Learning & Behavior
2010, 38 (2), 145-159
doi:10.3758/LB.38.2.145

Effect of outcome valence on positive and negative patterning in human causal reasoning

J. W. Whitlow, Jr.
Rutgers University, Camden, New Jersey

The importance of configural cues and whether a situation involves beneficent or maleficent outcomes was investigated in two experiments on human causal reasoning, based on experienced causal information. Participants learned positive and negative patterning discriminations involving either beneficent or maleficent outcomes in a health-reasoning task and in a social-reasoning task. With maleficent outcomes, positive patterning was consistently easier to learn than negative patterning, a positive patterning advantage that is predicted by current associative theories and commonly taken as evidence for configural cues. However, with beneficent outcomes, the two discrimination tasks were not significantly different in ease of learning, a result not predicted by current theories. The reliable positive patterning effect found with maleficent outcomes broadens the range of conditions in which the effect can be shown in causal reasoning. The novel effect of outcome valence poses an interesting theoretical challenge for attempts to account for the relation between learning about individual cues and combinations of those cues.

Patterning discriminations have been widely viewed as an effective means for demonstrating the role of configural cues in associative learning (e.g., Rescorla, 1972; Whitlow & Wagner, 1972), and this article reports two experiments that used patterning discriminations to provide evidence for configural cues in a human causal-reasoning task. Although researchers using laboratory-based approaches to understand human causal reasoning have emphasized the role of configural learning in this domain (e.g., Shanks, Charles, Darby, & Azmi, 1998; Shanks, Darby, & Charles, 1998; Williams, Sagness, & McPhee, 1994; Young, Wasserman, Johnson, & Jones, 2000), there has been little evidence for such learning in terms of this traditional measure.

Patterning discriminations establish differential responding between a set of stimulus constituents (say, A and B) and a compound stimulus made up of the constituents (AB). In negative patterning, trials in which the constituents occur by themselves are reinforced (i.e., followed by an identified outcome and designated A+ or B+), whereas trials in which the compound stimulus occurs are nonreinforced (i.e., followed by no outcome, here designated AB0).[1] In contrast, *in positive patterning*, trials in which the constituents occur by themselves are nonreinforced (i.e., followed by no outcome, here designated A0 and B0), whereas trials in which the compound stimulus occurs are reinforced (i.e., followed by an identified outcome and designated AB+).

Contemporary theories of associative learning generally agree in predicting that positive patterning is an easier discrimination to learn than negative patterning. This agreement in predicted outcome is impressive, because the theories offer widely divergent accounts of the nature of learning. For example, Pearce (1994, 2002) proposed a configural model of associative learning, which assumes that learning involves associating stimulus configurations with outcomes. In this account, negative patterning is more difficult than positive patterning because of similarities between the nonreinforced context and the stimulus constituents (A and B). In negative patterning, the nonreinforcement of the context is in opposition to the reinforcement of the constituents (A+ and B+), whereas in positive patterning, the nonreinforcement of the context matches the nonreinforcement of the constituents (A0 and B0). In contrast, Rescorla and Wagner (1972) proposed an elemental account of associative learning, which assumes that learning involves associating component stimuli individually with outcomes. According to their account, the negative patterning task can be learned only if the compound stimulus (AB) includes configural cues as well as the identifiable constituents (A and B), and the configural cues come to suppress response tendencies to the constituent cues. This account predicts that positive patterning will be easier, because it does not require use of configural cues for successful discrimination (although such cues might in fact be used; see Rescorla, 1973). More recently, several theorists (Harris, 2006; McLaren & Macintosh, 2002; Wagner, 2003) have proposed microelemental accounts of learning, which assume that learning involves associating microfeatures (McLaren & Macintosh, 2002) of individual stimuli with outcomes. In addition to making different assumptions about the functional stimulus

J. W. Whitlow, Jr., bwhitlow@camden.rutgers.edu

to which an association is attached, the various theories make different assumptions about the way learning is distributed across elements, what factors are relevant for learning to take place, and how associations generalize from one stimulus to another. Nonetheless, the theories agree that successful performance on the negative patterning task implies that configural cues of some kind contribute to the discrimination.

The design of the experiments to be reported here was guided by two considerations, one having to do with patterning discriminations in general and one having to do with the implementation of patterning discriminations in human learning.

A Patterning Discrimination Problem

The agreement among the competing theories in the predicted advantage of positive over negative patterning (a positive patterning advantage) obscures an inconvenient fact, which is that little difference between the two types of patterning discrimination has been reported in several empirical studies. For example, Woodbury (1943), who first demonstrated patterning discriminations with dogs learning positive and negative patterning discriminations of tones paired with food, found that negative patterning was learned as easily as positive patterning. Rescorla (1972), using rats in a discrete-trial barpress task with food reward, found only a small advantage for learning positive relative to learning negative patterning discriminations. More recently, Deisig, Lachnit, Giurfa, and Hellstern (2001) found similar learning curves for positive and negative patterning using honeybees in an olfactory discrimination task (but see also Deisig, Sandoz, Giurfa, & Lachnit, 2007). In a study of human eye-blink conditioning, Kinder and Lachnit (2002) even reported the intriguing suggestion of better discriminative learning for negative than for positive patterning at the relatively long interstimulus interval (ISI) of 1.2 sec.

In light of these results, in the experiments reported here, I attempted to evaluate whether the valence of the outcome might be a contributing factor in producing failures to find a positive patterning advantage. That is, an interesting feature of many studies in which a positive patterning advantage was not found is that they involve an outcome that is desirable (here termed *beneficent*), whereas studies in which a positive patterning advantage was found have involved an outcome that is undesirable (here termed *maleficent*). This feature is neatly illustrated in an article by Bellingham, Gillette-Bellingham, and Kehoe (1985), who compared positive and negative patterning in an appetitive approach task in rats and in an eye-blink conditioning task in rabbits. In the appetitive task with rats, Bellingham et al. found that initial discrimination learning was about equally as easy for the negative patterning task as for the positive patterning task, replicating the findings of Rescorla (1972). In the eye-blink task with rabbits, however, they found that the two discriminations exhibited striking differences in ease of learning, with good discrimination in a positive patterning task coupled with no evidence of discrimination in the negative patterning task.

Although the two paradigms studied by Bellingham et al. (1985) differ in many aspects (species, tasks, trial durations, etc.), it is notable that the outcome for the first paradigm was a beneficent one (food, to a hungry rat), whereas the outcome for the second paradigm was a maleficent one (shock to the skin about its eye, to a rabbit). Thus, one interpretation of these data is that a positive patterning advantage is readily found with maleficent outcomes but is not as readily found with beneficent outcomes. Such an interpretation would be consistent with the fact that the outcomes in most of the cases noted earlier in which investigators failed to find much difference between the two types of patterning discriminations were beneficent (e.g., Deisig et al., 2001; Rescorla, 1972; Woodbury, 1943).

To evaluate the role of outcome valence on the positive patterning advantage, in the present experiments, I used both beneficent and maleficent outcomes in a human causal-reasoning task by incorporating a contrast similar to one originally introduced by Le Pelley, Oakeshott, Wills, and McLaren (2005). In their study of causal reasoning about food, Le Pelley et al. asked participants to learn which food items led to enjoyment (a beneficent outcome), as well as to learn which food items led to allergic reactions (a maleficent outcome). In the experiments reported here, participants also learned about causal relations in which the outcomes were sometimes beneficent and sometimes maleficent.

A Patterning Discrimination Problem in Human Causal Reasoning

In the case of human causal reasoning, an additional consideration was that in most studies, a positive patterning advantage has not been found. Young et al. (2000, Experiment 1), for example, found similar rates of learning for positive and negative patterning in a medical diagnosis task, thereby replicating findings from earlier studies by Shanks, Charles, et al. (1998) and Shanks and Darby (1998), using a food allergy task. More recently, Harris and Livesey (2008) found no significant differences in rates of discrimination learning for positive and negative patterning in two different experiments, one involving an educational assessment task and the other involving a task assessing the potential of various wines to cause hangovers.

Young et al. (2000) suggested that two design features of patterning discriminations are critical for showing a positive patterning advantage. One feature is whether learning is measured in terms of response strength, rather than in terms of response choice. According to their simulations, some theories predict equivalent learning for positive and negative patterning when the response measure is one of choice between two alternatives. The second feature is whether conditions preclude the utility of an *opposites strategy*, in which participants give one response to single stimuli and the opposite response to the compound created by combining the single stimuli. This strategy takes advantage of the fact that most studies provide only two response alternatives (e.g., “no disease” or “disease”). In support of their proposal, Young et al. found that including a 50% noncontingent cue, which was either a single stimulus or a compound stimulus to which responses were sometimes reinforced and sometimes nonreinforced, led to a positive patterning advantage that was not seen without this cue. Young et al. argued that the noncontingent cue reduced the utility of an opposites strategy.

The experiments reported here were conducted in a way intended to address both of the design features identified by Young et al. (2000) as necessary to obtain a positive patterning advantage. First, participants responded with both a choice measure and a strength of association measure. According to Young et al.'s analysis, negative patterning should be harder than positive patterning when performance is assessed with the strength of association measure. Second, participants chose among three possible outcomes (neutral, beneficent, and maleficent), which was expected to mitigate the utility of an opposites strategy, because the outcome for one alternative would not predict the outcome for another alternative.

A Key Assumption

A final consideration needs to be mentioned. In the present experiments, I made the assumption, common to most prior studies of patterning in human causal reasoning, that describing an outcome as *neutral* or as having no effect corresponds to the absence of reinforcement in traditional animal learning designs. That is, in Bellingham et al. (1985), for example, a negative patterning discrimination involved presenting Cue A and Cue B followed by a reinforcer and presenting the compound AB followed by no reinforcer. In order to view the present experiments as investigations of positive and negative patterning, one must treat the neutral, or no effect, outcome as equivalent to the absence of a reinforcer. (This issue will be taken up further in the General Discussion section.)

EXPERIMENT 1
Health Reasoning

In Experiment 1, we embedded causal reasoning in a framework about food and health similar to those found in previous studies of causal learning (e.g., Shanks, Charles, et al., 1998). That is, participants were asked to assume the role of clinicians who, given information about the effects of different meals on a patient, try to determine how the patient reacted to various food items and combinations of food items. However, I incorporated a manipulation of outcome valence into Experiment 1 and included both conditions in which some meals led to a *beneficent* outcome (i.e., for some meals, the outcome was improved health, a desirable result of eating) and conditions in which some meals led to a *maleficent* outcome (i.e., for some meals the outcome was an allergic reaction, an undesirable result of eating). The outcomes most commonly used in prior research on causal reasoning, it may be noted, have been maleficent.

For each type of outcome, one set of materials was used to construct a negative patterning discrimination, and another set was used to construct a positive patterning discrimination (the sets of materials for one outcome were different from the sets for the other outcome). For the negative patterning discrimination, a compound was followed by a neutral outcome, and constituents of the compound were reinforced, with either a beneficent or maleficent outcome; for the positive patterning discrimination, constituents of a compound were followed by neutral outcomes, and the compound was reinforced with either a beneficent or maleficent outcome.

Having three response choices was expected to mitigate the utility of the kind of opposites strategy suggested by Shanks, Charles, et al. (1998) and Young et al. (2000). That is, knowing that one cue was followed by a particular outcome (e.g., no effect) would not help select the outcome for that cue when compounded with another, since the outcome might be better health or an allergic reaction. To the degree to which this mitigation was successful, it was expected that, as is predicted by current theories, positive patterning discriminations would be learned more readily than negative patterning discriminations.

To evaluate the claim of Young et al. (2000) that equivalent discriminations in positive and negative patterning would be obtained with a choice measure but not with a strength measure, learning was assessed with two kinds of response measures. On a trial-by-trial basis, participants chose among the three alternative outcomes (beneficent, neutral, maleficent), thereby providing a choice measure of learning. However, on a trial block basis (i.e., after a series of trials), they rated the associative strength of selected test stimuli with outcomes, thereby providing a strength measure. The participants also gave a final rating of associative strength with outcomes for all individual foods, all pairs, and all triads, after all study–test blocks were completed. Thus, it was possible to compare the relative difficulty of positive and negative patterning on both choice and strength measures.

One final feature of the design was intended to minimize the degree to which participants would try to solve the patterning tasks simply by responding in one way to any stimulus with a configural cue and responding in a different way to any stimulus without a configural cue. All meals contained at least two food items, and all of the patterning displays involved combinations of events. (This feature also served to make the meals somewhat more realistic, since no meal involved a single food item.) For example, the negative patterning discrimination involved a contrast between a nonreinforced compound of three items (ABC0) and its reinforced constituent pairs (AB+, BC+, and AC+) . A particular selection for the negative patterning task with a beneficent outcome might involve a meal consisting of *pork*, *peanuts*, and *peas* that was followed by a neutral outcome (no effect), and meals consisting of *pork–peanuts*, *peanuts–peas*, or *peas–pork* that were followed by improved health. Similarly, the positive patterning discrimination involved a contrast between a reinforced compound of three items (DEF+) and its nonreinforced constituent pairs (DE0, EF0, and DF0). For example, the positive patterning task with a beneficent outcome might involve a meal consisting of *lamb*, *tomatoes*, and *squash* that was followed by improved health, and meals consisting of *lamb–tomatoes*, *tomatoes–squash*, or *squash–lamb* that were followed by a neutral outcome. A similar set of comparisons was created for the negative and positive patterning task for the maleficent outcomes.

The two empirical questions of primary interest were, first, whether these discriminations would exhibit a positive patterning advantage, as is predicted by current theories and, second, whether any such advantage would depend on the valence of the outcomes. These questions were asked using a within-subjects design, which provided a direct comparison between the effects of beneficent outcomes and those of maleficent outcomes on patterning discriminations that was not confounded by potential differences in arousal or attentional focus.

Method

Participants. Thirty students, recruited from introductory psychology classes, participated in a session of approximately 1-h duration in partial satisfaction of a course requirement.

Materials. The stimulus materials consisted of a set of 12 food names, such as *lamb*, *spinach*, and *tomatoes*. During the study phases, the food names were presented as pairs or as triplets, with one above the other; during the test phases, items were presented singly, in pairs, or in triplets. The 12 food names were organized in four sets, with one set assigned to each discrimination problem for a given participant (see below). Across participants, the sets were assigned to conditions to ensure that each one was used in both positive and negative patterning and with both beneficent and maleficent outcomes.

Design. The primary manipulation of stimulus conditions involved a within-subjects, 2 × 2 factorial design combining, as one factor, the nature of the patterning discrimination (either positive or negative patterning) and, as the other factor, the nature of the reinforced outcome (either maleficent—an allergic reaction—or beneficent—improved health). This design generated four types of discrimination problems: negative patterning, maleficent outcome; positive patterning, maleficent outcome; negative patterning, beneficent outcome; and positive patterning, beneficent outcome. For all conditions, the nonreinforced outcome was designated as a *neutral* outcome.

The four discrimination problems were each represented by a different set of food items. Each set of food items consisted of the names of three foods, such as *lamb*, *spinach*, and *tomatoes*. To illustrate the structure of different presentations, consider the three foods to be A, B, and C. On compound trials, all three foods were presented as a triad (ABC), one above the other. On component trials, two foods were presented as a pair, also one above the other. Triads, such as ABC, are separable into three pairs (AB, BC, and AC), and these pairs were presented once each in every trial block. To balance the frequency of trials with reinforced and nonreinforced outcomes for a given discrimination problem, the triplets were presented three times each on every trial block.

A trial block consisted of 24 displays, with 6 displays from each discrimination problem. Half of the displays involved compound trials, on which a food triplet was shown, and half of the displays involved constituent trials, on which a pair of foods was shown. The order of the 24 displays was randomized separately for each participant at the start of each trial block

Procedure. On arrival in the laboratory, the participants were seated comfortably in front of the computer and given an instruction booklet that explained the nature of the tasks that they were to perform. After the participants read the instructions, the experimenter reviewed the task and answered any questions before starting the session.

During the session, the participants proceeded through a sequence of eight study–test cycles, followed by a final test block. The study period began with a brief review on the computer screen of the instructions for study, followed by the presentation of 24 displays characterized as "case studies." Each display included the first name of the patient (the patient was always the same person) and a list of items that constituted a meal. The display was accompanied by a request for the participants to predict the outcome of the meal on a 3-point scale (1, *improved health*; 2, *neutral*; 3, *allergic reaction*). As soon as a choice was made, the display changed to show the prediction and to indicate what outcome followed the meal, with feedback indicating "IMPROVED HEALTH!," "ALLERGIC REACTION!," or "neutral" being displayed for 3 sec.

Presentation of the case studies was followed by a block of 10 test trials. On each of these trials, the participants were asked to indicate the nature and strength of association between the outcome and a food item or combination of food items, using a scale ranging from −100 to +100. With this scale, a minus sign ("−") indicated an association with a maleficent outcome, and a plus sign ("+") indicated an association with a beneficent outcome; the number magnitude indicated the strength of the association.

The test block consisted of one test of the triplet from each of the four patterning problems, one test of a pair from each patterning problem, and a test of two individual food items. The selection of pairs and individual items alternated from one test block to the next.

After all eight study–test cycles were completed, a final test block of 28 tests was given, in which all food items, as well as all of the food pairs and food triplets presented in training, were each tested once for a measure of strength of association.

For the test blocks given after a study/prediction block and for the final test block given at the end, the order of test trials within a block was randomized at the start of the test block.

Results

Learning in this task was reflected in the outcome predictions made for each meal presented during a block of trials (on-line predictions that were a choice measure), in the ratings of associative strength given in test periods following each block of meal presentations (immediate off-line ratings that were a strength measure), and in the final series of ratings of associative strength given after all eight study–test blocks were completed (final off-line ratings that were also a strength measure). For all measures, the criterion for statistical significance was set at $\alpha = .05$.

Choice measures. Figure 1A depicts the mean prediction responses over trials for the positive and negative patterning tasks with the beneficent outcome (improved health). In the positive patterning task, the constituent pairs predicted a neutral outcome (DE0, EF0, DF0), and the compound triple predicted improved health (DEF+). In the negative patterning task, the constituent pairs predicted improved health (AB+, BC+, AC+), and the compound triple predicted a neutral outcome (ABC0). The *y*-axis is scaled so that predictions of a neutral outcome ("2") are at the bottom of the graph, and predictions of a beneficent outcome ("1") are at the top. Figure 1B depicts the corresponding data for the tasks with the maleficent outcome (the allergic reaction), but with the *y*-axis scaled so that predictions of a maleficent outcome ("3") are at the top of the graph, and predictions of a neutral outcome ("2") at the bottom.[2]

As can be seen in Figure 1A, positive and negative patterning discriminations with a beneficent outcome were both learned relatively quickly, with predictions for the nonreinforced displays in both discriminations remaining close to the neutral baseline of "2," whereas predictions for the reinforced displays in both moved toward a prediction of improved health (or "1"). According to Young et al. (2000), this equivalence of learning is predicted for choice measures by Pearce's (1994) configural model, for instance. However, a different picture of discrimination learning is shown in Figure 1B for the patterning tasks with the maleficent outcome. In these

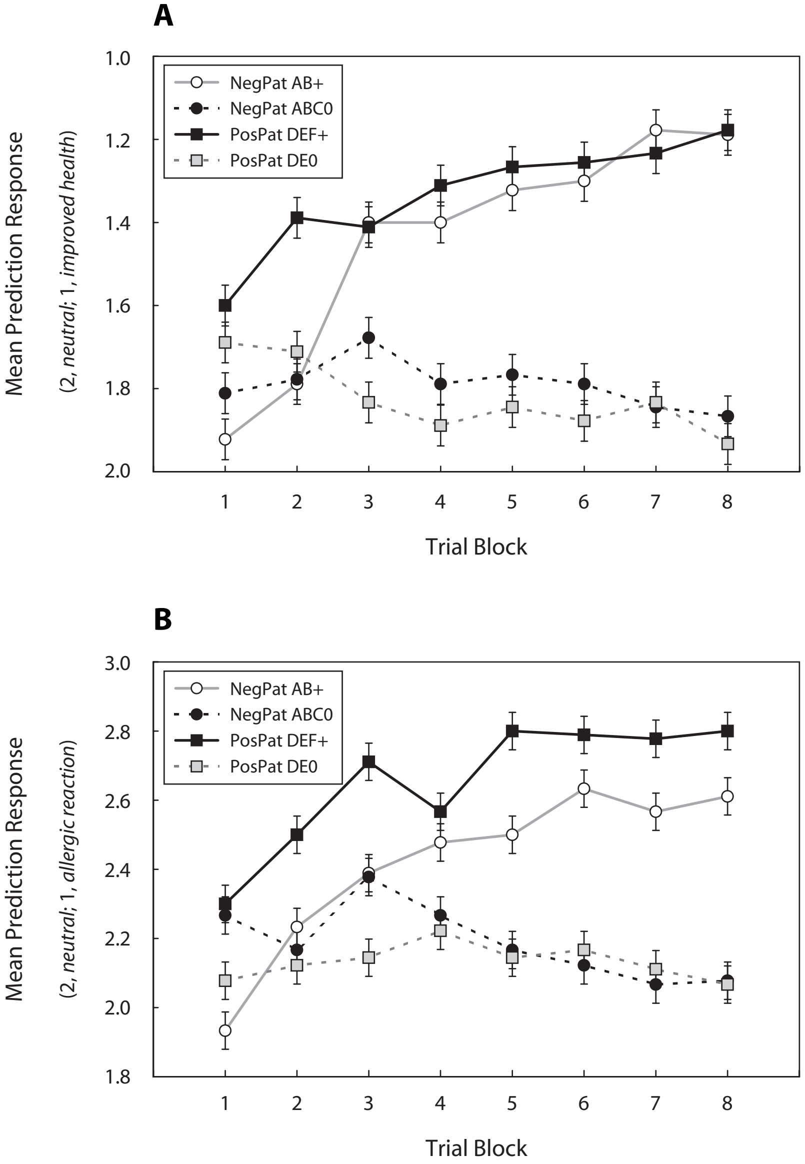

Figure 1. (A) Mean prediction responses in Experiment 1 as a function of trial block, shown separately for the positive (PosPat) and negative (NegPat) patterning tasks when the outcome was either beneficent (1) or neutral (2). The *y*-axis has been scaled so that predictions of the beneficent outcome are shown higher on the graph and predictions of the neutral outcome are shown lower. (B) Mean prediction responses in Experiment 1 as a function of trial block, shown separately for the positive (PosPat) and negative (NegPat) patterning tasks when the outcome was either maleficent (3) or neutral (2). Error bars show the standard errors of the means.

tasks, predictions for the nonreinforced displays also remained close to a neutral baseline of "2," but predictions for the reinforced displays were consistently closer to predictions of an allergic reaction (or "3") in the positive patterning discrimination than in the negative patterning discrimination.

For the purposes of analyzing the on-line prediction responses, data from the first trial, in which many responses were of necessity only guesses, were excluded. Data from each outcome condition were analyzed with $2 \times 2 \times 7$ repeated measures ANOVAs, with patterning task (positive or negative), cue reinforcement status (reinforced or nonreinforced), and trial as factors.

For the beneficent outcome, the ANOVA found significant effects of trial [$F(6,174) = 2.26$, $MS_e = 0.122$] and cue reinforcement status [$F(1,29) = 95.96$, $MS_e = 0.518$] and a significant interaction between these two factors [$F(6,174) = 12.11$, $MS_e = 0.086$], reflecting the increasing difference between reinforced and nonreinforced cues over trials. The nature of the patterning discrimination did not produce a main effect [$F(1,29) < 1$] or any significant interactions, except with trial [$F(6,174) = 3.63$, $MS_e = 0.086$], reflecting the crossover between Trial Blocks 2 and 3 in performance on the negative patterning discrimination.

For the maleficent outcome, the ANOVA also showed significant effects of trial [$F(6,174) = 2.56$, $MS_e = 0.148$] and cue reinforcement status [$F(1,29) = 74.36$, $MS_e = 0.542$] and a significant interaction [$F(6,174) = 11.09$, $MS_e = 0.078$], again reflecting the improved discrimination over trials. However, in the case of discriminations with the maleficent outcome, there was both a main effect of patterning task [$F(1,29) = 4.28$, $MS_e = 0.402$] and, more importantly, a significant interaction between patterning task and cue reinforcement status [$F(1,29) = 22.83$, $MS_e = 0.152$], reflecting the fact that the positive patterning discrimination was acquired significantly better than the negative patterning discrimination was.

An overall ANOVA that combined the factor of outcome with those of patterning task, cue reinforcement status, and trial found a significant interaction between patterning task and cue reinforcement status [$F(1,29) = 19.26$, $MS_e = 0.209$], reflecting an overall positive patterning advantage, in addition to the expected effects of trial and cue reinforcement status and their interaction. The overall ANOVA did not show the three-way interaction of outcome, patterning task, and cue reinforcement status to be significant [$F(1,29) = 1.79$, $MS_e = 0.219$, $p = .191$], however. Thus, strong evidence for the role of outcome valence on the difficulty of patterning discriminations in the prediction responses was lacking.

Associative strength ratings. Figure 2A shows mean ratings of associative strength for reinforced and nonreinforced cues in positive patterning and negative patterning for the beneficent outcome condition, with data averaged across two-trial blocks (the data were averaged in two-trial blocks for these ratings because only the tests across two blocks contained a full set of comparisons). Figure 2B displays data for the maleficent outcome condition in the same format, but with the *y*-axis rescaled so that greater association with the maleficent outcome is represented by higher scores, to better correspond to the data for the beneficent condition. (The participants indicated association with the maleficent outcome by using negative values.)

Consistent with the results for prediction responses, the association rating data show that both patterning tasks were mastered for each outcome, but the relative ease of positive and negative patterning depended on whether the outcome was beneficent or maleficent.

Discrimination learning in the negative patterning task was equal to that in the positive patterning task when the outcome was beneficent, but such learning was poorer for negative than for positive patterning when the outcome was maleficent.

Association rating data for the beneficent condition were analyzed using a $2 \times 2 \times 4$ ANOVA with patterning task (positive or negative), cue reinforcement status (reinforced or nonreinforced), and trial block as factors. This analysis showed significant effects of trial block [$F(3,87) = 2.81$, $MS_e = 1,027.7$] and of cue reinforcement status [$F(1,29) = 27.38$, $MS_e = 4,343.0$] and a significant interaction of the two factors [$F(3,87) = 6.92$, $MS_e = 928.9$]. These results confirmed the emergence of discriminative responding in the association ratings across trial blocks. However, the analysis showed no effect of the type of patterning task, with all *F* ratios involving patterning task being less than 1. A similar analysis for the maleficent condition also showed significant effects of trial block [$F(3,87) = 4.81$, $MS_e = 1,352.7$] and cue reinforcement status [$F(1,29) = 24.57$, $MS_e = 5,190.3$] and a significant interaction of these two factors [$F(3,87) = 8.89$, $MS_e = 915.4$]. However, in the case of the maleficent condition, there was also a significant interaction between patterning task and cue reinforcement status [$F(1,29) = 13.38$, $MS_e = 1,915.6$], reflecting the diminished learning to the reinforced cue in the negative patterning relative to the positive patterning condition.

To compare the two outcome conditions directly, a $2 \times 2 \times 2 \times 4$ ANOVA was carried out on data from the two conditions together, after the association ratings for the maleficent condition were multiplied by -1 to create a common scale for the two conditions. As with the prediction responses, the overall ANOVA showed a significant interaction between patterning task and cue reinforcement status [$F(1,29) = 9.79$, $MS_e = 1,428.5$], reflecting an overall positive patterning advantage. However, this analysis confirmed the differential learning of the two patterning tasks as a function of outcome valence, with a significant three-way interaction involving outcome, patterning task, and cue reinforcement status [$F(1,29) = 7.94$, $MS_e = 1,473.5$, $p = .009$], reflecting the observation that the positive patterning advantage was seen only for the maleficent outcome. The nature of the outcome was also associated with a main effect [$F(1,29) = 14.12$, $MS_e = 11,389.5$]; the rescaled ratings were higher for the beneficent than the maleficent outcomes.

Associative strength ratings: Final test. Mean association ratings from the final test series, given after all of the study–test blocks were completed, are depicted on the right-hand side of Figures 2A and 2B and also summarized in Table 1.[3]

As can be seen in Figures 2A and 2B, the participants demonstrated mastery of all the patterning discriminations by the time of the final association test. Mean ratings for stimuli associated with beneficent or maleficent outcomes were clearly different from the mean ratings for stimuli associated with neutral outcomes. In this final test, mean ratings for neutral items were nearly identical for the positive and negative patterning discriminations, in both outcome conditions, whereas ratings for reinforced items showed a slight discriminative advantage for positive over negative patterning tasks.

Table 1 also shows the same data with the addition of ratings for single food items in each condition. The table displays the mean ratings in different typefaces to indicate which comparisons were significantly different within each combination of outcome condition and test stimulus. The statistical analysis consisted of a $2 \times 2 \times 3$ ANOVA with outcome (beneficent or maleficent), patterning task (negative or positive), and test stimulus (single, pair, or triplet) as factors, followed by comparisons between individual means within each patterning task, using the pooled mean square error term from the outcome $\times$ patterning $\times$

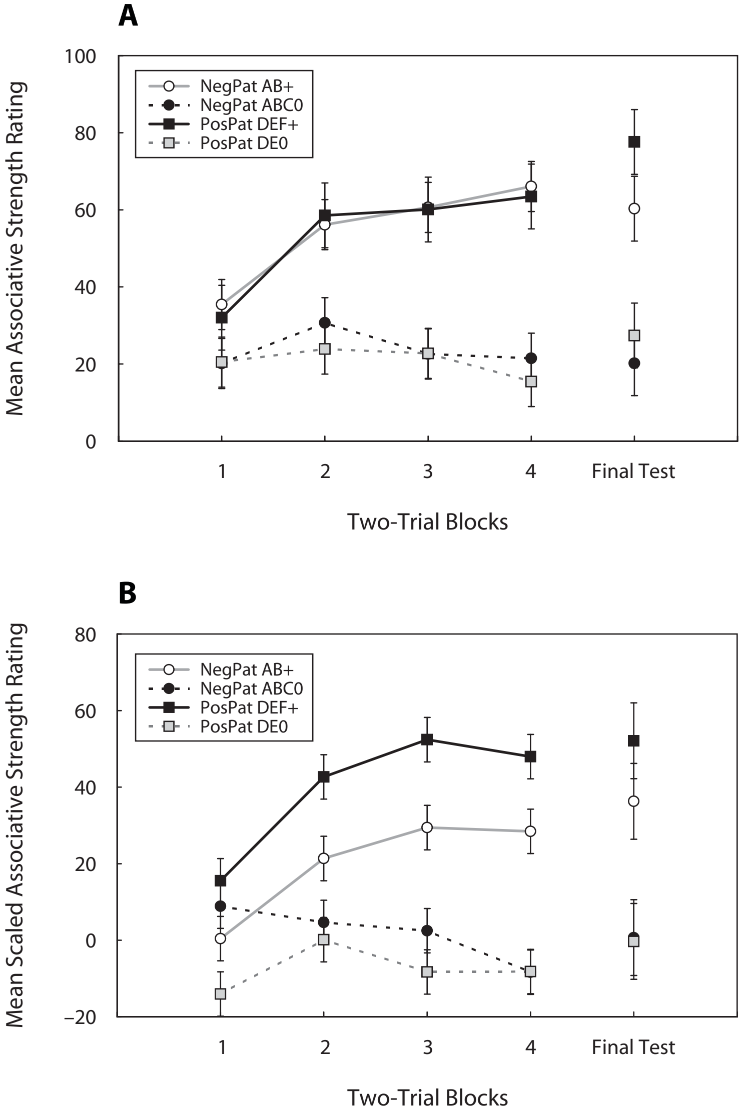

Figure 2. (A) Mean association ratings in Experiment 1 as a function of two-trial blocks for items paired with beneficent and neutral outcomes, shown separately for positive (PosPat) and negative (NegPat) patterning tasks. Data from the final association-rating task are shown on the right. (B) Mean association ratings in Experiment 1 as a function of two-trial blocks for items paired with maleficent and neutral outcomes, shown separately for positive (PosPat) and negative (NegPat) patterning tasks. Ratings have been multiplied by −1, so that the data are scaled in the same direction as in panel A. Data from the final association-rating task are shown on the right. Error bars show the standard errors of the means.

test stimulus interaction in the ANOVA to calculate a critical difference of 13.7.

For the beneficent outcome conditions, in the negative patterning task, the mean rating of the pair (AB) was significantly greater than the mean ratings for either the triple (ABC), or the single (A), and that for the single was significantly greater than that for the triple. In the positive patterning task, the mean rating for the triplet (DEF) was significantly greater than the ratings for either the single (D) or the pair (DE), which did not differ from each other. For the maleficent conditions, in the negative patterning task, the mean rating of the pair (AB) was significantly lower than the ratings for either the element (A) or the triplet (ABC), which did not differ, and in the positive patterning task, the mean rating of the triplet (DEF) was significantly lower than the mean ratings of either the single (D) or the pair (DE), which did not differ.

Discussion

The results of Experiment 1 contribute two important observations. First, in contrast to those of a number of prior studies, they clearly show a positive patterning advantage

Table 1
Mean Ratings of Associative Strength on the Final Association Test in Experiment 1 for Singles (A, D), Pairs (AB, DE), and Triplets (ABC, DEF) for the Positive and Negative Patterning Conditions

Test Stimulus	Beneficent Outcome	Maleficent Outcome
Negative Patterning		
A	**43.3**	−2.6
AB	**60.3**	**−36.3**
ABC	20.2	−0.7
Positive Patterning		
D	33.3	3.6
DE	27.4	0.3
DEF	**77.6**	**−52.1**

Note—Within each set of three, values in bold are significantly different from those not in bold. The means for the singles (A, D) and pairs (AB, DE) are based on 81 observations; those for the triplets (ABC, DEF) are based on 27 observations.

in human causal reasoning. Second, they also clearly show that the conditions for producing a positive patterning advantage are more complex than the set of design features identified by Young et al. (2000) for human causal-reasoning tasks. Specifically, the results suggest that the relative ease of learning positive and negative patterning depends partly on the type of outcome used in the discrimination. In a task concerned with assessing causal efficacy of different food items, a beneficent outcome (improved health) supported similar levels of learning in positive and negative patterning discriminations, whereas a maleficent outcome (allergic reaction) was associated with better learning in a positive than in a negative patterning task.

This finding of similar task difficulty for the two patterning tasks when the discrimination involved a beneficent outcome has no ready interpretation in terms of an opposites strategy. For one thing, the availability of three alternative responses was intended to make the role of an opposites strategy less plausible. More important, however, an opposites strategy is ruled out by the fact that positive patterning was learned more readily than negative patterning when the discrimination involved a maleficent outcome. If the participants used an opposites strategy to perform equally well in positive and negative patterning when the outcome was beneficent, one might expect them to also do so when the outcome was maleficent. Thus, it appears that use of an opposites strategy does not provide a general account of discrimination performance in patterning tasks, a conclusion reached for different reasons by Lachnit and his colleagues (e.g., Lachnit & Lober, 2001; Ludwig & Lachnit, 2003).

The results also offer no support for the idea that choice measures would be more likely than associative strength measures to show equivalent performance for positive and negative patterning. Instead, the choice measures and associative strength measures were in general agreement, showing an overall positive patterning advantage. Both measures also showed little difference between the two kinds of patterning tasks when the outcome was beneficent but a clear positive patterning advantage when the outcome was maleficent, although this interaction was only significant for the strength measure. Of course, one should be cautious about interpreting the equivalence of the behavioral measures without a developed theory of response generation for this task. The fact that the associative ratings were described to the participants as measures of associative strength does not mean that the underlying basis for their responding might nonetheless have been a choice process. Still, the results do show that the conditions needed to produce equivalent performance in positive and negative patterning involve more than the nature of the nominal response measure.

Finding that the relative difficulty of negative and positive patterning depends partly on outcome valence seems consistent with other research on patterning discriminations. That is, in many of the studies in which the positive patterning advantage predicted by current theory (e.g., Bellingham et al., 1985, Experiment 1; Deisig et al., 2001; Rescorla, 1972; Woodbury, 1943) was not found, outcomes that would here be termed *beneficent* were used; conversely, in most of the studies in which a clear advantage was shown of positive over negative patterning, outcomes that would be here termed *maleficent* were used.

In any case, one general conclusion supported by Experiment 1 is that researchers of human reasoning might be well advised to use beneficent as well as maleficent outcomes. However, it seemed reasonable to examine the generality of the effect of outcome valence in human causal reasoning before adopting such a general conclusion. Thus, a second experiment was conducted using a social-reasoning task that was based on a different explanatory framework.

EXPERIMENT 2
Social Reasoning

In Experiment 2, I replicated the design of the first experiment using a social-reasoning task in which individuals are assessed by a group. This task relies on the fact that people have causal power in social settings; for example, the individual members of a group each exert influence on decisions reached by the group. In other domains (e.g., Cheng & Holyoak, 1985; Cosmides, 1989), reasoning tasks that invoke social reasoning skills have often revealed types of judgments different from those seen with nonsocial tasks. Thus, it seemed reasonable to examine the role of the same design features studied in Experiment 1 in a task that tapped into social understanding.

Method

The method was the same as that used in Experiment 1, except as is noted. Thirty-five students from introductory psychology classes participated in a session of about 50 min in partial fulfillment of a course requirement.

The design of Experiment 2 was identical to that of the first experiment, but the orienting framework was different. In this experiment, participants were instructed that their task was to determine, from group assessments of a target individual, what sort of assessments of the target individual were held by group members, both individually and in combination with other group members. Assessments could be either *positive*, meaning that the group or group members had a positive assessment of the target person, *neutral*, or *negative*, meaning that the group or group members had a negative assessment of the target person. The positive assessments were treated as beneficent outcomes; the negative ones, as maleficent outcomes.

Eight study–test blocks were presented, with 24 displays in every study block. Each display included the first name of a target individual (the target was the same person for all displays) and a list of names that comprised a council. The display was accompanied by a request for the participants to predict the judgment of the council about the target (1, *positive*; 2, *neutral*; 3, *negative*). As soon as a choice was made, the display changed to show the prediction and to indicate the judgment of the council, with feedback indicating "POSITIVE!," "NEGATIVE!," or "neutral" being displayed for 3 sec.

Results and Discussion

As in the first experiment, the results will be summarized separately for the prediction responses, which provide a choice measure, and the association ratings, which provide a strength measure, both after each block and, in the final association ratings, after all blocks. A criterion of $\alpha = .05$ was used to determine statistical significance for all measures.

Choice measures. Using the same format as in Figures 1A and 1B, Figure 3A shows the mean prediction responses for the positive and negative patterning tasks with the beneficent outcome (positive assessment), and Figure 3B shows mean prediction responses for these tasks with the maleficent outcome (negative assessment).

It is readily apparent from comparing Figures 3A and 3B with Figures 1A and 1B that learning was more difficult in the social-reasoning task than in the health-reasoning task. The separation between prediction responses for cues followed by an identified outcome and responses for cues

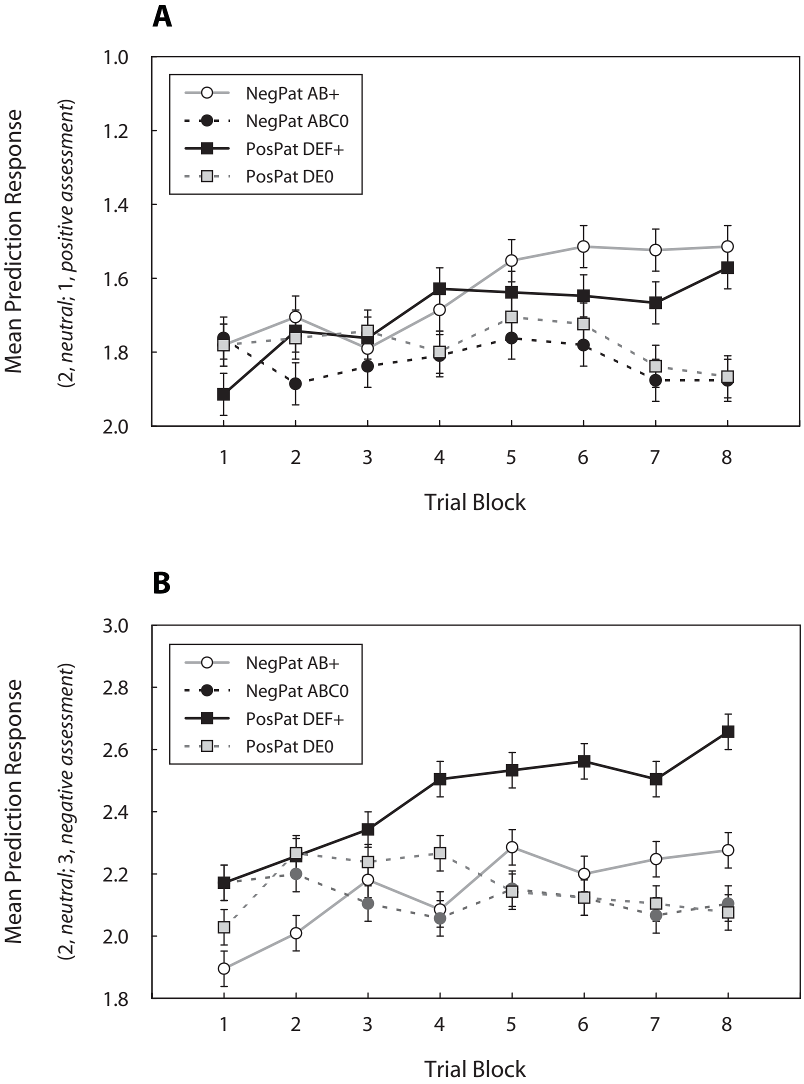

Figure 3. (A) Mean prediction responses in Experiment 2 as a function of trial block, shown separately for the positive (PosPat) and negative (NegPat) patterning tasks when the outcome was either beneficent (1) or neutral (2). The *y*-axis has been scaled so that predictions of the beneficent outcome are shown higher on the graph and predictions of the neutral outcome are shown lower. (B) Mean prediction responses in Experiment 2 as a function of trial block, shown separately for the positive (PosPat) and negative (NegPat) patterning tasks when the outcome was either maleficent (3) or neutral (2). Error bars show the standard errors of the means.

followed by a neutral outcome did not occur as rapidly or by as large an amount as was found in Experiment 1.

Nonetheless, it is clear that the participants learned the discriminations for both beneficent and maleficent outcomes. Interestingly, the negative patterning task appeared to be learned *more easily* than the positive patterning task when the outcome was beneficent. The results when the outcome was maleficent, in contrast, were similar to those seen with food items—namely, that the positive patterning task was learned more quickly than the negative patterning task.

The on-line prediction responses were analyzed as in Experiment 1, with data from the first trial excluded, using a $2 \times 2 \times 7$ repeated measures ANOVA, with patterning task (positive or negative), cue reinforcement status (reinforced or nonreinforced), and trial as factors.

For the beneficent outcome, the ANOVA showed a significant effect of cue reinforcement status [$F(1,34) = 18.97$, $MS_e = 0.356$], indicating that reinforced cues were given higher responses than nonreinforced ones. Cue reinforcement status did interact with patterning task [$F(1,34) = 4.76$, $MS_e = 0.153$], which reflected the fact of a greater difference between reinforced and nonreinforced cues for the negative patterning than for the positive patterning task. No other effects of patterning were significant (all Fs < 1).

For the maleficent outcome, the ANOVA showed an effect of cue reinforcement status [$F(1,34) = 26.14$, $MS_e = 0.328$] and an interaction with trial [$F(6,204) = 6.36$, $MS_e = 0.149$], indicating the clear improvement over trials in the participants' correctly predicting which meals were associated with the maleficent outcome (allergic reactions). There was also a main effect of patterning task [$F(1,34) = 12.81$, $MS_e = 0.603$] and, more importantly, a significant interaction between patterning task and cue reinforcement status [$F(1,34) = 12.54$, $MS_e = 0.277$]. The latter effects reflected the clear advantage for learning about meal outcomes in the positive patterning task compared with the negative patterning task.

As with the data from Experiment 1, an overall ANOVA was done in which data from both outcome conditions were combined, with prediction responses rescaled so that for both conditions, higher values represented predictions of the valenced outcome. This analysis confirmed the differential learning of the two patterning tasks as a function of outcome valence, yielding a significant three-way interaction involving outcome, patterning task, and cue reinforcement status [$F(1,34) = 21.22$, $MS_e = 0.226$, $p = .0005$]. This analysis also showed a main effect for patterning task [$F(1,34) = 6.84$, $MS_e = 0.494$] and an interaction between patterning and outcome [$F(1,34) = 10.30$, $MS_e = 0.402$], with both effects reflecting the lower mean rating in negative patterning tasks involving maleficent outcomes.

Associative ratings. Figures 4A and 4B show the mean ratings of associative strength for reinforced and nonreinforced cues in the positive and negative patterning conditions, with Figure 4A depicting results in the beneficent outcome condition and Figure 4B depicting the maleficent outcome condition. Consistent with the results for prediction responses, the association rating data show that, although both patterning tasks were mastered for each outcome, the relative ease of positive and negative patterning depended on whether the outcome was beneficent or maleficent. Discrimination learning in the negative patterning task was roughly equal to that in the positive patterning task when the outcome was beneficent, but it was much slower for negative than positive patterning when the outcome was maleficent.

ANOVAs on each outcome condition separately showed, for the beneficent outcome, a significant effect of trial [$F(3,102) = 5.46$, $MS_e = 1{,}259.98$] and cue reinforcement status [$F(1,34) = 4.12$, $MS_e = 2{,}456.04$], but there was no significant interaction between patterning and cue reinforcement status [$F(1,34) = 1.71$, $p > .20$]. For the maleficent outcome, the corresponding analysis also showed significant effects of trial [$F(3,102) = 5.99$, $MS_e = 1{,}560.59$] and cue reinforcement status [$F(1,34) = 8.70$, $MS_e = 2{,}491.27$]. However, this analysis also showed a significant interaction between patterning task and cue reinforcement status [$F(1,34) = 6.39$, $MS_e = 1{,}506.62$].

A $2 \times 2 \times 2 \times 4$ ANOVA was done on data from both conditions together (with association ratings for the maleficent condition multiplied by -1 to create a common scale). This analysis, like that for Experiment 1, confirmed an overall positive patterning advantage, with a significant interaction between patterning task and cue reinforcement status [$F(1,34) = 6.22$, $MS_e = 1{,}641.5$], but it did not show a significant three-way interaction involving outcome, patterning task, and cue reinforcement status [$F(1,34) = 1.37$, $MS_e = 1{,}036.27$, $p = .250$], presumably because of the low levels of learning in the beneficent condition.

Association ratings: Final test. Mean association ratings from the final test series, given after all study–test blocks were completed, are depicted in Figures 4A and 4B and are also summarized in Table 2, in the same manner as was done for Experiment 1. As can be seen in Figure 4A, the two types of patterning tasks showed similar levels of mastery in the final association ratings when the outcome was beneficent. However, as can be seen in Figure 4B, only the positive patterning task showed strong evidence of mastery when the outcomes were maleficent.[4]

Analyses of the final associative ratings were done in the same manner as before, using an ANOVA with outcome (beneficent or maleficent), patterning task (negative or positive), and test stimulus (single, pair, or triplet) as factors to estimate a critical difference of 12.7 for comparisons within each patterning task. The resulting picture of differences among the trial types was slightly different from what was found in Experiment 1. With the beneficent outcome conditions, the negative patterning task had a mean rating for the pair (AB) that was significantly greater than that for the triplet (ABC) but, unlike in Experiment 1, not for the single (A), although the rating for the single was also significantly greater than that for the triplet. The ratings in the positive patterning task resembled those in Experiment 1: The mean rating for the triplet (DEF) was significantly greater than the ratings for both the single (D) and the pair (DE), which did not differ. With

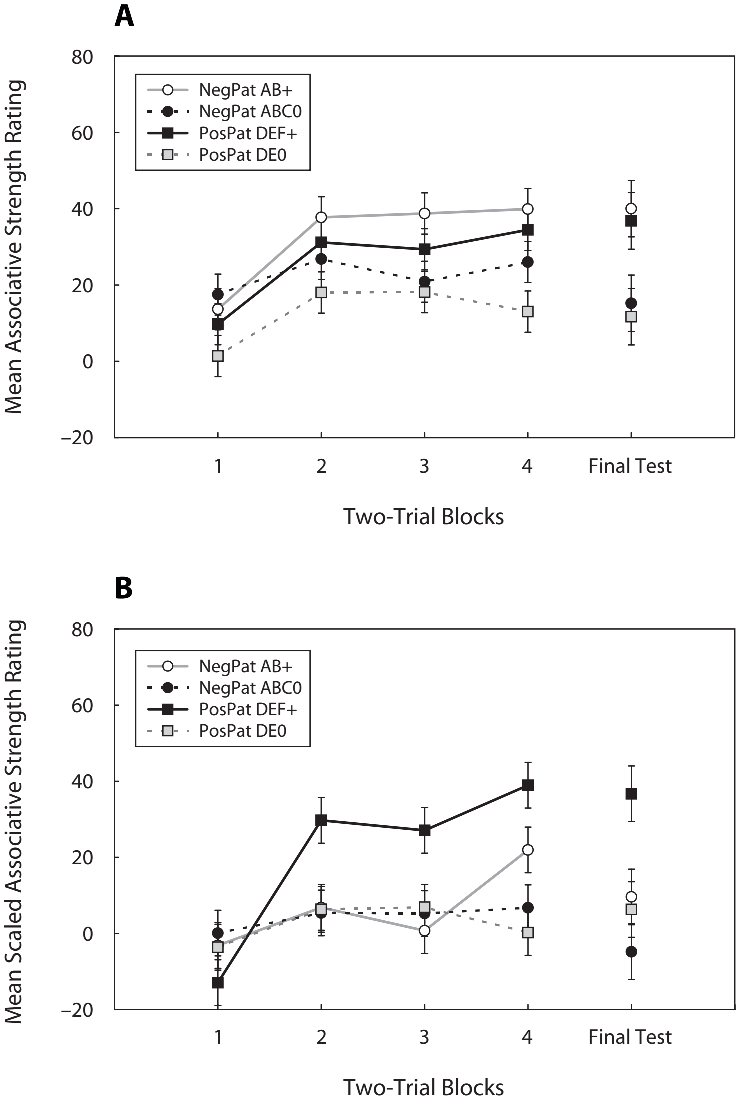

Figure 4. (A) Mean association ratings in Experiment 2 as a function of two-trial blocks for items paired with beneficent and neutral outcomes, shown separately for positive (PosPat) and negative (NegPat) patterning tasks. Data from the final association-rating task shown on the right. (B) Mean association ratings in Experiment 2 as a function of two-trial blocks for items paired with maleficent and neutral outcomes, shown separately for positive (PosPat) and negative (NegPat) patterning tasks. Ratings have been multiplied by −1 so that the data are scaled in the same direction as in panel A. Data from the final association-rating task are shown on the right. Error bars show the standard errors of the means.

the maleficent conditions, the negative patterning task had a mean rating for the pair (AB) that was significantly different from the mean rating for the triplet (ABC), as was found in Experiment 1, but not from that for the single element (A), which also did not differ from that for the triplet. Indeed, as is evident in Figure 4B, the three test stimuli received very similar ratings. In contrast, results for the positive patterning task clearly replicated that seen in Experiment 1, with the mean rating of the triplet (DEF) being significantly lower than the ratings of either the single (D) or the pair (DE), which did not differ.

Discussion

The results of Experiment 2, in which a different framing scenario was used, generally confirmed the findings of Experiment 1. In particular, the results showed a positive patterning advantage for both prediction responses and association ratings when the outcome was maleficent and a different picture when the outcome was beneficent.

One difference between the experiments was that in Experiment 2, the role of outcome valence was significant in an overall analysis of the choice measure (the on-line prediction responses) but not in the overall analysis of the strength

Table 2
Mean Ratings of Associative Strength on the Final Association Test in Experiment 2 for Singles (A, D), Pairs (AB, DE), and Triplets (ABC, DEF) for the Positive and Negative Patterning Conditions

Test Stimulus	Beneficent Outcome	Maleficent Outcome
Negative Patterning		
A	***30.0***	*−1.7*
AB	***40.0***	***−9.6***
ABC	15.2	4.8
Positive Patterning		
D	11.2	−6.4
DE	11.7	−6.3
DEF	**36.8**	**−36.7**

Note—Within each set of three, values in bold are significantly different from those that are in plain font but not from those in italics; however, values in italics are not significantly different from each other. The means for the singles (A, D) and pairs (AB, DE) are based on 87 observations; those for the triplets (ABC, DEF) are based on 29 observations.

measure (the association ratings). The essential statistical support for this claim—a three-way interaction among outcome, patterning task, and cue reinforcement status—was not significant for prediction responses in Experiment 1, but it was significant in that experiment for association ratings. Whether this difference is meaningful requires further study. A not unreasonable view is that the two experiments show the same trends, with the differences between the two measures in terms of which shows significant effects attributable to sampling error. On this view, it would make sense to combine the separate probabilities from each experiment to get an overall assessment for prediction responses and associative ratings. Doing so gives $z = 2.91$ ($p < .01$) for the prediction responses and $z = 2.15$ ($p < .02$) for the association ratings. On the other hand, there may be meaningful differences between the two task-orienting frameworks with respect to whether prediction responses or associative ratings are the better measure. Nevertheless, it is clear that for each framework, outcome valence interacts with the difficulty of patterning on at least one measure.

Another difference between the experiments was that the associative ratings in Experiment 2 showed only limited mastery of the negative patterning discrimination in the maleficent condition, despite clear evidence for mastery of the positive patterning discrimination. In Experiment 1, associative ratings in the poststudy blocks showed mastery, albeit at a lower level than for positive patterning, and showed similar levels of performance in the final associative ratings.

The results from the two experiments taken together suggests that the effect of outcome valence could be on both the asymptotes and the rates of learning the patterning discriminations. That is, the results of Experiment 1 showed relatively rapid learning, with differences between positive and negative patterning apparent during the learning process. When tested at the end of the session, however, the participants showed clear evidence of learning both patterning tasks. The final association ratings suggested that negative patterning might have a lower learning asymptote than positive patterning, but none of the observed differences were reliable. By way of contrast, the results of Experiment 2 showed much slower learning, which did not appear to have reached asymptote by the end of the training session, with the result that the data from the final session still showed an advantage for positive over negative patterning in the case of maleficent outcomes.

GENERAL DISCUSSION

The experiments reported in this article make two notable contributions to the literature on human causal reasoning. First, they demonstrate a robust positive patterning advantage with both a choice response measure and an associative strength measure, extending the findings of Young et al. (2000) with respect both to response measures and to procedures. Second, they show that this positive patterning advantage can be moderated by the valence of the outcome. The positive patterning advantage was consistently found for discriminations involving maleficent outcomes. Indeed, for maleficent outcomes under the conditions of measurement used for the association rating task in Experiment 2, the participants showed limited evidence of learning a negative patterning discrimination, despite clear evidence of learning a positive patterning discrimination. However, with beneficent outcomes, the advantage of positive over negative patterning was significantly diminished.

These results encourage renewed consideration of current accounts of patterning, and not only for the case of causal reasoning. The common prediction of most associative learning theories is that positive patterning should be easier than negative patterning, but this prediction was confirmed in the present experiments only when the outcomes were maleficent. The lack of agreement between theoretical prediction and the results for beneficent outcomes is a distinctive finding. However, it seems best viewed as part of a larger set of findings from a variety of paradigms in which the predicted advantage of positive over negative patterning either has not been found or has been very small. An important task for theoretical development is to specify the boundary conditions over which a positive patterning advantage holds. Of particular interest in relation to the present research, of course, is whether effects of outcome valence can be demonstrated in tasks other than human causal reasoning.

The results also support two conclusions regarding methodological considerations in the causal-reasoning task. In terms of response measurement, the results were generally similar whether causal reasoning was assessed by a choice response or by an associative rating scale. This similarity is encouraging, given the heterogeneity of measures currently in use. Nonetheless, continued attention to response measures seems warranted, given the demonstration by Young et al. (2000) that predictions from exemplar models differ depending on whether the behavioral measure is a choice response or a rating response. The need to attend to possible differences between response measures is also reinforced by the apparent negative patterning advantage observed with the choice measure in Experiment 2.

In terms of the scenarios used to study human causal reasoning, the present results suggest that investigators would do well to look more extensively at the way people learn about the causes of beneficent outcomes. With few exceptions (e.g., Le Pelley et al., 2005), the recent literature on causal reasoning in humans has been based on situations that involve maleficent outcomes, such as adverse reactions to food (e.g., Larkin, Aitken, & Dickinson, 1998; Shanks, Charles, et al., 1998; Shanks, Darby, & Charles, 1998), allergic reactions to medicine (e.g., Matute, Arcediano, & Miller, 1996), detection or diagnosis of disease (e.g., Shanks, 1991), destruction of tanks (e.g., De Houwer, Becker, & Glautier, 2002), process failures in industrial plants (e.g., Cobos, López, Caño, Almaraz, & Shanks, 2002; Waldmann, 2001), and anticipation of electric shock (Lachnit & Kimmel, 1993; Mitchell & Lovibond, 2002). This focus on maleficent outcomes may distort our characterization of how people learn about complex causal scenarios. It is worth noting, for example, that Fredrickson and Branigan (2005) argued that positive affective states support different types of attentional processes than do negative affective states. Fredrickson's (2001) broaden-and-build theory of positive emotions claims that positive emotions lead to a broadening of attentional focus and a greater inclusiveness of thought–action tendencies. To the extent that beneficent outcomes are linked to the elicitation of positive emotions and maleficent outcomes are linked to negative emotions, one might expect to find differences in learning for the two types of outcome, especially for complex tasks in which attentional focus is likely to matter.

Why Should Outcome Valence Matter? Possible Explanations

The results found here pose an obvious theoretical challenge of explaining why the valence of the outcome should alter the relative difficulty of patterning discriminations or, at least, appear to alter their relative difficulty. At present, one can only offer speculative suggestions, but two different approaches to the problem suggest themselves (one of which suggested itself to an anonymous reviewer).

Approach 1: Direct Application of Associative Theorizing

The design of Experiments 1 and 2 assumes an isomorphism between the human causal-reasoning task and animal conditioning paradigms. Specifically, it assumes that a display followed by a neutral outcome is equivalent to a conditioning trial in which the stimuli are nonreinforced, meaning that no outcome occurred. With this identification, the positive and negative patterning tasks are clearly differentiated. As was described in the introduction, negative patterning trials involve presenting constituent events with an outcome (e.g., A+, B+) and a compound of those constituents with no outcome (AB0), whereas positive patterning trials consist of presenting constituent stimuli without an outcome (e.g., A0, B0) and a compound of those constituents with an outcome (AB+). All other events (e.g., the intertrial events) also occur without an outcome, making the two patterning discriminations distinct.

This way of describing the events in a human causal-reasoning task is characteristic of much of the literature that has drawn on theoretical approaches from animal conditioning studies. With this approach, there are three possible modifications to existing theory that could accommodate the findings reported here on the effects of outcome valence: retrospective attention to configural cues, retrospective reweighting of cue salience, and differential processing dynamics.

Retrospective attention to configural cues. One possibility is that beneficent outcomes lead to a retrospective revaluation of the cues present at the time of an experience, such that configural cues are given more prominence on the occasion of a beneficent outcome, for example. (Equivalently, simple cues could be given more prominence on the occasion of a maleficent outcome.) This could be viewed as a variation of the retrospective evaluation suggested by Kamin (1969) in his classic account of blocking.

Retrospective reweighting of cue salience. A second possibility is that beneficent outcomes lead to a retrospective reweighting of weak cues relative to strong cues. That is, rather than focusing attention on configural cues, beneficent outcomes might give a greater boost to weak cues than to stronger cues. In terms of Harris's (2006) attentional buffer model, for example, one might suggest that the attentional boost given to a cue was dependent on the salience of the cue, with weaker cues getting more of an attentional boost.

Differential processing dynamics. A third possibility picks up on the proposal of Krane and Wagner (1975) that there may be advantages for certain cue–outcome combinations because of their greater similarity in regard to the temporal dynamics of processing. Specifically, suppose that configural cues take longer to be fully processed than simple cues, as is suggested by the optimal ISI data of Kinder and Lachnit (2002), and that beneficent outcomes take longer than maleficent outcomes to be processed fully. Then there may be a special advantage for configural cues to develop associations with beneficent outcomes.

Approach 2: Recasting the Nature of the Problems

If one does not accept the isomorphism of the causal-reasoning task with animal conditioning preparations, the identification of these problems as positive and negative patterning can be problematic. A particular concern arises if the outcomes are designated Outcome 1 and Outcome 2, rather than as reinforcement and nonreinforcement. With such a designation, the negative patterning task can be represented as $A \rightarrow 1$, $B \rightarrow 1$, and $AB \rightarrow 2$, and the positive patterning task as $A \rightarrow 2$, $B \rightarrow 2$, and $AB \rightarrow 1$. These are obviously equivalent problems, and there is no reason to expect negative patterning to be more difficult than positive patterning.

The potential equivalence of positive and negative patterning that occurs if the outcomes are interchangeable may account for some of the failures to find a positive patterning advantage in earlier studies of patterning in causal reasoning (e.g., Harris & Livesey, 2008; Shanks, Charles, et al., 1998; Shanks & Darby, 1998). The analysis does not seem immediately applicable to the present experiments, because there was, in fact, a positive patterning advantage in the maleficent conditions, which implies that the problems were not equivalent. However, a more general point raised by such an analysis is that participants may not see the problem in the same way that the experimenter has characterized them. This point was made cogently by an anonymous reviewer, who proposed an account along the following lines:

Suppose that for the health-reasoning task, participants viewed the outcomes of *better health* and *no effect* as equally salient, but viewed the *allergic reaction* as more salient than either of them. In this case, the beneficent outcome condition might have involved learning that some cues were associated with one outcome and other cues were associated with another, equally salient outcome, producing the kind of problem equivalence between positive and negative patterning described above. The maleficent condition, in contrast, would have involved learning to associate some cues with a very salient outcome (allergic reaction) and to associate other cues with its absence, thereby producing a parallel with the presence and absence of reinforcement in animal conditioning studies, and, correspondingly, a positive patterning advantage.

A related account can be offered for the social-reasoning task, with the change that a beneficent outcome was viewed as an expected (and therefore less salient) outcome, whereas the neutral outcome was unexpected (and therefore more salient), and the maleficent outcome was even more unexpected (and even more salient). In this case, the experienced situation for the beneficent condition would be that some cues are associated with a salient outcome (neutral reaction) and other cues are associated with its absence (liking), leading the participants to experience the nominal positive patterning task as a negative patterning one and the negative patterning task as a positive patterning one. This might account for the apparent negative patterning advantage in the prediction responses of Experiment 2 (although this advantage was not seen in the association ratings). For the maleficent condition, this analysis leads to the same prediction as was made for the health-reasoning task, with some cues associated with a salient outcome (disliking) and other cues associated with its absence (neutral).

Approach 1 and Approach 2 each offer speculative but testable ideas about how to interpret the effects of outcome valence, and further pursuit of both approaches should be fruitful.

Summary

The present research has demonstrated a robust positive patterning advantage in human causal reasoning that is moderated by the valence of the outcome. The presence of a positive patterning advantage strongly suggests that configural cues can play the kind of role in causal reasoning that they are presumed to play in animal conditioning. It also broadens the range of conditions, beyond those identified as critical by Young et al. (2000), in which configural cues contribute to causal reasoning.

The research has also demonstrated that outcome valence is one determinant of the positive patterning advantage in the human causal-reasoning task. One task for future research is to determine how broadly applicable this result is to other examples of patterning discriminations, and another task is to evaluate possible explanations for the finding. It seems clear, however, that the patterning discrimination task remains an interesting theoretical puzzle for associative learning theories.

AUTHOR NOTE

I express my appreciation to Ines Meier, John Beecroft, Melissa Anderson, Danielle Heaton, Elizabeth Cline, and Jennifer Osborn for assistance in collecting these data. I also thank Mary Bravo, Tony Dickinson, Geoffrey Hall, Harald Lachnit, and several anonymous reviewers for helpful comments on earlier versions of the article. Correspondence concerning this article should be addressed to J. W. Whitlow, Jr., Rutgers University, 311 N. Fifth Street, Camden, NJ 08102 (e-mail: bwhitlow@camden.rutgers.edu).

REFERENCES

Bellingham, W. P., Gillette-Bellingham, K., & Kehoe, E. J. (1985). Summation and configuration in patterning schedules with the rat and rabbit. *Animal Learning & Behavior*, **13**, 152-164.

Cheng, P. W., & Holyoak, K. J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, **17**, 391-416. doi:10.1016/0010-0285(85)90014-3

Cobos, P. L., López, F. J., Caño, A., Almaraz, J., & Shanks, D. R. (2002). Mechanisms of predictive and diagnostic causal induction. *Journal of Experimental Psychology: Animal Behavior Processes*, **28**, 331-346. doi:10.1037/0097-7403.28.4.331

Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how humans reason? Studies of the Wason selection task. *Cognition*, **31**, 187-276. doi:10.1016/0010-0277(89)90023-1

De Houwer, J., Becker, T., & Glautier, S. (2002). Outcome and cue properties modulate blocking. *Quarterly Journal of Experimental Psychology*, **55A**, 965-985. doi:10.1080/02724980143000578

Deisig, N., Lachnit, H., Giurfa, M., & Hellstern, F. (2001). Configural olfactory learning in honeybees: Negative and positive patterning discrimination. *Learning & Memory*, **8**, 70-78. doi:10.1101/lm.8.2.70

Deisig, N., Sandoz, J.-C., Giurfa, M., & Lachnit, H. (2007). The trial-spacing effect in olfactory patterning discriminations in honeybees. *Behavioural Brain Research*, **176**, 314-322. doi:10.1016/j.bbr.206.10.019

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, **56**, 218-226.

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought–action repertoires. *Cognition & Emotion*, **19**, 313-332. doi:10.1080/02699930441000239

Harris, J. A. (2006). Elemental representations of stimuli in associative learning. *Psychological Review*, **113**, 584-605. doi:10.1037/0033-295X.113.3.584

Harris, J. A., & Livesey, E. J. (2008). Comparing patterning and biconditional discriminations in humans. *Journal of Experimental Psychology: Animal Behavior Processes*, **34**, 144-154. doi:10.1037/0097-7403.34.1.144

Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 279-296). New York: Appleton-Century-Crofts.

Kinder, A., & Lachnit, H. (2002). Responding under time pressure: Testing two animal learning models and a model of visual categorization. *Quarterly Journal of Experimental Psychology*, **55A**, 173-193.

Krane, R. V., & Wagner, A. R. (1975). Taste aversion learning with a delayed shock US: Implications for the "generality of the laws of learning." *Journal of Comparative & Physiological Psychology*, **88**, 882-889. doi:10.1037/h0076417

Lachnit, H., & Kimmel, H. D. (1993). Positive and negative patterning in human classical skin conductance response conditioning. *Animal Learning & Behavior*, **21**, 314-326.

Lachnit, H., & Lober, K. (2001). What is learned in patterning discrimination? Further tests of configural accounts of associative learning in human electrodermal conditioning. *Biological Psychology*, **56**, 45-61. doi:10.1016/50301-0511(00)00087-9

Larkin, M. J. W., Aitken, M. R. F., & Dickinson, A. (1998). Retrospective revaluation of causal judgments under positive and negative contingencies. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1331-1352. doi:10.1037/0278-7393.24.6.1331

Le Pelley, M. E., Oakeshott, S. M., Wills, A. J., & McLaren, I. P. L. (2005). The outcome specificity of learned predictiveness effects: Parallels between human causal learning and animal conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, **31**, 226-236. doi:10.1037/0097-7403.31.2.226

Ludwig, I., & Lachnit, H. (2003). Asymmetric interference in patterning discriminations: A case of modulated attention? *Biological Psychology*, **62**, 133-146. doi:10.1016/S0301-0511(02)00124-2

Matute, H., Arcediano, F., & Miller, R. R. (1996). Test question modulates cue competition between causes and effects. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **22**, 182-196. doi:10.1037/0278-7393.22.1.182

McLaren, I. P. L., & Macintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, **30**, 177-200.

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. doi:10.1080/02724990244000025

Pearce, J. M. (1994). Similarity and discrimination: A selective review and connectionist model. *Psychological Review*, **101**, 578-607. doi:10.1037/0033-295X.101.4.587.

Pearce, J. M. (2002). Evaluation and development of a connectionist theory of configural learning. *Animal Learning & Behavior*, **30**, 73-95.

Rescorla, R. A. (1972). "Configural" conditioning in discrete-trial bar pressing. *Journal of Comparative & Physiological Psychology*, **79**, 307-317. doi:10.1037/h0032553

Rescorla, R. A. (1973). Evidence for "unique stimulus" account of configural conditioning. *Journal of Comparative & Physiological Psychology*, **85**, 331-338. doi:10.1037/h0035046

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

Shanks, D. R. (1991). Categorization by a connectionist network. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **17**, 433-443. doi:10.1037/0278-7393.17.3.433

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. doi:10.1037/0278-7393.24.6.1353

Shanks, D. R., & Darby, R. J. (1998). Feature- and rule-based generalization in human associative learning. *Journal of Experimental Psychology: Animal Behavior Processes*, **24**, 405-415. doi:10.1037/0097-7403.24.4.405

Shanks, D. R., Darby, R. J., & Charles, D. (1998). Resistance to interference in human associative learning: Evidence of configural processing. *Journal of Experimental Psychology: Animal Behavior Processes*, **24**, 136-150. doi:10.1037/0097-7403.24.2.136

Wagner, A. R. (2003). Context-sensitive elemental theory. *Quarterly Journal of Experimental Psychology*, **56B**, 7-29.

Waldmann, M. R. (2001). Predictive versus diagnostic causal learning: Evidence from an overshadowing paradigm. *Psychonomic Bulletin & Review*, **8**, 600-608.

Whitlow, J. W., Jr., & Wagner, A. R. (1972). Negative patterning in classical conditioning: Summation of response tendencies to isolable and configural components. *Psychonomic Science*, **27**, 299-301.

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. doi:10.1037/0278-7393.20.3.694

Woodbury, C. B. (1943). The learning of stimulus patterns by dogs. *Journal of Comparative Psychology*, **35**, 29-40. doi:10.1037/h0054061

Young, M. E., Wasserman, E. A., Johnson, J. L., & Jones, F. L. (2000). Positive and negative patterning in human causal learning. *Quarterly Journal of Experimental Psychology*, **53B**, 121-138. doi:10.1080/027249900392922

NOTES

1. A common convention is to designate the occurrence of an outcome with a "+" sign and the absence of an outcome with a "−" sign, making a negative patterning task an A+, B+, C+, ABC− discrimination. However, this convention would be confusing for the present experiments, which have positive outcomes, negative outcomes, and neutral outcomes. Hence, the occurrence of an outcome, whether beneficent or maleficent, is indicated by a "+" sign, as in A+, and the occurrence of no outcome or a neutral outcome is designated by a "0," as in AB0. Thus, the positive patterning task was DE0, EF0, DF0, and DEF+, whereas the negative patterning task was AB+, BC+, AC+, and ABC0.

2. The data display was designed to maximize the ease of comparing learning of the two discrimination tasks, since that was the focus of the research, but this choice may give the incorrect impression that responses to beneficent and maleficent outcomes were on the same scale. Other display choices are also reasonable. Plots of mean response choice, in which beneficent predictions tend toward "1" and maleficent predictions tend toward "3," with neutral choices varying between below "2" to above "2," better display the recorded behavior but make comparisons of the two discriminations difficult.

3. Data for 3 participants, who apparently misunderstood the instructions for the final rating to mean that they were to predict the outcome using "1," "2," and "3," were excluded from this analysis.

4. Data for 5 participants who used "1," "2," and "3" rather than the associative rating scale were excluded from this analysis.

(Manuscript received November 3, 2009;
accepted for publication November 5, 2009.)