8), I ^{r-} (8), T+ (8), N ⁻ (16), LM ⁻ (8), M ⁻ (8)	TI TI ^r TN TM P P ^r T N I C	34.6 37.5 38.6 39.7 90.2 86.3 91.5 6.5 6.1 43.9

Note—P, positive cue; I, inhibitory cue; P^r, reduced positive cue; I^r, reduced inhibitory cue; T, positive transfer cue; N, negative contingency control cue; L, and M, additional control cues for the potential inhibitory effects of I and I^r on test; C, novel control cue. Presentation of the “migraine” outcome is denoted by “+,” and presentation of the “no migraine” outcome is denoted by “-.” The total number of presentations of each training cue is given in parentheses. Ratings were converted to a scale of 0 (“Migraine certain not to occur”) to 100 (“Migraine certain to occur”) and averaged over three test trials.

tively; $F_s(1,44) = 16.18$ and 19.41 , respectively]. The reduced inhibition score (39.0) was significantly lower than all three of the above scores [smallest $F(1,44) = 16.13$].

A similar pattern was found for scores based only on the last trial for each cue. The means for the N- simple discrimination, the LM- simple discrimination, the standard, and the partial inhibition were 88.1, 85.3, 79.5, and 65.5, respectively. There was no difference between the two simple discriminations [$F(1,44) = 2.64$]. The standard inhibition score was lower than the N- simple discrimination [$F(1,44) = 4.31$] but was not significantly lower than the LM- simple discrimination [$F(1,44) = 1.70$]. The partial inhibition score was statistically the lowest [smallest $F(1,44) = 8.43$], indicating that reducing the number of P+ trials had produced poorer learning of the reduced inhibition discrimination than of the standard inhibition discrimination.

The mean ratings given to each test cue are shown in Table 4. The ratings given to P^r (presented only four times during training) were lower than the ratings given to P (presented eight times during training), this difference being marginally significant [$F(1,44) = 3.04, p = .088$]. The summation test showed that the standard inhibition design produced reliable inhibition against both controls, since the compound TI was rated significantly lower than both TN [$F(1,44) = 7.07$] and TM [$F(1,44) = 9.87$]. There was no significant difference between these two controls [$F(1,44) < 1$].

The most important question was how cue I^r would function in the summation test. The compound TI was rated significantly lower than TI^r [$F(1,44) = 5.92$], indicating that I^r gained less inhibitory strength than I. However, no reliable inhibition was found on I^r since TI^r was not rated significantly lower than either of the summation test controls [$F_s(1,44) = 0.50$ and 1.50].

Discussion

The main finding from this experiment—that inhibitory training using a positive cue given limited training was less effective than inhibitory training using a positive cue given twice as much training—provides further evidence that the effectiveness of an inhibitory cue depends on the strength of its positive cue. As previously, no inhibitory effect of the amended inhibitor was detected. However, unlike in the previous experiments, the strength of the reduced positive cue (P^r), was high and not significantly weaker than the full positive cue (P). This is not necessarily a problem for the claim that inhibitory learning to I develops in proportion to the positive strength of P, since, according to associative learning theories, inhibitory learning can occur only after the positive cue has acquired some strength. Because the four P+ trials were scattered throughout the training phase, there may not have been sufficient opportunity for the development of inhibitory learning supported by the final positive strength of P^r.

GENERAL DISCUSSION

As was noted in the introduction, contingency accounts of how humans come to assign inhibitory value to a cue deny the traditional assumption that this depends on the strength of the positive cue with which the putative inhibitor is compounded. The main motivation for the present study was to test this rarely examined assumption. All four experiments provided clear evidence that it is correct and, thus, suggest that contingency models do not provide a satisfactory account of inhibitory learning.

The probabilistic contrast model (Cheng & Novick, 1992) was designed specifically to explain how participants judge the causal strength of cues. Although in the present experiments the participants were not asked to predict causal strength, it is highly likely that they used a notion of causality to guide their predictions. If the contingency calculation for the inhibitory cue is confined to the focal set containing P, as is described by this model, the participants should judge I as an inhibitor because I holds a contingency of -1 in this focal set. Note that a smaller number of P+ trials in the same focal set does not change the contingency between I and the outcome. Because I and I^r in Experiment 4 had identical focal sets and equal negative contingencies with the outcome, they should have had equal inhibitory strength. Consequently, the difference between I and I^r found in Experiment 4 appears to contradict this model.

The probabilistic contrast model does, however, make the assumption that there is a monotonic relationship between the number of trials witnessed and confidence in rating the cue. Following this assumption, ratings for a particular cue approach the contingency for that cue as the number of trials increases. However, it is not clear how this assumption applies to cues such as P^r in the present experiments, where P^r defines the focal set for a particular target cue rather than act as the target cue itself. An interpretation is possible in which a smaller number of P+ trials may reduce confidence in the presence or suitability of the focal set. In turn, this may reduce confidence in the judgment of I^r. Therefore, the focal set aspect of the probabilistic contrast model might provide an account of the lack of inhibitory learning on I^r in Experiment 4 by recourse to a confidence explanation.

A focal set account of the results obtained from the overshadowed inhibition design used in Experiments 1, 2, and 3 is more complicated. The first thing to be considered is whether there is a focal set containing a positive cause available for I^o. Notice that P^o holds a positive contingency with the outcome. However, when P^o is followed by the outcome, it occurs always alongside X. This additional cue is a possible alternative cause of the outcome and, therefore, there is no focal set available for P^o that excludes alternative causes. Thus, participants should consider the causal status of P^o to be unknown. Similarly, because there is no known positive cause on trials in

which I^0 occurs, the lack of an appropriate focal set would prevent participants from forming a judgment about I^0 . If it is assumed that participants rate cues with unknown causal strength at the midpoint of the rating scale, the probabilistic contrast model might provide an account along these lines for the ineffectiveness of the overshadowed inhibitor in the first three experiments.

It is interesting to follow the predictions made by the probabilistic contrast model for other cues in the overshadowed inhibition design. As was noted above, the strength of P^0 should be considered unknown because an alternative cause X is constantly present. The same prediction would apply to the judgment of X because P^0 is present whenever X is followed by the outcome. Similarly, after training in the compound $OS+$, judgment of the simple overshadowed cue O also should be withheld due to the constant presence of cue S . This account suggests that all three cues should be given the same rating. In contrast to this prediction, Experiment 3 found that the participants made different ratings of the predictive strengths of the three cues.

The probabilistic contrast model nonetheless may allow for the detection of a cue–no outcome relation for P^0 . This may happen because of P^0I^0- trials, in which a positive relation for P^0 would be precluded. In this case, no alternative cause would be present on the P^0X+ trials and the positive contingency between X and the outcome would be used as the basis for the judgment. This particular version of the probabilistic contrast model could account for the pattern of ratings given to cues P^0 , X , and O in Experiment 3. Power probabilistic contrast theory (Cheng, 1997; but see Allan, 2003) is a development of the probabilistic contrast model that includes a base rate normalization function. Because the contingencies between the target cues and outcome in the present experiments were always the same, the interpretation above is also required by the power theory in order to explain the present findings.

A final point to be made about the probabilistic contrast model is that when designs allowing equal confidence of judgment are compared, the model proposes that participants use contingency as a measure of causal strength or else judge the causal status of a cue as unknown. Therefore, once learning—or confidence—has reached asymptote, the model is basically an all-or-none theory of learning. With regard to inhibitory effects, the presence of a positive cue in the focal set of a potential inhibitor merely permits the interpretation of inhibition. Because the model is normative in its basis, the *amount* of inhibitory strength is determined by the negative contingency. This means that the model can account for null effects but not for partial effects. Although this theoretical difference between contingency-based and associative theories is testable in principle, an experimental test for a reduced but significant inhibitory effect would need much more statistical power than that of the present experiments.

Whereas the probabilistic contrast model needs assumptions auxiliary to the basic contingency account in order to explain the present results, in general the results fit comfortably within the framework provided by associative learning theories. It is therefore of interest to examine the question of what kind of associative theory—in particular, elemental or configural—provides the most satisfactory account.

As was noted above, elemental associative learning theories (see, e.g., Rescorla & Wagner, 1972) assume that in the design $P+$, $PI-$, $I-$, the positive strength held by P proportionally determines the inhibitory strength acquired by I . As well as predicting this main result from the present study, such an approach also is able to account for related effects produced by the overshadowed inhibition design, P^0X+ , P^0I^0- , and I^0- . Because such models assume that the elements share a limited amount of associative strength, X gains extra strength at the expense of P^0 . In contrast, a simple overshadowing design, as that used in the training of $OS+$, produces equal strength for both cues. As is consistent with this analysis, in Experiments 2 and 3 P^0 was rated lower than the simple overshadowed cue, and in Experiment 3 X was rated higher than the overshadowed cue.

As was noted earlier, Pearce's (1994) configural account of inhibitory learning also predicts that this depends on the strength of the positive cue. Once the P configural unit has started to gain positive strength, this will generalize to the configural unit for PI , but there is no outcome on such trials and, thus, PI will gain inhibitory strength. In a reciprocal manner, the P unit will increase its positive strength above asymptotic level in order to counteract generalized inhibitory strength from the PI unit. Importantly, the generalization assumption is necessary both for inhibitory learning on the PI unit and for the detection of conditioned inhibition effects in the summation test due to similarity between the PI and TI units. Thus, if generalization between these units is weak, then only weak inhibitory effects will be found on test. It should be noted that the effect of the standard inhibitor (I) as measured by the difference between TI and the control compound TN , TC , or TM , was consistently small in all four experiments. Elemental theory would predict much lower ratings of TI .

The generalization mechanism proposed by Pearce (1994) also explains why inhibitory learning is reduced in the overshadowed inhibition design; it is because there is less generalization between P^0X and P^0I^0 than between P and PI . A similar effect occurs when the degree of similarity is the same but the amount of positive strength is smaller. Thus, the partial positive strength of P^0 in Experiment 4 should support a smaller inhibitory effect. Pearce's configural theory predicts the reduction in inhibitory effects observed in this series of experiments as readily as do elemental theories.

Configural theory also can explain the pattern of overshadowed cues in the overshadowed inhibition design, as

long as an assumption is added that is appropriate for human predictive judgments. Pearce's (1994) model was developed to account for the results from experimental contexts in which—as in animal experiments—a novel cue elicits little or no response on test. However, in human predictive judgments a novel cue is typically rated around 50 (on a 0–100 scale; see Tables 2, 3, and 4). Therefore, a novel unit's default mode of a rating of about 50 must be taken into account when calculations are made for similarity or generalization functions. For example, if test cue O receives generalized strength from the training compound OS+, then one would expect O to be rated about midway between the ratings given to OS and to a novel cue. This is close to the value we observed for O and O/S in Experiments 3 and 4. Furthermore, the configural unit P^oX+ will have greater strength than a simple positive cue because of the need to counteract generalized inhibitory strength from the compound P^oI^o. This excess positive strength should produce a slightly higher rating for X than for O. Again, this prediction closely matches the findings in Experiment 3.

Pearce (1994) also predicted that P^o should generalize to both P^oX+ and the slightly inhibitory cue P^oI^o-. This will lead to a rating slightly lower than 50 for P^o, even with a regression toward a novel cue's rating. This prediction is consistent with the direction in which P^o was rated. However, what is problematic for this account is that P^o was consistently given very low ratings in Experiments 1, 2, and 3, as is more consistent with elemental theories.

Additional support for a configural analysis comes from the training data. In Experiments 1, 2, and 3, the discrimination scores for the overshadowed inhibition design tended to be higher than those for the standard inhibition design, and significantly so in Experiment 3. This indicates that the overshadowed inhibition discrimination was the easier of the two to learn. An elemental analysis needs to take into account three elements in the overshadowed discrimination but only two in the standard discrimination, and this should mean that the overshadowed discrimination is learned more slowly. On the other hand, according to a configural analysis both discriminations involve two units, but the units in the overshadowed discrimination are more distinct and, as a result, this discrimination is easier to learn.

As a final comment on the critical role of generalization rules in explaining inhibitory learning in terms of configural learning theories, the results from Group X- in Experiment 2 should be noted. This group showed an atypical and complete absence of the standard inhibitory effect despite standard inhibitory training. As was suggested earlier, the only obvious explanation is that, when faced with a particularly complicated set of compound cues, these participants adopted a strategy of configural encoding that allowed little generalization between units. Thus, in human predictive judgment tasks the degree of generalization may not be determined by fixed rules of the kind proposed by Pearce (1994), but may vary with the task conditions. Such models have been developed

(Kruschke, 1992) but are yet to fully explain the conditions that determine generalization adjustments (see Shanks et al., 1998). This suggests the unappealing but plausible possibility that in experiments with a complex set of single and compound cues, as those used throughout the present study, the mean scores conceal a variety of learning strategies that vary across participants with respect to both elemental or configural encoding of compounds and generalization from one compound to another (also see Williams, Sagness, & McPhee, 1994).

A very different approach to the analysis of human predictive judgments is the logical reasoning or propositional model (see, e.g., De Houwer & Beckers, 2003; Lovibond & Shanks, 2002; Mitchell & Lovibond, 2002; but see also Livesey & Boakes, 2004). If participants use an "if-then" proposition to form conclusions about individual cues, they could arrive at the proposition that if no outcome follows P^oI^o, then neither P^o nor I^o leads to the outcome. If this is the case, then X alone must be responsible for the outcome on the P^oX+ trials. These propositions naturally lead to the pattern of ratings given in the test phases. However, this is only one of two possible solutions; the other is that P^o is responsible for the outcome and I^o is an inhibitor. Therefore, it may seem that a solution not requiring inhibition is favored over one requiring inhibition. However, it must be pointed out that the control cue N, trained as N-, also has two possible solutions because it cannot be ruled out that N is not an inhibitor. Therefore, the pattern of ratings could be made to fit a logical analysis, if only with hindsight. As with the addition of confidence assumptions to the probabilistic contrast model, a reasoning model could make recourse to a confidence variable in order to explain the findings of Experiment 4. Overall, though, until such an approach is made more precise, it remains incapable of providing a detailed account of inhibitory learning.

In summary, associative theories, a logical reasoning approach, and the probabilistic contrast model can all account for the absence of inhibition on the overshadowed inhibitor—given some supplementary assumptions—and associative theories can in addition provide a straightforward explanation for the differences in strengths of the supporting cues in the overshadowed inhibition design. Among associative theories, elemental models can explain most of the test results found in the present series of experiments, but only Pearce's (1994) configural model can account for both test and training data. However, this particular version of configural theory with its specific generalization rules fails to account for the lack of inhibitory effects in the more complicated design given to Group X- in Experiment 2. This result may require a configural model allowing flexibility in generalization. Aside from its theoretical implications, in terms of empirical contribution the present study shows that varying the strength of positive cue in two quite different ways in a human predictive learning task influences the effectiveness of an inhibitor with which it is presented in training.

REFERENCES

- ALLAN, L. G. (2003). Assessing power PC. *Learning & Behavior*, **31**, 192-204.
- CHAPMAN, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Learning, Memory, & Cognition*, **17**, 837-854.
- CHAPMAN, G. B., & ROBBINS, S. J. (1990). Cue interaction in human contingency judgment. *Journal of Experimental Psychology: Memory & Cognition*, **18**, 537-545.
- CHENG, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review*, **104**, 367-405.
- CHENG, P. W., & NOVICK, L. R. (1992). Covariation in natural causal induction. *Psychological Review*, **99**, 365-382.
- CHENG, P. W., PARK, P., YARLAS, A. S., & HOLYOAK, K. J. (1996). A causal power theory of focal sets. In D. R. Shanks, K. J. Holyoak, & D. R. Medin (Eds.), *The psychology of learning and motivation: Vol. 34. Causal learning* (pp. 315-355). San Diego: Academic Press.
- DE HOUWER, J., & BECKERS, T. (2003). Secondary task difficulty modulates forward blocking in human contingency learning. *Quarterly Journal of Experimental Psychology*, **56B**, 345-357.
- DICKINSON, A. (2001). Causal learning: An associative analysis. *Quarterly Journal of Experimental Psychology*, **54B**, 3-25.
- HEARST, E., BOTTJER, S. W., & WALKER, E. (1980). Conditioned approach-withdrawal behavior and some signal-food relations in pigeons: Performance and positive vs. negative "associative strength." *Bulletin of the Psychonomic Society*, **16**, 183-186.
- HEARST, E., & FRANKLIN, S. R. (1977). Positive and negative relations between a signal and food: Approach-withdrawal behavior to the signal. *Journal of Experimental Psychology: Animal Behavior Processes*, **3**, 37-52.
- KARAZINOV, D. M., & BOAKES, R. A. (2004). Learning about cues that prevent an outcome: Conditioned inhibition and differential inhibition in human predictive learning. *Quarterly Journal of Experimental Psychology*, **57B**, 153-178.
- KRUSCHKE, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **99**, 22-44.
- LE PELLEY, M. E., & McLAREN, I. P. L. (2001). Retrospective reevaluation in humans: Learning or memory? *Quarterly Journal of Experimental Psychology*, **54B**, 311-352.
- LIVESEY, E. J., & BOAKES, R. A. (2004). Outcome additivity, elemental processing and blocking in human causality judgements. *Quarterly Journal of Experimental Psychology*, **57B**, 361-379.
- LOLORDO, V. M., & FAIRLESS, J. L. (1985). Pavlovian conditioned inhibition: The literature since 1969. In R. R. Miller & N. E. Spear (Eds.), *Information processing in animals: Conditioned inhibition* (pp. 1-49). Hillsdale, NJ: Erlbaum.
- LOVIBOND, P. F., & SHANKS, D. R. (2002). The role of awareness in Pavlovian conditioning: Empirical evidence and theoretical implications. *Journal of Experimental Psychology: Animal Behavior Processes*, **28**, 3-36.
- MILLER, R. R., HALLAM, S. C., HONG, J. Y., & DUFORE, D. S. (1991). Associative structure of differential inhibition: Implications for models of conditioned inhibition. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 141-150.
- MITCHELL, C. J., & LOVIBOND, P. F. (2002). Backward and forward blocking in human electrodermal conditioning: Blocking requires an assumption of outcome additivity. *Quarterly Journal of Experimental Psychology*, **55B**, 311-329.
- PAVLOV, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, Trans.). London: Oxford University Press.
- PEARCE, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, **101**, 587-607.
- RESCORLA, R. A. (1969). Conditioned inhibition of fear resulting from CS-US contingencies. *Journal of Comparative & Physiological Psychology*, **67**, 504-509.
- 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.
- SHANKS, D. R., CHARLES, D., DARBY, R. J., & AZMI, A. (1998). Configural processes in human associative learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1353-1378.
- WAGNER, A. R. (1971). Elementary associations. In H. Kindler & J. T. Spence (Eds.), *Essays in neobehaviorism: A memorial to Kenneth W. Spence* (pp. 187-213). New York: Appleton-Century-Crofts.
- WARD, W. C., & JENKINS, H. M. (1965). The display of information and the judgment of contingency. *Canadian Journal of Psychology*, **19**, 231-241.
- WILLIAMS, D. A. (1995). Forms of inhibition in animal and human learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **21**, 129-142.
- WILLIAMS, D. A. (1996). A comparative analysis of negative contingency learning in humans and nonhumans. In D. R. Shanks, K. J. Holyoak, & D. R. Medin (Eds.), *The psychology of learning and motivation: Vol. 34. Causal learning* (pp. 89-131). San Diego: Academic Press.
- WILLIAMS, D. A., & DOCKING, G. L. (1995). Associative and normative accounts of negative transfer. *Quarterly Journal of Experimental Psychology*, **48A**, 976-988.
- WILLIAMS, D. A., SAGNESS, K. E., & MCPHEE, J. E. (1994). Configural and elemental strategies in predictive learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **20**, 694-709.
- WITCHER, E. S., & AYRES, J. J. B. (1980). Systematic manipulation of CS-US pairings in negative CS-US correlation procedures in rats. *Animal Learning & Behavior*, **8**, 67-74.

(Manuscript received August 28, 2003;

revision accepted for publication March 17, 2004.)