Effects of pretraining on acquisition of novel configural discriminations in human predictive learning

RICK MEHTA AND EMILY RUSSELL
Acadia University, Wolfville, Nova Scotia, Canada

In two experiments, participants acquired one of two target configural discriminations (a biconditional or negative patterning discrimination) in a predictive learning task. In Experiment 1, participants were pretrained with either a configural or an elemental discrimination; in Experiment 2, they were pretrained with a configural discrimination, an elemental discrimination, or a control discrimination that was not expected to bias them toward elemental or configural processing. In both experiments, acquisition of the target configural discriminations was faster after configural pretraining than after elemental pretraining. In addition, the negative patterning discrimination was acquired faster than the biconditional discrimination. Finally, the results of Experiment 2 were more consistent with the notion that elemental pretraining hindered acquisition of the target discriminations than with the notion that configural pretraining enhanced their acquisition. Implications of these findings are discussed.

An important question for researchers in animal and human cognition is the extent to which compound stimuli are processed elementally or configurally. On the one hand, elemental theories such as the Rescorla–Wagner model (Rescorla & Wagner, 1972) treat compound stimuli as the sum of their constituent elements (see also Harris, 2006; Mackintosh, 1975; McLaren & Mackintosh, 2000, 2002; Pearce & Hall, 1980; Wagner, 2003; Wagner & Brandon, 2001). For example, if the elements A and B are reinforced individually, the associative strength accrued to the AB compound is thought to be equal to the sum of the associative strengths acquired by each element (i.e., AB = A + B). On the other hand, configural theories (e.g., Pearce, 1987, 1994, 2002; see also Kruschke, 1992) treat the AB compound as an entity that is distinct from the sum of its elements; in other words, the compound stimulus AB is not viewed as "A + B," but is instead thought to be a discrete event related to A and B only through physical similarity.

Although elemental and configural theories are sometimes viewed as being opposed to one another, recent theoretical efforts have moved away from the notion that humans and animals process complex stimuli either elementally or configurally toward the notion that both types of processing are required, but under different conditions (Fanselow, 1999; Melchers, Shanks, & Lachnit, 2008; Pearce & Bouton, 2001; Rudy & Sutherland, 1989, 1992; Williams, Sagness, & McPhee, 1994; Wynne, 1996). For example, Pearce's configural theory (Pearce, 1987, 1994) is better than the Rescorla–Wagner model (Rescorla & Wagner, 1972) in accounting for resistance to interference (Pearce & Wilson, 1991; Shanks, Charles, Darby, & Azmi,

1998; Shanks, Darby, & Charles, 1998; Williams, Gawel, Reimer, & Mehta, 2005) and for salience effects in discrimination learning (Pearce & Redhead, 1993; Redhead & Pearce, 1995), whereas the Rescorla–Wagner model is better able to account for summation (Myers, Vogel, Shin, & Wagner, 2001; Rescorla, 1997, 1999; Wagner, 2003) and relative validity (Wagner, Logan, Haberlandt, & Price, 1968). Thus, there is growing acceptance that the extent to which stimuli are processed configurally or elementally is flexible, rather than representing mutually exclusive alternatives.

Support for the view that configural and elemental strategies are used under different circumstances comes from previous research in animal and human learning identifying a number of variables that influence the processing of complex stimuli composed of separable elements (see Melchers et al., 2008, for a review). In animal learning, for example, Kehoe (1986; Kehoe & Shereurs, 1986) has shown that mere exposure to a compound stimulus can alter the extent to which that stimulus is processed elementally or configurally. Using a rabbit nictitating membrane conditioning preparation, Kehoe found that the amount of generalized excitation received by the elements of the reinforced AB decreased as the number of reinforced compound trials increased. Conversely, in human learning, Williams et al. (1994) have shown that instructions emphasizing the role of individual cues induce more elemental processing of compound stimuli in a predictive learning task, relative to control conditions. Thus, there is converging evidence from both the animal and human learning literatures that stimulus encoding is malleable, rather than fixed.

R. Mehta, rick.mehta@acadiau.ca

The focus of this article is on the role of prior learning, or pretraining, on the processing of compound stimuli in human predictive learning. Recent experiments have shown that experience with a discrimination that can be solved elementally encourages elemental processing of novel compounds, whereas experience with a discrimination that can be solved configurally encourages configural processing of novel compounds (Williams & Braker, 1999; Williams et al., 1994). For example, Mehta and Williams (2002, Experiments 2 and 3) used a two-stage transfer design to demonstrate that prior learning affects the processing of novel compounds. In Stage 1, participants acquired a discrimination in which the compound provided information that was either consistent with its elements (elemental: A+, B+, AB+, where "+" stands for trials on which an outcome occurs in predictive learning tasks or to reinforced trials in electrodermal conditioning experiments, and where "−" stands for no-outcome trials in predictive learning tasks or to unreinforced trials in electrodermal conditioning experiments) or inconsistent with its elements (configural: A+, B+, AB−). Participants then learned to differentiate two novel compounds (WX+, YZ−) in Stage 2, and were presented with the elements W, X, Y, and Z at test. The results showed that experience with the discrimination of Stage 1 strongly influenced responses to the test cues. Participants in the elemental group made a substantially higher proportion of outcome predictions to W and X, which were derived from WX+, than to Y and Z, which were derived from YZ−. However, participants in the configural group exhibited similar levels of responding to W, X, Y, and Z, which suggests that generalization between compounds and their constituent elements was reduced in the configurally pretrained group but not in the elementally pretrained group.

In addition to demonstrating that pretraining influences the way in which novel compounds are processed, Mehta and Williams (2002) identified terminal learning levels as a mediating variable that influences the effects of pretraining. That is, differences in responding to the elements presented at test were observed when learning levels in the last block of Stage 1 were high, but were not observed when the terminal learning levels were low. Furthermore, Williams and Braker (1999, Experiments 2 and 3) have identified time constraints as a mediating variable. Specifically, these authors found that participants' response patterns to the novel stimuli presented at test were affected by pretraining when the participants had ample time to make a response during the task, but that the response patterns to the novel cues were the same for the elemental and configural groups when participants were required to respond as quickly as possible.

The effect of pretraining has been observed not only in predictive learning tasks, but also in electrodermal conditioning experiments that have used letters presented on a computer screen as conditioned stimuli and an electric shock presented to the skin as an unconditioned stimulus. In one such study (Melchers, Lachnit, & Shanks, 2004, Experiment 2), participants were trained on either an elemental (A+, AB+, C−, CB−) or a configural (A−, AB+, C+, CB−) discrimination in Stage 1. The elemental discrimination could easily be solved by learning that A was a predictor of reinforcement and that C was a predictor of nonreinforcement. However, the configural discrimination could not be solved with an elemental strategy because B could not at the same time be an excitor on the A−, AB+ trials and an inhibitor on the C+, CB− trials; thus, a configural strategy had to be used to solve this discrimination. In Stage 2, participants received EX+ and FX− trials and were then presented with unreinforced presentations of E and F at test in order to measure transfer of prior learning. The results revealed that the elemental group showed more skin conductance responding to E than to F, whereas the configural group showed similar levels of responding to these cues. Thus, the effects of pretraining appear to be consistent across a range of learning tasks.

Although the findings described thus far support the notion that prior learning affects the processing of novel compounds, Melchers, Lachnit, Üngör, and Shanks (2005) were concerned about the generality of these results because confirmation for the effect of configural pretraining depended on the observation of no differences in responding to the novel cues presented at test (i.e., a null result for the configurally pretrained group). To address this issue, the authors used an alternative experimental design to measure the effects of prior learning by first training participants with an elemental or configural discrimination in Stage 1, and then presenting the participants with a novel configural discrimination in Stage 2. The fundamental difference between this design and the designs employed by Melchers et al. (2004) is that Melchers et al. (2005) used a target discrimination in Stage 2 that could only be solved configurally, whereas Melchers et al. (2004) used a target discrimination in Stage 2 that could be solved either elementally or configurally. According to the approach used by Melchers et al. (2005), pretraining with a configural discrimination should lead to faster acquisition of the novel discrimination, relative to pretraining with an elemental discrimination.

Melchers et al. (2005) have presented results from two electrodermal conditioning experiments that are consistent with this hypothesis. In Experiment 1, participants were pretrained with either an elemental (A+, B−, AB+, CD−) or a configural discrimination (AB+, BC−, CD+, DA−) in Stage 1, with 6 presentations of the four stimuli in both discriminations. In the elemental discrimination, A should become excitatory because it is consistently reinforced, whereas B should become neutral because it is reinforced and unreinforced an equal number of times; in this case, the AB compound can be thought of as the sum of its elements. In the configural discrimination, however, each element is equally reinforced and unreinforced; thus, the only way to solve this discrimination is by processing the configuration of cues that is presented on each trial. This discrimination has been called a *biconditional discrimination* because reinforcement is conditional on both cues present on a given trial (Saavedra, 1975). In Stage 2, participants were administered a negative patterning target discrimination (E+, F+, EF−), with 6 presentations of each reinforced element and 12 presentations of the unreinforced compound. The negative patterning discrimination

has long been thought to be a configural one because the associative strength of the unreinforced compound EF can never be equal to the combined associative strength of the two reinforced elements (Whitlow & Wagner, 1972). The authors reported that the negative patterning target discrimination was acquired faster after configural pretraining than after elemental pretraining.

In Experiment 2, Melchers et al. (2005) attempted to extend the results of Experiment 1 to a different target discrimination. As well, the discriminations used in pretraining were changed so that (1) the elemental and configural groups received equal exposure to elements and compounds, and (2) the same cues were presented for both the elemental and configural groups, the only difference residing in which cues were reinforced and which ones were unreinforced. In Stage 1, the elemental group was pretrained with an A+, AB+, C−, CB− discrimination, whereas the configural group was pretrained with an A−, AB+, C+, CB− discrimination; these discriminations were the same ones used in Experiment 2 of Melchers et al. (2004). In Stage 2, both groups were administered a biconditional target discrimination (DE+, EF−, FG+, GD−). Both groups experienced six trials with each cue in Stage 1 and in Stage 2. The authors reported an effect of pretraining, but noted that the effect was weaker than that observed in Experiment 1. Specifically, the results revealed no overall effect of pretraining when the analysis included all of the data from the six trials of Stage 2. However, when the authors restricted their analysis to only the first three presentations of the four stimuli, they found that the biconditional discrimination was acquired faster in the configural group. The authors reasoned that the effects of pretraining were weaker in Experiment 2 than in Experiment 1 because the biconditional discrimination was easier to learn, and thus would have been less sensitive to the effects of pretraining.

In summary, Melchers et al. (2005) have demonstrated that pretraining with configural discrimination leads to faster acquisition of target configural discriminations, relative to pretraining with an elemental discrimination. Their data also suggest that the effects of pretraining are stronger for a negative patterning target discrimination than for a biconditional target discrimination. However, the comparison for the effects of pretraining was made between experiments, rather than being compared directly; in addition, different pretraining cues were used across Experiments 1 and 2. These differences in the designs of the two experiments make it difficult to assess whether pretraining effects are similar for biconditional and negative patterning target discriminations. Furthermore, an interesting theoretical issue arises when one contemplates the results of Melchers et al. (2005)—namely, whether pretraining with the configural discrimination *enhanced* acquisition of the target discriminations or whether pretraining with an elemental discrimination *hindered* acquisition of the target discriminations. Without a control group, it is impossible to distinguish between these alternatives.

These issues provided the impetus for the present set of experiments. A broad goal of these experiments was to as-

sess the generality of the results of Melchers et al. (2005) by replicating their findings in the context of a predictive learning task. Replication of the main results of Melchers et al. (2005) with a different experimental preparation would demonstrate that their findings were not specific to the task, instructions, and procedures that they used. This finding, in turn, would provide stronger evidence for the notion that pretraining influences acquisition of novel configural discriminations. With regard to the goals of the specific experiments, Experiment 1 was designed to extend the work of Melchers et al. (2005) by directly comparing the effects of prior learning on acquisition of the negative patterning and biconditional target discriminations; Experiment 2 was designed to address the issue of whether the target configural discriminations are acquired faster after configural pretraining or are acquired more slowly after elemental pretraining.

EXPERIMENT 1

The design of Experiment 1, which was adapted from Melchers et al. (2005), is displayed in Table 1. As shown in the table, Stage 1 (the pretraining stage) consisted of either a configural discrimination (A−, AB+, C+, CB−) or an elemental discrimination (A+, AB+, C−, CB−). The specific discriminations that were administered in Stage 1 were adopted from Experiment 2 of Melchers et al. (2005) because they allowed the number of trials on which elements and compounds were followed by an outcome and by no outcome to be equated across conditions. In Stage 2, participants were presented with one of two target discriminations: negative patterning (E+, F+, EF−) or biconditional (DE+, EF−, FG+, GD−). A potential problem with this design was that there were two cues followed by an outcome and two cues followed by no outcome in the biconditional discrimination, whereas there were two cues followed by an outcome and only one cue followed by no outcome in the negative patterning discrimination. To equate the number of cues followed by an outcome and no outcome across conditions and to avoid using filler trials, twice as many EF− trials were used in the negative patterning discrimination (i.e., E+, F+, EF−, EF−). It should be noted that this procedure for equating outcome and no-outcome trials in the negative patterning discrimination was also employed by Melchers et al. (2005).

Table 1
Design of Experiment 1

Stage 1	Stage 2
Configural (A−, AB+, C+, CB−)	Negative patterning (E+, F+, EF−, EF−)
	Biconditional (DE+, EF−, FG+, GD−)
Elemental (A+, AB+, C−, CB−)	Negative patterning (E+, F+, EF−, EF−)
	Biconditional (DE+, EF−, FG+, GD−)

Note—The letters in parentheses represent the cues that were presented (chemicals); "+" refers to cues followed by an outcome (bacteria lived); "−" refers to cues not followed by an outcome (bacteria did not live).

Finally, the discriminations in Stages 1 and 2 were presented over blocks of trials. This aspect of the experimental design differs from that of Melchers et al. (2005), who presented each cue a total of six times within each stage. To increase the likelihood that the respective discriminations would be learned and to determine the nature of the effects of pretraining over a longer time, the number of cue presentations was increased and the stimuli were presented over blocks of trials. Each stage consisted of four blocks of trials, with each cue being presented five times in each block. Because the raw data were going to be transformed to proportional data (see the Results and Discussion section of Experiment 1 for more detail), each cue was presented five times per block (instead of six times per block) so that the transformed data would be expressed in intervals of 0.2 (1/5).

To summarize, Experiment 1 used a mixed design in which pretraining (configural or elemental) and target discrimination (negative patterning or biconditional) were between-subjects factors, and the cues presented in each stage and blocks of trials were within-subjects factors. On the basis of the results of Melchers et al. (2005), it was hypothesized that acquisition of the target configural discriminations of Stage 2 would be faster after configural pretraining than after elemental pretraining, and that this effect would be more pronounced for the negative patterning discrimination than for the biconditional discrimination.

Method

Participants. One hundred undergraduates from Acadia University participated for course credit or were paid $5 for their participation. Participants were tested individually or in groups of two and were randomly assigned to one of the four experimental conditions ($n = 25$ per cell).

Task. After reading and signing the consent form, participants read two screens of instructions, reproduced below.

Screen 1

Imagine that a new strain of bacteria, called E. Tremus, exists in the mammalian digestive system. Scientists plan to study the extent to which different chemicals influence the strain's survival. To do this, the bacterial sample will first be placed in culture (petri dishes). After that, one chemical (e.g., chemical 1) or a pair of chemicals (e.g., chemicals 1 and 2) will be added to the culture. A few hours later, the scientists will verify whether or not the bacterial sample survives.

Your job is to learn what effect each of the chemicals has on E. Tremus, and how combinations of chemicals influence its survival. That is, chemicals will be tested individually as well as in pairs.

Screen 2

You will be presented with the results from this study. There will be a total of 160 trials, and on each trial you will be told whether a single chemical (e.g., chemical 1), or a pair of chemicals (e.g., chemicals 1 and 2) was added to the bacterial sample. After predicting whether the culture survived, you will be informed whether you are correct and whether the bacteria survived. Use this information to learn what each chemical or pair of chemicals does.

When doing the task, try to keep track of what happened for each individual chemical and for each pair of chemicals. However, do not write down this information.

It is important that you concentrate during the experiment. To make the task simpler for you, you will not be told the actual names of the chemicals because their names are complicated. Instead, they will be labeled with letters such as A, B, C, D, E, and F.

After reading the instructions and having their questions answered, participants proceeded with the task. Information about the cues (chemicals added to the bacterial strain) and trial outcome (whether or not the bacteria lived) was presented on personal computers. The trial number and number of correct predictions were printed on a gray box across the top half of the screen. On each trial, participants were informed which chemicals were added to the bacterial strain. After entering whether they thought the bacteria would live, they were informed of the trial's outcome. Participants then pressed the space bar to commence the next trial. This procedure remained in place for all 160 trials.

To help ensure that participants could differentiate the cues, each chemical's letter name was displayed in a different color. When two chemicals were presented, the cues were presented on the same line and were separated by the word "and." For example, on an AB trial the participants were presented with the statement "A and B were added." Cue A was always on the left side of the word "and," and Cue B was always on the right side.

Procedure. A total of seven cues were presented (A, B, C, D, E, F, G). The letters A through L were used as letter names for the cues and were randomly assigned to their stimulus roles for each participant. Although there were seven cues in this experiment, 12 letter names (A–L) were used to ensure that the results of this experiment could be generalized across a range of letter names.

Stage 1 contained 80 trials that were divided into four blocks of trials. Four cues (A, AB, C, CB) were randomly presented five times within each block, resulting in 20 trials per block. In the configural discrimination, the outcome (the bacteria's survival) occurred when AB and C were presented but did not occur when A and CB were presented, and in the elemental discrimination, the outcome occurred when A and AB were presented but did not occur when C and CB were presented.

Stage 2 also consisted of 80 trials divided into four blocks of trials. In the negative patterning discrimination, the outcome occurred when the cues E and F were presented but not when EF was presented. The cues E and F were presented 5 times within each block, whereas EF was presented 10 times within each block. In the biconditional discrimination, the cues presented were DE, EF, FG, and GD, and each cue was presented 5 times in each block. The outcome occurred in the presence of DE and FG but not in the presence of EF and GD. Cue presentation was random within each block for both discriminations, and no event was used to mark the transition between Stages 1 and 2.

Results and Discussion

On each trial, an outcome prediction was recorded as "1" if participants predicted the bacterial sample would survive and as "0" if they predicted it would not. The proportion of outcome predictions was then calculated for each cue within each block. The variable of interest was the difference in responding to the outcome versus no-outcome trials. For the pretraining data from Stage 1, responding to AB+ and C+ was averaged and was compared with average responding to A− and CB− in the configural group, and responding to A+ and AB+ was averaged and was compared with average responding to C− and CB− in the elemental group. For the target discriminations of Stage 2, responding to E+ and F+ was averaged and compared with the average response to the two sets of EF− trials in the negative patterning discrimination, and responding

to DE+ and FG+ was averaged and compared with the average response to EF− and GD− in the biconditional discrimination. In all cases, there were no differences in responding to the sets of cues being averaged. A rejection criterion (α) of .05 was used for all statistical analyses. Unless stated otherwise, significant interactions were investigated with tests of simple main effects.

The Stage 1 discriminations were easily acquired. For brevity, only the data from the final block will be described. The elementally pretrained group mastered its discrimination, as evidenced by a ceiling effect on the outcome trials and a floor effect on the no-outcome trials. Ceiling effects were defined as means greater than .95, and floor effects were defined as means smaller than .05. The elemental group made a high proportion of outcome predictions

to the cues paired with the bacteria's survival ($M = .99$, $SE = .004$) and a low proportion of outcome predictions to the cues paired with the bacteria not surviving ($M = .01$, $SE = .005$). The group pretrained configurally did not exhibit ceiling and floor effects. However, participants in this group clearly acquired the discrimination: A higher proportion of outcome predictions was made to the cues paired with the bacteria's survival ($M = .94$, $SE = .02$) than to the cues paired with the bacteria not surviving ($M = .07$, $SE = .02$). Not surprisingly, this difference was statistically significant [$F(1,49) = 895.30, p < .001$].

The acquisition curves for the Stage 2 discriminations are shown in Figure 1. Although the overall patterns appear similar across the two levels of pretraining, interesting differences can be seen early in training. In particular,
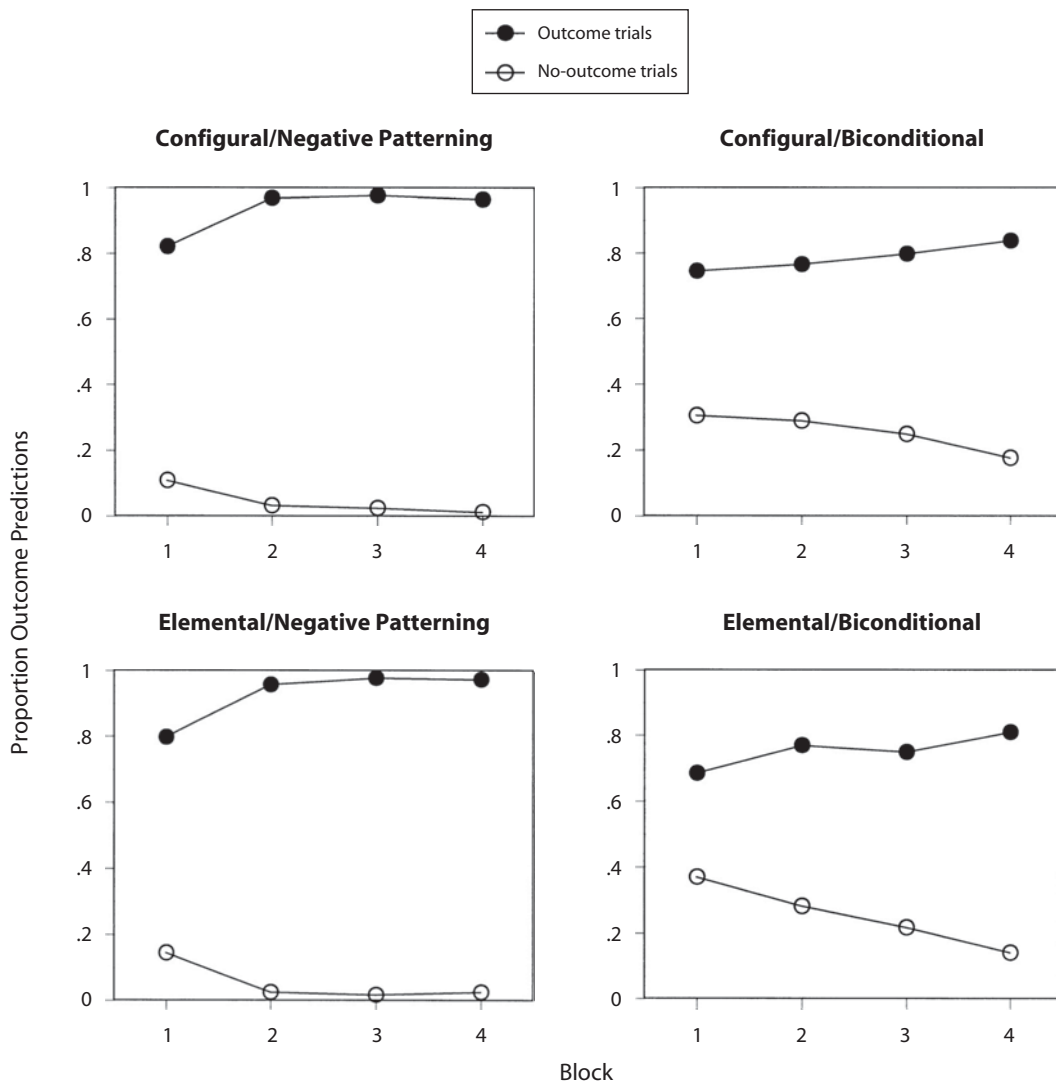
Figure 1. Mean proportion of outcome predictions for the outcome and no-outcome trials in Stage 2 of Experiment 1. The condition before the slash denotes whether pretraining in Stage 1 was configural or elemental, and the condition after the slash denotes whether the target discrimination in Stage 2 was a negative patterning or a biconditional discrimination. In the negative patterning discrimination, average responding to the E+ and F+ trials was compared with average responding on the two sets of EF− trials. In the biconditional discrimination, average responding on the DE+ and FG+ trials was compared with average responding on the EF− and GD− trials.

inspection of Block 1 suggests that the target discriminations were acquired faster after configural pretraining than after elemental pretraining, which is consistent with the results of Melchers et al. (2005). The data in the figure also suggest that the negative patterning discrimination was acquired much faster than the biconditional discrimination.

As for the statistical analyses of Stage 2, the data from Block 1 were analyzed separately from those of the subsequent blocks. Because Melchers et al. (2005) detected effects of pretraining in six or fewer trials, the first block of trials was of primary interest because it allowed for a reasonably direct comparison of the present results with those of Melchers and colleagues. The data from the second to fourth blocks were explored in order to assess the extent to which there was further learning of the target discriminations and to determine whether pretraining had any further effect on acquisition of these discriminations in the subsequent blocks of Stage 2.

A 2 (pretraining: configural or elemental) $\times$ 2 (target discrimination: negative patterning or biconditional) $\times$ 2 (outcome: outcome vs. no-outcome trials) ANOVA on the predictions made in Block 1 revealed main effects of target discrimination [$F(1,96) = 9.03, p < .01$] and outcome [$F(1,96) = 526.44, p < .001$], as well as pretraining $\times$ outcome [$F(1,96) = 3.97, p < .05$] and target discrimination $\times$ outcome [$F(1,96) = 43.30, p < .001$] interactions. The pretraining $\times$ outcome interaction attained significance because the difference in responding to the outcome versus no-outcome trials was larger after configural pretraining [$F(1,49) = 226.35, p < .001$] than after elemental pretraining [$F(1,49) = 148.41, p < .001$]. However, the three-way interaction (pretraining $\times$ target discrimination $\times$ outcome) failed to attain significance [$F(1,96) < 1$], which suggests that the benefits of configural pretraining were similar for both target discriminations.

Finally, the target discrimination $\times$ outcome interaction attained significance because the difference in responding to the outcome versus no-outcome trials was larger in the negative patterning discrimination [$F(1,49) = 524.22, p < .001$] than in the biconditional discrimination [$F(1,49) = 109.88, p < .001$]; in other words, the negative patterning discrimination was acquired more quickly than the biconditional discrimination, and this difference was apparent in the first block of Stage 2.

With regard to the statistical analysis of Blocks 2–4 of Stage 2, the data from the negative patterning discrimination had to be omitted from the analyses because there were clear ceiling and floor effects in the negative patterning discrimination. From an intuitive point of view, it was apparent that pretraining had no further influence on acquisition of this discrimination. Furthermore, from a statistical standpoint, the data from this condition could not be included in any analyses because they violated the homogeneity-of-variance assumption that needs to be satisfied in order to perform an ANOVA (Howell, 2002). For these reasons, only the data from the biconditional discrimination were examined. A 2 (pretraining) $\times$ 3 (block) $\times$ 2 (outcome) ANOVA on the predictions in the biconditional discrimination revealed a main effect of outcome [$F(1,48) = 229.24, p < .001$] and a block $\times$ outcome interaction [$F(2,96) = 9.87,$

$p < .001$]. However, the main effect of pretraining failed to attain significance, and this factor did not interact with any other variable [largest $F(1,48) = 2.76, p = .10$]. This analysis suggests that acquisition of the biconditional discrimination continued to improve over Blocks 2–4, but that acquisition was no longer influenced by pretraining. Taken together with the analyses of Block 1, the data imply that pretraining influenced acquisition of the target discriminations only in the first block of Stage 2.

To summarize, the main finding from Experiment 1 was that acquisition of the novel configural discriminations in Stage 2 was faster after configural pretraining than after elemental pretraining. This result replicates the findings of Melchers et al. (2005) and extends the generality of their results to a predictive learning task. However, contrary to Melchers et al. (2005), who found that pretraining effects were stronger for a negative patterning discrimination than for a biconditional discrimination, the effect of pretraining in the present experiment was similar for both target discriminations. Thus, there was general agreement in the finding that pretraining influences acquisition of the target configural discriminations, but a discrepancy in the nature of this influence. This discrepancy provided the first motivation for Experiment 2, which was to replicate Experiment 1 and assess the reliability of the finding that configural pretraining has a similar effect on acquisition of biconditional and negative patterning discriminations.

The second goal of Experiment 2 was to investigate whether pretraining with a configural discrimination in Stage 1 improved acquisition of the target discriminations in Stage 2, or whether pretraining with an elemental discrimination impeded acquisition of the target discriminations. This goal was achieved by including a control group with which the elementally and configurally pretrained groups could be compared.

The third impetus for Experiment 2 came from a secondary finding in Experiment 1—namely, that the negative patterning discrimination was acquired much faster than the biconditional discrimination. This difference was apparent in Block 1 and became more pronounced in the subsequent blocks. Furthermore, inspection of Figure 1 suggests that the difference in responding to the outcome versus no-outcome trials in the fourth block of the biconditional discrimination was similar to the level of differentiation observed in the first block of the negative patterning discrimination. A similar result has recently been reported by Harris and Livesey (2008). This finding is interesting because it contradicts the predictions of some associative learning theories, such as Pearce's configural theory (Pearce, 1987, 1994) and the unique-cue version of the Rescorla–Wagner model (Rescorla, 1972, 1973), but is more readily anticipated by the model of Harris (2006); the predictions from these models are described in the General Discussion section. Because of the theoretical significance of the differential learning rates for the two discriminations, the final goals of Experiment 2 were to replicate this finding and to rule out plausible alternative accounts for the present results. One possibility is that the negative patterning discrimination was acquired faster than the biconditional discrimination because there

were only three trial types in the negative patterning discrimination (E+, F+, EF−) but four trial types in the biconditional discrimination (DE+, EF−, FG+, GD−). A second possibility is that the negative patterning discrimination was acquired faster because there were twice as many EF− trials than E+ and F+ trials. Experiment 2 was designed to rule out these possibilities.

EXPERIMENT 2

The design for Experiment 2 is shown in Table 2. As in Experiment 1, participants were pretrained with a configural or an elemental discrimination. However, Experiment 2 also included a control group presented with the elements A− and C+ and with the compounds WX+ and YZ− (i.e., A−, WX+, C+, YZ−); this group thus had experience with elements and with compounds, but the compounds were not composed of previously seen elements. The control group was added in order to determine whether configural pretraining would facilitate acquisition of the target discriminations or whether elemental pretraining would hinder acquisition of the target discriminations.

In Stage 2, the biconditional discrimination was identical to that used in Experiment 1. However, the negative patterning discrimination was modified by adding a new compound (GD−) in order to decrease the number of presentations of EF− while equating the number of outcome and no-outcome trials across conditions. The main advantage of using GD− was that it served as a filler cue in the negative patterning discrimination but was part of the target discrimination in the biconditional discrimination.

Method

Participants and Task. One hundred twenty undergraduates from Acadia University participated for course credit or were paid $5 for their participation. Participants were tested individually or in groups of two and were randomly assigned to one of the six experimental conditions (n = 20 per cell). The task was identical to the one used in Experiment 1.

Procedure. A total of 11 cues were presented (A, B, C, D, E, F, G, W, X, Y, Z). The letters A through L were used as letter names

Table 2
Design of Experiment 2

Stage 1	Stage 2
Configural (A−, AB+, C+, CB−)	Negative patterning (E+, F+, EF−, GD−) Biconditional (DE+, EF−, FG+, GD−)
Elemental (A+, AB+, C−, CB−)	Negative patterning (E+, F+, EF−, GD−) Biconditional (DE+, EF−, FG+, GD−)
Control (A−, WX+, C+, YZ−)	Negative patterning (E+, F+, EF−, GD−) Biconditional (DE+, EF−, FG+, GD−)

Note—The letters in parentheses represent cues that were presented (chemicals); "+" refers to cues followed by an outcome (bacteria lived); "−" refers to cues not followed by an outcome (bacteria did not live).

for the cues and were randomly assigned to their stimulus roles for each participant.

Stage 1 contained 80 trials divided into four blocks of 20 trials. The procedure in this stage was identical to Stage 1 of Experiment 1 for the configural and elemental groups. However, the cues presented to the control group were A, WX, C, and YZ, which were randomly presented five times within each block. The outcome occurred in the presence of WX and C, but not in the presence of A and YZ.

Stage 2 also consisted of 80 trials divided into four blocks of 20 trials. The biconditional discrimination was identical to that of Experiment 1. However, in the negative patterning discrimination, the outcome occurred in the presence of E and F, but not in the presence of EF and GD. All four cues were randomly presented five times within each block, and no event was used to mark the transition between Stages 1 and 2.

Results and Discussion

For the data from Stage 1 (pretraining), responding to WX+ and C+ was averaged and compared with the average response to YZ− and A− in the control group. The Stage 1 data for the configural and elemental groups were treated in the same manner described in Experiment 1. For the target discriminations of Stage 2, the data for the biconditional discrimination were treated identically to the corresponding data of Experiment 1. As in Experiment 1, there were no differences in responding to the sets of cues being averaged. In the negative patterning discrimination, responding to E+ and F+ was averaged and compared with that to EF−. Because the cue GD− was a filler cue in the negative patterning discrimination, it was omitted from the statistical analysis. However, it was noted that the mean proportion of outcome predictions made to GD− was virtually identical to the average proportion of outcome predictions made to EF−.

Data from 3 participants in the biconditional treatment were removed from the analyses because these participants showed little evidence of having acquired their respective tasks. Two of the participants (1 from the configural/biconditional and 1 from the control/biconditional group) responded identically on the outcome and no-outcome trials during the experiment. Specifically, the 1st participant exhibited alternating strings of "yes" responses followed by strings of "no" responses in both Stage 1 and Stage 2, and the 2nd participant made an equal proportion of outcome predictions to the outcome and no-outcome trials across both stages of the experiment. The 3rd participant (who was in the control/biconditional group) acquired the Stage 1 discrimination, but made an equal proportion of outcome predictions to both the outcome and no-outcome trials in the first three blocks of Stage 2 and a *higher* proportion of outcome predictions on the *no-outcome* trials than on the outcome trials in the fourth block of Stage 2. Inspection of the data from these 3 participants suggested that they were outliers because their responses were random or because they had failed to pay attention to the stimuli in the experiment; for this reason, the data from these participants were of little interest. The final sample sizes ranged from 18 to 20 per group.

As in Experiment 1, the Stage 1 discriminations were easily acquired. Only the data from the final block will be described. The elemental group mastered its discrimination, and exhibited a ceiling effect on the outcome

trials ($M = .995$, $SE = .003$) and a floor effect on the no-outcome trials ($M = .003$, $SE = .003$). Although the configural and control groups did not exhibit ceiling and floor effects, they clearly learned their respective discriminations at similar levels by making a higher proportion of predictions on the outcome trials (configural, $M = .92$, $SE = .02$; control, $M = .94$, $SE = .02$) than on the no-outcome trials (configural, $M = .06$, $SE = .02$; control, $M = .10$, $SE = .03$). A 2 (pretraining: configural or control) $\times$ 2 (outcome) ANOVA on these data confirmed this impression by revealing only a main effect of outcome [$F(1,75) = 939.19$, $p < .001$]. Neither the main effect of pretraining nor the pretraining $\times$ outcome interaction attained significance, [largest $F(1,75) = 2.68$, $p = .11$]. Thus, both the configural and control groups clearly learned their respective discrimination, although acquisition was best in the elementally pretrained group.

The acquisition curves for the Stage 2 discriminations are shown in Figure 2. As in Experiment 1, the overall acquisition curves appear similar across the three levels of pretraining, but interesting differences can be seen in the early part of this stage. Specifically, the data from Block 1 suggest that the target discriminations were learned faster after configural pretraining than after elemental pretraining, and that the configural group acquired the target discriminations at a rate similar to that of the control group. In addition, the data in the figure show that the negative patterning discrimination was acquired faster than the biconditional discrimination. Consistent with Experiment 1, there were ceiling and/or floor effects in Blocks 2–4 in the negative patterning discrimination, which were not apparent in the biconditional discrimination. The statistical analyses that follow confirm the impressions just described.

To be consistent with the analyses used in Experiment 1, the data from Block 1 and from Blocks 2–4 were investigated separately. A 3 (pretraining: configural, elemental, or control) $\times$ 2 (target discrimination: negative patterning or biconditional) $\times$ 2 (outcome) ANOVA on the predictions made in Block 1 revealed a main effect of outcome [$F(1,111) = 330.83$, $p < .001$], as well as pretraining $\times$ outcome [$F(2,111) = 6.03$, $p < .01$] and target discrimination $\times$ outcome [$F(1,111) = 7.96$, $p < .01$] interactions. To discover the nature of the pretraining $\times$ outcome interaction, the difference between the proportion of predictions made to the outcome and no-outcome trials was first calculated for each participant. Tukey's HSD test was then conducted on the difference scores to examine the group differences; a post hoc test was used because the hypothesis regarding the pattern of the mean difference scores was nondirectional. The analysis revealed that the mean difference scores for the configural and control groups were not significantly different from one another and that the mean difference scores for these two groups (collectively) were larger than for the elemental group—(configural = control) > elemental. Consistent with Experiment 1, the pretraining $\times$ target discrimination $\times$ outcome interaction was not statistically significant [$F(2,111) < 1$], which suggests that the effects of pretraining were similar for acquisition of the biconditional and negative patterning

discriminations. The target discrimination $\times$ outcome interaction attained significance because the difference in responding to the outcome versus no-outcome trials was larger in the negative patterning discrimination [$F(1,59) = 278.96$, $p < .001$] than in the biconditional discrimination [$F(1,56) = 189.46$, $p < .001$].

With regard to Blocks 2–4, the data from the negative patterning discrimination could not be analyzed because of ceiling and floor effects. Thus, statistical analyses were restricted to the biconditional discrimination. A 3 (pretraining) $\times$ 3 (block) $\times$ 2 (outcome) ANOVA on the predictions made in Blocks 2–4 revealed a main effect of outcome [$F(1,54) = 276.84$, $p < .001$] and a block $\times$ outcome interaction [$F(2,108) = 10.92$, $p < .001$]. In consonance with Experiment 1, the data suggested that acquisition of the biconditional discrimination improved over Blocks 2–4, and that pretraining did not have an impact on acquisition of this discrimination after the first block of trials.

In summary, Experiment 2 replicated the main results of Experiment 1: (1) Acquisition of the novel configural discriminations in Stage 2 was faster after configural pretraining than after elemental pretraining, and (2) the effect of pretraining was similar for both the biconditional and negative patterning target discriminations. Experiment 2 extended the results of Experiment 1 by demonstrating that the configurally pretrained group did not perform better than the control group, which suggests that acquisition of the novel configural discriminations of Stage 2 was hindered by elemental pretraining, as opposed to being facilitated by configural pretraining. Finally, Experiment 2 replicated the finding that the negative patterning target discrimination was acquired much faster than the biconditional discrimination. In Experiment 1, it could have been argued that this result arose because the negative patterning group experienced twice as many EF$-$ trials (in relation to E+ and F+ trials) or had experience with only three rather than four trial types. However, Experiment 2 ruled out these alternative accounts by modifying the negative patterning discrimination so that participants experienced E+, F+, EF$-$, GD$-$ trials. This modification reduced the number of EF$-$ trials while ensuring that participants in this condition still had experience with four trial types. Although the finding of faster acquisition of the negative patterning discrimination was a secondary finding, it was replicated in two experiments.

GENERAL DISCUSSION

One goal of the present experiments was to assess the effects of prior learning on acquisition of novel configural discriminations. Consistent with Melchers et al. (2005), these discriminations were learned faster after configural pretraining than after elemental pretraining. The present experiments also employed a larger number of trials in Stages 1 and 2, and the data from Stage 2 confirmed that pretraining influences the processing of novel stimuli for a small number of trials. Specifically, the configurally pretrained group was better than the elementally pretrained group for only the first block of five trials. This result mir-
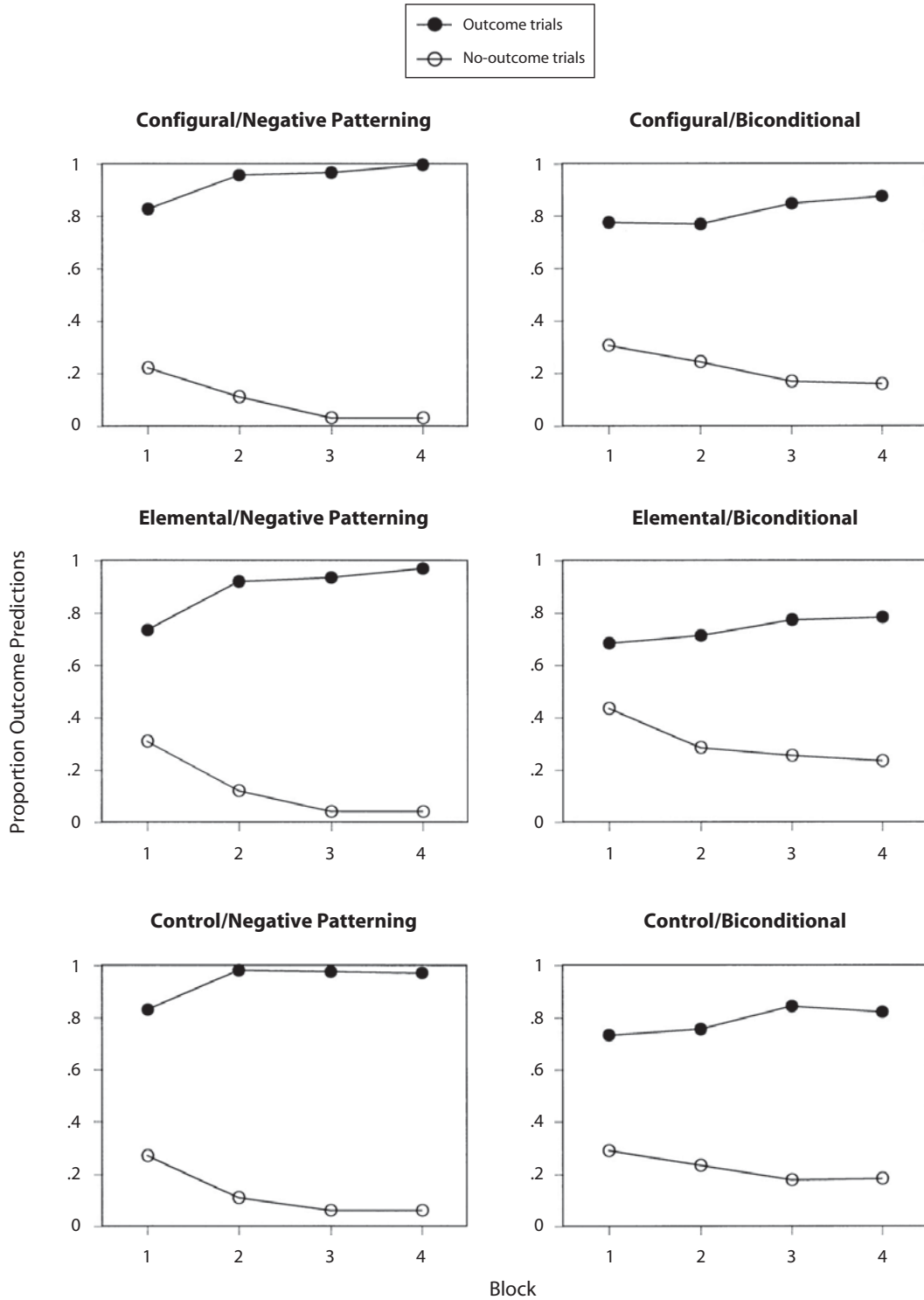
Figure 2. Mean proportion of outcome predictions on the outcome and no-outcome trials in Stage 2 of Experiment 2. The condition before the slash denotes whether Stage 1 was a configural, an elemental, or a control discrimination. The condition after the slash denotes whether the target discrimination in Stage 2 was a negative patterning or a biconditional discrimination. In the negative patterning discrimination, average responding on the E+ and F+ trials was compared with average responding on the EF− trials. In the biconditional discrimination, average responding on the DE+ and FG+ trials was compared with average responding on the EF− and GD− trials.

rors those of Melchers et al. (2005), who found that configural pretraining led to faster acquisition across all six cue exposures in their first experiment and was restricted to the first three trials with each cue in their second experiment. The present findings thus replicate the main results of Melchers et al. (2005) and extend them to a predictive learning task.

Although our results were generally consistent with those of Melchers et al. (2005), a notable discrepancy was that the benefits of configural pretraining were similar for both target discriminations in the present experiments, whereas Melchers et al. (2005) found that configural pretraining had a stronger influence on acquisition of the negative patterning target discrimination than of the biconditional target discrimination. Differences in experimental design and procedure may account for this discrepancy. Specifically, the present experiments were designed to compare the effects of pretraining on acquisition of the two target discriminations directly and thus used the same pretraining discriminations across conditions, whereas the experiments of Melchers et al. (2005) were not designed to address this issue and had to rely on between-experiment comparisons; furthermore, it would be difficult to compare the effects of pretraining across their experiments because each experiment employed different pretraining discriminations. Both of these differences in the experimental designs could have contributed to the discrepant results attained.

With regard to experimental procedure, it is possible that configural pretraining led to faster acquisition of both target discriminations in the present experiments because participants had 20 exposures to each stimulus in Stage 1 (4 blocks of trials × 5 presentations of each cue per block = 20 presentations) whereas participants in the experiments of Melchers et al. (2005) had 6 exposures to each stimulus. This increase in exposure to the cues of Stage 1 may have increased the likelihood of observing a transfer effect for both of the target discriminations, regardless of their difficulty. Consistent with this notion, Nakagawa (1999) trained rats on a discrimination to criterion or overtrained them for 10 or 20 days prior to presenting them with a novel discrimination at test, and found that overtrained rats showed more transfer than those who were trained to criterion; furthermore, transfer effects were larger with increased overtraining. Given this finding, it seems reasonable to assert that that the increased experience with the discriminations of Stage 1 may have led to faster acquisition of both target discriminations in the present experiments, even if one was more difficult than the other.

The second goal of the present experiments was to extend the findings of Melchers et al. (2005) by assessing whether configural pretraining aids or elemental pretraining hinders acquisition of the target discriminations. The finding from Experiment 2 that the elementally pretrained group acquired the target configural discriminations at a slower rate than did the configural and control groups supports the latter interpretation. This result is consistent with findings from other research suggesting that participants who are pretrained with a control discrimination tend to either rely on a configural strategy or exhibit less generalization than does an elementally pretrained group when presented with novel cues at test (Experiment 2 of Williams & Braker, 1999; see also Shanks, Charles, et al., 1998).

The findings from the present experiments thus add to a growing body of literature suggesting that the extent to which compound stimuli are processed elementally or configurally is flexible, rather than fixed (Kehoe, 1986; Mehta & Williams, 2002; Melchers et al., 2008; Pearce & Bouton, 2001; Rudy & Sutherland, 1989, 1992, 1995; Williams & Braker, 1999; Williams et al., 1994). Presumably, this flexibility would allow an information processing system to adapt to its environment by allowing it to use elemental or configural processing, depending on which one is required or is better able to solve the problem at hand. For this reason, models that are fixed in terms of whether stimuli are processed strictly configurally or elementally do not readily account for effects of prior learning because of their adherence to one approach (Harris, 2006; Pearce, 1987, 1994; Rescorla & Wagner, 1972). However, there are two schools of thought that can address the issue of flexibility of processing. According to the first view, stimuli are processed along a continuum, with purely elemental being at one extreme and purely configural being at the other extreme. The alternative view is that there are two information processing systems (one elemental and one configural), and that the relative activation of the two systems determines which one is used (Fanselow, 1999). Within each framework, pretraining can influence the way in which cues are processed. The view of stimuli being processed along a continuum will be considered first.

The hypothesis that stimuli are processed along a configural–elemental continuum can provide a reasonable fit to the present results by either (1) modifying a configural approach so that the amount of generalization between elements and compounds can vary or (2) modifying the nature of the interactions between elements and compounds in elemental models. With regard to the first approach, Williams and Braker (1999) stated that Pearce's (1987, 1994) configural model has difficulty accounting for pretraining effects because the amount of generalization between elements and compounds in this model is fixed; they suggested that a parameter can be added to the model that would allow the amount of generalization to vary. The addition of this parameter would increase the flexibility of the model and would make it similar to Kruschke's (1992) ALCOVE model. This modification would allow the model to provide a good fit to the present data by assuming that there is little generalization between elements and compounds when participants are pretrained with a control discrimination and by predicting that elemental pretraining will lead to an increase in the generalization parameter. The assumption of configural processing as a default strategy would imply that the control group and the configurally pretrained group would process the novel discrimination in Stage 2 with a configural strategy and that this strategy would allow the target discriminations to be solved. However, after elemental pretraining, the value of the generalization parameter should increase,

which would result in an elemental strategy initially being used to solve the target configural discriminations. Consequently, the elementally pretrained group should be impaired on acquisition of the discriminations. As this group gains more experience with the configural discrimination, the value of the generalization parameter would decrease in fairly short order, which would allow this group to perform at levels similar to the control and configurally pretrained groups after the first block of trials.

In general, the idea of adding a generalization parameter to Pearce's (1987, 1994) model allows it to account for the data from the present experiments. However, there are two main challenges that need to be overcome before this parameter can be integrated into Pearce's model. First, to minimize the number of additions and modifications to the model, the generalization parameter would have to encode a variety of factors that can influence the processing of compound stimuli, such as the cover story and instructions employed (Melchers et al., 2008). A consequence of doing so is that the model's predictive power would become situation specific. Second, the main disadvantage of incorporating this parameter is that it reduces the testability of the model by making it too flexible. Unless there are firm guidelines that dictate the amount of the generalization between elements and compounds as a function of variables such as cover stories and instructions, and that state the conditions under which the parameter will change, it is difficult to use this approach to make predictions about future experimental results.

One might enquire whether more recent elemental approaches, such as the replaced-elements version of the Rescorla–Wagner model (Wagner, 2003; Wagner & Brandon, 2001) or McLaren and Mackintosh's (2000, 2002) model, can provide a reasonable fit to the present data. For brevity, this discussion will focus on the replaced-elements model. According to this approach, individual cues activate representations that contain a fixed number of elements. When cues are presented in combination, the compound activates a representation of a unique cue that encodes specifically for a compound stimulus and is able to inhibit a proportion of the elements. Depending on the proportion of inhibited elements, a compound stimulus can be processed along a configural–elemental continuum. Wagner (2003) has discussed how this approach can be used to describe and predict when summation will be observed in animal learning.

Within the context of the present experiments, it seems reasonable to hypothesize that there would be a large amount of inhibition between the unique cue and the elements after experience with a control or configural discrimination but that the amount of inhibition would decrease after elemental pretraining. Thus, when participants are administered the configural target discrimination, those pretrained elementally would initially be slower at solving the discrimination than those pretrained configurally. However, a mechanism would need to be in place that would allow the amount of inhibition to increase as a result of experience with the target discrimination; furthermore, this increase would have to happen fairly quickly to allow the elementally pretrained group to perform at levels similar to those of the configural and control groups after the first block of trials.

An advantage of the replaced-elements model is that the model is flexible because it already has a built-in mechanism that allows the amount of inhibition between the unique cue and the elements to vary; presumably, the amount of inhibition that is observed in a predictive learning task can also be influenced by other variables (e.g., cover story, instructions). However, as in the critiques made to the notion of modifying Pearce's (1987, 1994) theory, the factors that dictate the extent to which there is an increase or decrease in inhibition need to be established in order for this approach to be testable.

Although the view of compound stimuli varying along a configural–elemental continuum can account for the data of the present experiments, an alternative framework that can also account for the present results is the conceptualization of multiple learning systems working in parallel (Fanselow, 1999). According to this proposal, there are two learning systems: a configural system and an elemental one. The goal of these systems is to process the events presented to them and to use this information to make predictions about the relationships among these events. This approach assumes that the configural and elemental systems operate independently and that both systems are thus able to acquire and store information. Although the two systems process information in parallel, they interact with each other by competing for association with an outcome. The system that more accurately predicts the occurrence of the outcome will become associated with the outcome and will inhibit the other system.

This account can explain the results of the present experiments by postulating that activation of the elemental system is suppressed when participants undergo configural pretraining. Because the configural system is more active at the start of Stage 2, participants in this group can be expected to acquire their target discrimination faster than would participants pretrained with an elemental discrimination. In addition, because the configural and elemental learning systems are operating in parallel, activity in the configural system should start to increase when the elemental system is not easily able to solve the target discriminations; this change in activity should allow for the target discriminations to be solved. Although this approach can account for the difference in acquisition rates between the elemental and configural groups, it does not readily explain why the group pretrained with a control discrimination would use a configural strategy when presented with the target configural discrimination.

Future research should assess the reliability and generality of the findings from the present experiments. This goal can be achieved by using a different task (e.g., skin conductance preparation; predictive learning task with a different cover story), as well as by manipulating the complexity of the pretraining and/or target discriminations. Although the comparison of interest is usually the difference between the elemental and configurally pretrained groups, the results from the present experiments suggest

that a control condition is necessary to serve as a baseline condition with which the elementally and configurally pretrained groups can be compared. The main advantage of doing so is that the control condition provides useful information about what type of processing is used on a task when participants are pretrained on a discrimination that cannot be solved elementally or configurally, and how pretraining shifts the processing of stimuli from this baseline condition. This approach should putatively be useful for future theory development.

The main topic of concern in this article was how pretraining affects acquisition of novel configural discriminations. However, a secondary issue that arose was the finding that the negative patterning target discrimination was acquired faster than the biconditional target discrimination. A similar finding was recently reported by Harris and Livesey (2008). In their experiments, participants acquired a negative patterning, a positive patterning (e.g., A−, B−, AB+; the outcome occurs only when both cues are present on a given trial), and a biconditional discrimination in two within-subjects experiments. The authors reported that the two patterning discriminations were acquired faster than the biconditional discrimination. The main difference between the results of the present experiments, which used the target discriminations as a between-subjects factor, and the experiments of Harris and Livesey is that the negative patterning discrimination was mastered much faster in the present experiments.

The finding that the negative patterning discrimination was learned faster than the biconditional discrimination is interesting from a theoretical perspective because it is not predicted by Pearce's (1987, 1994) configural theory, the unique-cue version of the Rescorla–Wagner model (Rescorla, 1972, 1973), or the replaced-elements version of the Rescorla–Wagner model (Wagner, 2003; Wagner & Brandon, 2001), but is predicted by Harris's (2006) elemental theory. The remaining discussion will describe the predictions from these models. Pearce's theory will be considered first because it was designed to account for the acquisition of configural discriminations, which were not readily explained by the original formulation of the Rescorla–Wagner model (Rescorla & Wagner, 1972). The elemental approaches will be considered next because Harris's theory makes a correct prediction in a situation where a configural theory and other elemental models anticipate faster acquisition for the biconditional discrimination than for the negative patterning discrimination.

According to Pearce's (1987, 1994) configural theory, the entire stimulus complex present on a given trial acquires or loses associative strength, depending on whether it is reinforced (outcome occurs) or unreinforced (outcome does not occur); the stimulus complex also receives generalized excitatory or inhibitory associative strength from physically similar stimuli. Applying this concept to a biconditional discrimination (DE+, EF−, FG+, GD−), the amount of associative strength accrued to DE, for example, would be equal to the sum of DE's own associative strength and generalized inhibitory strength received from

EF. The amount of generalized inhibition received by DE from EF($_{EF}e_{DE}$) is determined by Equation 1:

$$_{EF}e_{DE} = (_{EF}S_{DE})(E_{EF}). \qquad (1)$$

According to this equation, the amount of associative strength that DE receives from EF is equal to the product of the similarity of EF to DE ($_{EF}S_{DE}$) and EF's associative strength (E_{EF}). The similarity of two cues is determined by the extent to which they share common elements. Assuming that DE and EF are equally salient, the common element is E, which comprises one half of EF and one half of DE. The product of these two proportions is their similarity ($_{EF}S_{DE} = .5 \times .5 = .25$). Because the similarity of DE and EF is small, there will be little generalized inhibition passed from EF to DE, and the cues will be relatively easy to differentiate. The concepts just described can be applied to the other compounds in the biconditional discrimination.

One should now consider what happens when Pearce's (1987, 1994) theory is applied to a negative patterning discrimination (E+, F+, EF−). According to the model, the amount of associative strength accrued to EF on an EF− trial will be equal to the sum of (1) EF's own associative strength and (2) generalized excitation received from the elements E and F. The amount of generalized excitation that EF receives from E is determined by the similarity of E to EF. The common element E will comprise 100% of itself and one half of EF; the product of these two proportions will thus be .5 ($_{E}S_{EF} = 1 \times .5 = .5$). The same reasoning applies to the amount of generalized excitation that EF will receive from F. When one compares the amount of similarity among the cues in the biconditional and negative patterning discriminations, one can see that there is much more similarity in the negative patterning discrimination than in the biconditional discrimination (.5 vs. .25). By extension, there will be larger amounts of generalization among the stimulus complexes in the negative patterning discrimination, which will make cue differentiation more difficult in this discrimination. Consequently, the difference in the asymptotic associative strengths for the outcome versus no-outcome trials will be larger in the negative patterning discrimination, and more trials will be required to reach asymptote. Computer simulations of Pearce's (1987, 1994) model confirm these predictions by revealing that (1) the asymptotic weights for the negative patterning discrimination are $+2\lambda$ for the reinforced elements (E+, F+) and -2λ for the unreinforced compound (EF−), whereas the asymptotic weights for the biconditional discrimination are $+1.33\lambda$ for the reinforced compounds (DE+, FG+) and -0.67λ for the unreinforced compounds (EF−, GD−), and (2) that acquisition is predicted to be faster for the negative patterning discrimination.

Turning to the predictions of the unique-cue version of the Rescorla–Wagner model (Rescorla, 1972, 1973), this theory treats compounds as the sum of their elements; however, it is able to solve configural discriminations by including cues that encode for the presence of compound stimuli. In a negative patterning discrimination (E+, F+, EF−), E and F each gain 1λ of associative

strength because both cues are consistently reinforced or followed by an outcome. The EF compound is encoded as EFX, with X representing the unique combination of E with F. The cue X will accumulate -2λ of associative strength to offset the excitation attained by E and F, so that the associative strength of the unreinforced EF− trials sums to 0. In the biconditional discrimination (DE+, EF−, FG+, GD−), a unique cue is assigned to each of the four compounds (i.e., W, X, Y, Z); the discrimination can thus be depicted as DEW+, EFX−, FGY+, and GDZ−. The cues D, E, F, and G do not gain associative strength because they are followed by the occurrence and nonoccurrence of the outcome an equal number of times. However, W and Y will quickly gain large increments in associative strength and will attain 1λ of associative strength at asymptote, whereas Y and Z will not gain associative strength and will stay neutral with 0λ of associative strength at asymptote (Harris, 2006; Rudy & Sutherland, 1995). In a manner similar to that of Pearce's configural theory, the unique-cue model predicts that the biconditional discrimination will be acquired faster than the negative patterning discrimination because the difference in the asymptotic weights of the reinforced versus unreinforced cues (outcome vs. no-outcome trials) is smaller in the biconditional discrimination, making fewer trials necessary to reach asymptote. A similar prediction is also made by the replaced-elements version of the Rescorla–Wagner model (Wagner, 2003; Wagner & Brandon, 2001) because it relies on the presence of unique cues to solve the discriminations.

Although Pearce's (1987, 1994) model and variations of the Rescorla–Wagner (Rescorla & Wagner, 1972) model do not make a correct prediction in this situation, Harris's (2006) model can account for the negative patterning discrimination being acquired faster than the biconditional discrimination. According to this theory, a stimulus is composed of a number of elements, some of which are strong (or salient) and some of which are weak. The key idea in the theory is that only the most salient elements of a stimulus can enter a limited-capacity attention buffer. When an individual cue is presented, its most salient elements will enter the buffer; however, the buffer will still have enough capacity to allow the weaker elements to enter. But when a compound stimulus is presented, only the most salient elements from each cue that composes the compound will enter the buffer. Applied to a negative patterning discrimination (E+, F+, EF−), only the strongest elements of E and F will enter the buffer on EF− trials; however, the weaker elements of E and F will enter the buffer on E+ and F+ trials (i.e., when these cues are presented individually). The strongest elements of E and F should become inhibitory on the EF− trials, and the weaker elements should become excitatory on the E+ and F+ trials. This mechanism allows the discrimination to be solved relatively quickly. However, in a biconditional discrimination (DE+, EF−, FG+, GD−), it is more difficult for the elements of D, E, F, and G to enter the buffer because each cue is followed by an outcome and by no outcome an equal number of times. In order for the discrimination to be solved, the model relies on inciden-

tal differences in the distribution of salient and weak elements in the cues in order to allow for a small number of elements to enter the buffer upon presentation of certain compounds, but not others (Harris & Livesey, 2008). This mechanism allows for the discrimination to eventually be solved, but does so with more difficulty than the negative patterning discrimination does.

REFERENCES

FANSELOW, M. S. (1999). Learning theory and neuropsychology: Configuring their disparate elements in the hippocampus. *Journal of Experimental Psychology: Animal Behavior Processes*, **25**, 275-283. doi:10.1037/0097-7403.25.3.275

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

HOWELL, D. C. (2002). *Statistical methods for psychology* (5th ed.). Pacific Grove, CA: Duxbury.

KEHOE, E. J. (1986). Summation and configuration in conditioning of the rabbit's nictitating membrane response to compound stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, **12**, 186-195. doi:10.1037/0097-7403.12.2.186

KEHOE, E. J., & SHEREURS, B. J. (1986). Compound-component differentiation as a function of CS–US interval and CS duration in the rabbit's conditioned nictitating membrane response. *Animal Learning & Behavior*, **14**, 144-154.

KRUSCHKE, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, **99**, 22-44. doi:10.1037/0033-295X.99.1.22

MACKINTOSH, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, **82**, 276-298. doi:10.1037/h0076778

McLAREN, I. P. L., & MACKINTOSH, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, **28**, 211-246.

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

MEHTA, R., & WILLIAMS, D. A. (2002). Elemental and configural processing of novel cues in deterministic and probabilistic tasks. *Learning & Motivation*, **33**, 456-484. doi:10.1016/S0023-9690(02)00008-5

MELCHERS, K. G., LACHNIT, H., & SHANKS, D. R. (2004). Past experience influences the processing of stimulus compounds in human Pavlovian conditioning. *Learning & Motivation*, **35**, 167-188. doi:10.1016/S0023-9690(03)00044-4

MELCHERS, K. G., LACHNIT, H., ÜNGÖR, M., & SHANKS, D. R. (2005). Prior experience can influence whether the whole is different from the sum of its parts. *Learning & Motivation*, **36**, 20-41. doi:10.1016/j.lmot.2004.06.002

MELCHERS, K. G., SHANKS, D. R., & LACHNIT, H. (2008). Stimulus coding in human associative learning: Flexible representations of

parts and wholes. *Behavioural Processes*, **77**, 413-427. doi:10.1016/j .beproc.2007.09.013

Myers, K. M., Vogel, E. H., Shin, J., & Wagner, A. R. (2001). A comparison of the Rescorla–Wagner and Pearce models in a negative patterning and summation problem. *Animal Learning & Behavior*, **29**, 36-45.

Nakagawa, E. (1999). Acquired equivalence of discriminative stimuli following two concurrent discrimination learning tasks as a function of overtraining in rats. *Psychological Record*, **49**, 327-348.

Pearce, J. M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review*, **94**, 61-73.

Pearce, J. M. (1994). Similarity and discrimination: A selective review and a 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.

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology*, **52**, 111-139. doi:10.1146/ annurev.psych.52.1.111

Pearce, J. M., & Hall, G. (1980). A model of Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, **87**, 532-552. doi:10.1037/0033 -295X.87.6.532

Pearce, J. M., & Redhead, E. S. (1993). The influence of an irrelevant stimulus on two discriminations. *Journal of Experimental Psychology: Animal Behavior Processes*, **19**, 180-190. doi:10.1037/0097 -7403.19.2.180

Pearce, J. M., & Wilson, P. N. (1991). Failure of excitatory conditioning to extinguish the influence of a conditioned inhibitor. *Journal of Experimental Psychology: Animal Behavior Processes*, **17**, 519-529. doi:10.1037/0097-7403.17.4.519

Redhead, E. S., & Pearce, J. M. (1995). Similarity and discrimination learning. *Quarterly Journal of Experimental Psychology*, **48B**, 46-66.

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. (1997). Summation: Assessment of a configural theory. *Animal Learning & Behavior*, **25**, 200-209.

Rescorla, R. A. (1999). Summation and overexpectation with qualitatively different outcomes. *Animal Learning & Behavior*, **27**, 50-62.

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

Rudy, J. W., & Sutherland, R. J. (1989). The hippocampal formation is necessary for rats to learn and remember configural discriminations. *Behavioural Brain Research*, **34**, 97-109. doi:10.1016/S0166 -4328(89)80093-2

Rudy, J. W., & Sutherland, R. J. (1992). Configural and elemental associations and the memory coherence problem. *Journal of Cognitive Neuroscience*, **4**, 208-216. doi:10.1162/jocn.1992.4.3.208

Rudy, J. W., & Sutherland, R. J. (1995). Configural association theory and the hippocampal formation: An appraisal and reconfiguration. *Hippocampus*, **5**, 375-389. doi:10.1002/hipo.450050502

Saavedra, M. A. (1975). Pavlovian compound conditioning in the rabbit. *Learning & Motivation*, **6**, 314-326. doi:10.1016/0023 -9690(75)90012-0

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., & 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. doi:10.1080/ 02724990244000133

Wagner, A. R., & Brandon, S. E. (2001). A componential theory of Pavlovian conditioning. In R. R. Mower & S. B. Klein (Eds.), *Handbook of contemporary learning theories* (pp. 23-64). Mahwah, NJ: Erlbaum.

Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of Experimental Psychology*, **76**, 171-180. doi:10.1037/h0025414

Whitlow, J. W., & 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., & Braker, D. S. (1999). Influence of past experience on the coding of compound stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, **25**, 461-474. doi:10.1037/0097 -7403.25.4.461

Williams, D. A., Gawel, J. D., Reimer, D. S., & Mehta, R. (2005). Resistance to interference in complex negative patterning. *Learning & Behavior*, **33**, 417-427.

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

Wynne, C. D. L. (1996). Transverse patterning in pigeons. *Behavioural Processes*, **38**, 119-130. doi:10.1016/S0376-6357(96)00032-0