Journal of Experimental Psychology: General
2003, Vol. 132, No. 2, 266–276

0096-3445/03/$12.00 DOI: 10.1037/0096-3445.132.2.266

Acquired Equivalence and Distinctiveness in Human Discrimination Learning: Evidence for Associative Mediation

Geoffrey Hall
University of York

Chris Mitchell
Unilever Research

Steven Graham
University of York

Yvonna Lavis
University of New South Wales

In the first stage of Experiments 1–3, subjects learned to associate different geometrical figures with colors or with verbal labels. Performance in Stage 2, in which the figures signaled which of 2 motor responses should be performed, was superior in subjects required to make the same response to figures that had shared the same Stage 1 associate. A third stage of testing showed that the events used as associates in Stage 1 were capable of evoking the motor response trained in Stage 2, an outcome predicted by an associative interpretation of such transfer effects. Experiment 4 provided evidence that the relevant associations can be effective in controlling motor responding even when subjects report an antagonistic relationship between the events.

It has long been thought (see, e.g., James, 1890) that prior training can modify the ease with which two stimuli can be discriminated from one another. In particular, training in which both have been associated with a common event has been said to retard subsequent discrimination (or to enhance generalization) between them—the acquired equivalence effect. Discrimination training, in which the cues are associated with differing outcomes, is thought to enhance performance on further tasks requiring discrimination between them—the acquired distinctiveness effect. These effects were first investigated experimentally for animal discrimination learning by Lawrence (1949), although the bulk of subsequent research has been carried out with human subjects (for a review, see Hall, 1991). It must be acknowledged that many of these later experiments produced results that were not wholly convincing as demonstrations of the effects of interest; also that they failed to shed much light on the mechanisms responsible. Students of animal discrimination learning, however, have recently returned to the issue, making use of experimental designs that successfully establish the reality of acquired equivalence/distinctiveness effects for the species studied (e.g., see Kaiser, Sherburne, Steirn, & Zentall, 1997, for work with pigeons; see also Bonardi, Rey, Richmond, & Hall, 1993). Further, theoretical developments in the analysis of associative learning in animals (e.g., Hall, 1996; Holland, 1990) suggest a possible account of the mechanisms involved. The aim of the work reported here is to confirm the reality of acquired equivalence and distinctiveness effects in human subjects using the experimental design successfully used with animals and to extend the procedure to test the implications of a theoretical interpretation derived from theories based on studies of animal conditioning.

Earlier studies with human subjects concentrated, for the most part, on the acquired distinctiveness effect and made use of a two-stage, transfer-of-training procedure, in which subjects learn to apply different verbal labels to stimuli in the first stage followed by a discrimination involving overt motor responses to the same stimuli in the second (e.g., Gagné & Baker, 1950). It was usually found (for reviews, see, e.g., Cantor, 1965; Gibson, 1969; Hall, 1991) that Stage 1 training facilitated learning of the second task. This result is of potential theoretical importance—it provides, for example, the empirical basis for the proposal that the way in which events have been categorized will influence perceptual discrimination (see Goldstone, 1994, 1998). It remains to establish, however, that the effect obtained in these experiments is in fact a specific consequence of the discrimination training given with the stimuli in the first stage. In the experiment by Gagné and Baker (1950), and in several others (e.g., Battig, 1956; Goldstone, 1994; Holton & Goss, 1956), comparison was made with a control condition given no Stage 1 training, leaving open the possibility that the positive transfer obtained in the experimental condition arose simply because the first stage of training produced *general* facilitatory effects.

What is needed is to compare the effects of discrimination training with those of some other Stage 1 training procedure that is equally effective in producing general transfer effects but which does not result in the stimuli becoming linked to distinctively different associates. There are a few reports of experiments that

Geoffrey Hall and Steven Graham, Department of Psychology, University of York, York, United Kingdom; Chris Mitchell, Unilever Research, Wirral, United Kingdom; Yvonna Lewis, School of Psychology, University of New South Wales, Sydney, New South Wales, Australia.

Chris Mitchell is now at the School of Psychology, University of New South Wales, Sydney, New South Wales, Australia; Steven Graham is now at the National University of Singapore and Cognitive Neuroscience Laboratory, Singapore General Hospital, Singapore.

This work was supported by Unilever Research. We thank C. Bonardi and P. Quinlan for their insightful comments and criticisms.

Correspondence concerning this article should be addressed to Geoffrey Hall, Department of Psychology, University of York, York YO10 5DD, United Kingdom. E-mail: gh1@york.ac.uk

attempt to achieve this, although none has been wholly without problems. One strategy has been to give the control subjects a same/different discrimination in Stage 1, using the same stimuli as those to which the experimental subjects were required to apply different verbal labels (Norcross & Spiker, 1957; see also Kurtz, 1955). The problem with this procedure is that the different Stage 1 tasks are unlikely to be matched in the demands they make on the subjects, and a difference in the difficulty of these tasks might be enough in itself to produce a difference between the groups in their test performance. The strategy used by Goldstone (1994) was to present compound stimuli in the first stage of training, one aspect of the stimuli (e.g., their size) being irrelevant to the discrimination being trained (this being based, e.g., on differences in brightness), so that when it came to the test phase (involving a further brightness discrimination), all subjects were familiar with the stimuli and all had received discriminative pretraining. The results obtained indicated an acquired distinctiveness effect of a sort, in that the test task was performed more readily by those subjects for whom the same dimension (*brightness* in this example) was relevant in both stages. This effect did not depend, however, on the test stimuli having been linked with different associates in the first stage, as superior performance was obtained even when the two brightnesses to be discriminated on test had been associated with the same response in Stage 1 training.

The experiments to be reported here adopt a different strategy, a version of which has previously been used in studies of human discrimination learning by Reese (1972) and by Norcross (1958; see also Grice, 1965). It is based directly on that used in studies of animal discrimination learning, with the design of Experiment 1 being conceptually identical to that used in the experiments by Bonardi et al. (1993) and Kaiser et al. (1997). In the second of these experiments, pigeons received initial training with four stimuli, *A, B, C,* and *D,* in which presentations of *A* and of *B* were each followed by a further stimulus, *X,* and presentations of *C* and *D* were followed by a different stimulus, *Y.* In the next stage, the birds were required to learn a successive discrimination involving the original four stimuli. The discrimination was acquired readily when the reinforced stimuli were *A* and *B* and the nonreinforced stimuli were *C* and *D* (an arrangement consistent with the treatment given to the stimuli in the first stage). Acquisition occurred more slowly in the inconsistent condition, with *A* and *C* reinforced, and *B* and *D* nonreinforced. These results were interpreted as showing that generalization occurred readily between stimulus pairs that had been treated in the same way in the first stage of training (the acquired equivalence effect), or that generalization was poor between pairs that had been treated differently in the first stage (the acquired distinctiveness effect), or both. It must be acknowledged that a shortcoming of this experimental design is that it cannot tell us whether the effect observed depends on positive transfer in the distinctiveness condition, negative transfer in the equivalence condition, or both of these. However, this design, unlike that used in many previous experiments, does allow the conclusion that at least one of these effects must be a reality. Furthermore, this design can be exploited to test an interpretation of the effect that can be derived from associative learning theory.

Experiment 1

In Experiment 1 (the design of which is summarized in Table 1), subjects received initial training with four visual stimuli (shapes), two of which (*A* and *B*) were followed by one outcome (presentation of a red rectangle), and two (*C* and *D*) followed by another (a green rectangle). No overt response was required in Stage 1. All subjects then underwent Stage 2 discrimination training in which the same shape stimuli were used to indicate which of two different motor responses should be made. The focus of interest was the performance of these subjects on two different versions of the Stage 2 task. Subjects in the inconsistent condition were required to discriminate between stimuli that had had different associates in Stage 1; in the consistent condition, the discrimination required was between stimuli that had shared an associate. Superior performance in the consistent condition would constitute evidence for the operation of an acquired equivalence/distinctiveness effect in this situation.

The Stage 2 result of Experiment 1 (to anticipate) successfully confirmed the reality of the effect in this training situation. In a final stage of testing, we went on to test an associative interpretation of this effect. At first sight, associative principles may seem ill-fitted to explain transfer effects of this sort. Associative learning seems likely to occur in the first stage of training; that is, the subjects will form associations between *A* and red and between *B* and red, between *C* and green and between *D* and green. However, the design used in experiments of this type was specifically introduced with the intention of ensuring that the associations acquired in the first phase of training would be irrelevant to the solution of the second discrimination with its new response requirement; transfer must, it was suggested, be based on some other learning process. Such was the argument put forward by Lawrence (1949) in explaining his demonstration of acquired distinctiveness in rats (see also Sutherland & Mackintosh, 1971). This other learning process was taken to be one that modulates the attention paid to various aspects of the stimuli. In the Stage 1 task, the subject might learn to focus attention on some feature shared by *A* and *B* but not

Table 1
Design: Experiment 1

Stage 1	Stage 2	Stage 3
	Group consistent	
A → red	*A* → left	
B → red	*B* → left	red → left/right?
C → green	*C* → right	green → left/right?
D → green	*D* → right	
	Group inconsistent	
A → red	*A* → left	
B → red	*B* → right	red → left/right?
C → green	*C* → left	green → left/right?
D → green	*D* → right	

Note. Feedback was given after responses in Stage 2. All subjects in a given group received all types of trial listed under a given stage of training. *A, B, C,* and *D* represent visual stimuli (see Figure 1) presented on a computer monitor; red and green refer to colored rectangles. Left and right refer to keyboard response required (left = *back slash;* right = *forward slash*).

by *C* and *D,* as this feature would uniquely predict the outcome of the trial. Similarly attention might be boosted to a feature shared by *C* and *D,* but not present in *A* or *B.* A tendency to attend to these features would put the consistent group at an advantage in Stage 2, as the features being attended to would reliably indicate which response should be made.

It turns out, however, that given certain assumptions, associative theory can be extended to explain both acquired equivalence and acquired distinctiveness. As long ago as 1941, N. E. Miller and Dollard suggested that training that attaches the same associate to two different stimuli (as when the same verbal label is applied to both) could provide a common element that might mediate generalization between them (see also Hull, 1939). N. E. Miller and Dollard expressed their idea in the stimulus–response terminology then current, but in more modern terminology, the associative account may be applied to the present procedure as follows. When, in Stage 2, the subject learns to perform a given response to a particular stimulus (e.g., to respond left to *A*) we assume that the presentation of *A* will evoke some central representation of the associate attached to it in Stage 1. This associate will thus also become a cue for performing that particular response. By virtue of its Stage 1 training, stimulus *B* will also evoke this same associate and thus a tendency to make the same response (i.e., to respond left to *B*) will be elicited immediately. For subjects in the consistent condition, this tendency will facilitate Stage 2 performance. But for subjects in the inconsistent condition (who are required to respond right to B), this tendency will need to be overcome and will detract from efficient performance on the test. Similar considerations will apply to stimuli *C* and *D*. Acquired equivalence between *A* and *B* and between *C* and *D* will thus generate a difference in test performance between the consistent and inconsistent conditions.

Central to this account is the assumption that the associatively activated representation of an event can acquire associative strength as a signal for some other event with which it is paired. Direct evidence to support the validity of this assumption has become available only recently, from work on animal conditioning by Holland (1990). In one of his experiments, Holland gave rats initial training in which presentation of a tone was followed by a particular type of food. The rats then received training in which the tone was associated with gastric nausea induced by injection of a mild toxin. A final test showed that the rats tended to reject the food-type that had been presented in the first phase of training. Holland concluded that an association had been formed in the second phase of training between the representation of the food (associatively activated by the tone) and the state of nausea. The effect has been referred to as *mediated conditioning,* with the acquisition of the aversion to the food being mediated by conditioning with another event (in this case the tone).

Stage 3 of the present experiment was designed to provide a test of a mediated-conditioning interpretation of the results of Stage 2. After completing Stage 2, all subjects were given trials on which the red and green rectangles were presented, and they were required to choose between the left and right response keys. If some central representation of an associated color is activated by the presence of a particular shape during Stage 2, then this representation of the color should, according to the theory under test, become associated with the keypress required for that shape. Thus it can be predicted that subjects in the consistent group should respond on one key (that associated with shapes *A* and *B*) for the red rectangle and on the other key (that associated with shapes *C* and *D*) for the green rectangle. Subjects in the inconsistent group, by contrast, received Stage 2 training in which there was no consistent correlation between the response required and the associate of the cue. There are thus no grounds to predict anything other than random responding by these subjects in Stage 3.

Method

Subjects. Twelve individuals between the ages of 18 and 40, the majority of whom were students at the University of York, York, United Kingdom, took part in the experiment. They were randomly allocated to two groups (n = 6): Group Consistent and Group Inconsistent.

Apparatus. Experimentation was conducted in a sound-attenuated testing room using a personal computer, complete with mouse and keyboard, and with the screen positioned at eye level, about 0.5 m from the subject. The geometrical shapes (stimulus set I, Figure 1) were black and 5 cm in height. The colored patches were presented as horizontal rectangles measuring 9 cm × 8 cm. All the stimuli appeared in the center of the screen on a gray background.

Procedure. The subject was initially informed that a series of visually distinct geometrical shapes would be presented on the screen. In Stage 1, each of the shapes appeared on the screen for 0.5 s, and at the offset of each shape, a colored rectangle. The subject was told to note which color followed which shape, and then to click on the colored rectangle using the left button of the electronic mouse to initiate the next trial. The subjects were informed that the speed of responding was not important during this stage. Following a 2-s intertrial interval (ITI), the shape for the next trial then appeared. Stage 1 comprised 32 trials (eight each of the four types depicted in Table 1), each of which consisted of one of the four geometrical shapes followed by one of the two colored rectangles. Two of the shapes (*A* and *B* in Table 1) were consistently followed by the red rectangle and the remaining two shapes (*C* and *D*) were followed by the green rectangle. Which of the shapes was assigned to a given trial type was determined at random. The sequence of trials was random.

At the end of Stage 1, instructions appeared on the screen telling the subjects that they would now be required to respond by pressing a key on the computer keyboard. During Stage 2 there were 32 trials (eight of each type) in which one of the geometrical shapes was presented on the screen. The subjects were required to categorize the geometrical shapes by press-

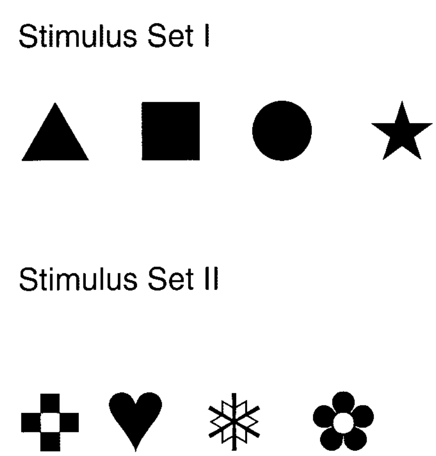

Figure 1. Experiments 1, 2, and 3 made use of stimulus set I; in addition, subjects in Experiment 2 also experienced stimulus set II. All stimuli were black figures on a gray background.

ing either the *back slash* key, which was located on the left of the keyboard (response "left" in Table 1), or the *forward slash* key, which was located on the right of the keyboard (response "right" in Table 1). At the outset of Stage 2, the subjects could not know which keypress was the appropriate response for each shape, but it was explained that feedback would allow them to learn this task. After the subject had pressed one of these keys, the stimulus disappeared and was replaced with a word—"Correct" or "Wrong"—during the 2-s interval that preceded presentation of the next shape. For half the subjects in the consistent group, the left response was required to shapes *A* and *B*, and the right response to shapes *C* and *D;* for the remaining subjects, the response assignments were reversed. For subjects in the inconsistent group, response to the left key was required after presentations of *A* and of *C,* and response to the right key after presentations of *B* and of *D.* The subjects were told that they should try to respond quickly, but as accurately as possible.

Stage 3 comprised eight trials (four of each type) in which one of the colored rectangles was presented on the screen. The subjects were required to categorize the geometrical shapes by pressing either the *back slash* key, on the left of the keyboard, or the *forward slash* key, on the right of the keyboard, just as in Stage 2. They were informed that no feedback would be given and that they should use their experience in the previous stages to decide which keypress was the most appropriate response. After the subject had pressed one of these keys, the stimulus disappeared and, following the 2-s ITI, the next shape appeared. The subjects were again told that they should try to respond quickly, but as accurately as possible.

Results and Discussion

All subjects completed Stage 1 without incident. No data were collected during this stage. The results of central interest are for the discrimination performance of the two groups during Stage 2. The upper part of Table 2 shows the group mean percent correct scores for this stage. It is evident that the consistent group showed a superiority over the inconsistent group, and this difference proved to be statistically reliable, $F(1, 10) = 9.08$, $MSE = 86.10$ $\eta^2 = .48$. (Here, and elsewhere a rejection criterion of $p < .05$ was adopted.) The mean latencies of correct responses were 0.71 s for the consistent group and 1.10 s for the inconsistent group, a difference that is not statistically reliable, $F(1, 10) = 2.55$. $MSE = 0.19$, $p = .14$. There is thus no reason to think the difference between the groups in the percent correct scores might be the consequence of a speed–accuracy trade-off.

The results for Stage 3 were analyzed according to whether the response was appropriate in terms of the training given in the previous stages. Thus, for example, some subjects in the consistent group had experienced the red rectangle following two of the shapes that were consistently associated with one of the keys, say on the left-hand side of the keyboard. For these, the appropriate Stage 3 response to the red rectangle was defined as pressing the left key; a response to the right key would be appropriate for the green rectangle. In scoring the responses for the inconsistent group, for half the subjects it was arbitrarily deemed that a response to the left given red should be deemed "appropriate," as should a response to the right given green; these assignments were reversed for the remaining subjects.

Table 2 gives group mean scores for "appropriate" responding in Stage 3. As might be expected, subjects in group inconsistent showed no consistent pattern of responding. Four subjects tended to choose one response key (some chose left and some chose right) and stick to it; the others appeared to respond quite at random. For the group as a whole, the mean score was 50%. Subjects in group consistent, however, made the appropriate choice in terms of their prior training, and the mean score was close to 100%. Statistical comparison of the two groups showed them to differ reliably, $F(1, 10) = 6.58$, $MSE = 958.33$, $\eta^2 = .39$. Group mean latencies over all eight trials of this stage were 0.76 s for group consistent and 1.32 s for group inconsistent; these scores did not differ reliably, $F(1, 10) = 3.19$, $MSE = 0.29$, $p = .10$.

The results for Stage 2 of this experiment confirm that discrimination training with a given set of stimuli will result in transfer to a new task involving the same stimuli, but having different response requirements. The nature or extent of this transfer was found to depend on the specific form of training given in the first phase. When the test was arranged so that the subjects had to make the same response to stimuli that had shared an associate in the first phase, performance was enhanced relative to that shown by subjects required to make different responses to such stimuli. We conclude that the first phase of training established either an acquired equivalence effect, an acquired distinctiveness effect, or both of these.

Table 2
Percent Correct Results

Experiment	Stage 2	Stage 3
Experiment 1		
Group consistent	96.35 (1.89)	95.83 (4.17)
Group inconsistent	80.21 (5.02)	50.00 (17.38)
Experiment 2		
Consistent condition	76.20 (4.24)	85.58 (5.45)
Inconsistent condition	67.31 (4.41)	46.15 (8.54)
Experiment 3		
Group consistent	93.23 (1.27)	Reversal: 89.58 (2.08)
Group inconsistent	88.02 (1.62)	Nonreversal: 96.88 (2.14)
Experiment 4		
Group consistent	91.67 (2.09)	
Group inconsistent	84.38 (2.20)	

Note. The standard error of the mean is shown in parentheses. The results for Stage 3 of Experiment 4 are presented in Figure 2.

The results for Stage 3 are what would be expected on the basis of the associative account of the transfer effects obtained in Stage 2. Central to this account, it will be recalled, was the proposal that generalization between the shape cues could be mediated by the acquisition of response-controlling power by representations of the outcomes that had been paired with these cues in Stage 1. It was assumed, for instance, that the representation of a red rectangle, activated associatively by the shape that had preceded it in initial training, would come to function as a cue for the response being trained to that shape. Stage 3 tested this idea by assessing explicitly what response the red rectangle (now actually presented) tended to evoke. The results for group consistent confirmed that the colors were capable of evoking the response expected on the basis of this account. Further implications of this analysis and a consideration of possible alternatives will be taken up in Experiment 4, after the next two experiments, which were designed to confirm the reliability and extend the generality of the findings of Experiment 1, have been described.

Experiment 2

This experiment was designed to allow a within-subject comparison of the consistent and inconsistent training procedures. The design is summarized in Table 3. All subjects experienced two sets of four shape stimuli in Stage 1 training. As in Experiment 1, associations were established between the shapes and colors, with a different pair of colors being used for each set of shapes. In Stage 2, the subjects were trained on the left–right discrimination task in the consistent arrangement for one set of shapes and in the inconsistent arrangement for the other set. Replication of the transfer effect obtained in Experiment 1 would be evident as more accurate performance in the consistent condition than in the inconsistent condition. In Stage 3, the subjects were asked to make a left–right choice in response to each of the colors. On the basis of the results obtained for the final test in Experiment 1, it might be expected that each of the colors trained in the consistent condition would be able to evoke the response trained in Stage 2 to the shape with which that color had been associated in Stage 1.

Table 3
Design: Experiment 2

Stage 1	Stage 2	Stage 3
	Consistent condition	
$A \rightarrow$ red	$A \rightarrow$ left	
$B \rightarrow$ red	$B \rightarrow$ left	red $\rightarrow$ left/right?
$C \rightarrow$ green	$C \rightarrow$ right	green $\rightarrow$ left/right?
$D \rightarrow$ green	$D \rightarrow$ right	
	Inconsistent condition	
$E \rightarrow$ yellow	$E \rightarrow$ left	
$F \rightarrow$ yellow	$F \rightarrow$ right	yellow $\rightarrow$ left/right?
$G \rightarrow$ purple	$G \rightarrow$ left	purple $\rightarrow$ left/right?
$H \rightarrow$ purple	$H \rightarrow$ right	

Note. Feedback was given after responses in Stage 2. All subjects received all types of trial listed under a given stage of training. Letters *A–H* represent eight different visual stimuli (see Figure 1); red, green, yellow, and purple refer to colored rectangles presented on a computer monitor. Left and right refer to keyboard response required (left = *back slash;* right = *forward slash*).

Method

The subjects were 13 individuals from the same population as served in Experiment 1. The apparatus was identical to that used in Experiment 1. In addition to the shapes (Figure 1, set I) used in Experiment 1, a second set of shapes (Figure 1, set II) was used. Set II also consisted of black geometrical shapes 5 cm high. As well as the red and green rectangles used in Experiment 1, yellow and purple rectangles could also be displayed. One set of stimuli was used as stimuli *A–D* of Table 3; that is, they were trained in the consistent condition with red and green rectangles as their associates. The other set were trained in the inconsistent condition (stimuli *E–H* of Table 3), with purple and yellow rectangles as the associates. For 6 subjects, set I was used for stimuli *A–D* and set II for *E–H;* for the other subjects, the reverse arrangement was used. Stages 1 and 2 each consisted of 64 trials (eight of each of the trial types shown in Table 3). Stage 3 consisted of 16 trials (four of each type). In details not specified here, the procedure was the same as that described for Experiment 1.

Results and Discussion

Group mean scores for the Stage 2 discrimination are shown in Table 2. Although the absolute size of the effect was small, the performance shown in the consistent condition was significantly superior to that shown in the inconsistent condition, $F(1, 12) = 4.90$, $MSE = 104.85$, $\eta^2 = .22$. Mean latencies for correct responses (1.43 s for the consistent condition and 1.36 s for the inconsistent condition) did not differ significantly ($F < 1$). This outcome matches that obtained in Experiment 1 and indicates the operation of acquired equivalence/distinctiveness effects in this within-subject procedure.

Table 2 also shows "correct" mean scores for the Stage 3 task. During Stage 2, each of the shapes that had previously been associated with one of the consistent colors (red and green) was consistently associated with a given response. There was thus an appropriate response available for the subject to make when these colors were presented in Stage 3. For these colors, a choice was scored as "correct" according to the feedback that had been received in Stage 2. In scoring Stage 3 responding to the inconsistent colors (purple and yellow), the assignment of a color to a given response was made at random for each subject. Appropriate responding was found to occur only to the consistent colors. Comparing the consistent and inconsistent scores summarized in the figure revealed a significant difference between them, $F(1, 12) = 12.96$, $MSE = 779.25$, $\eta^2 = .39$. The mean response latencies to the two sets of stimuli (2.02 s for consistent stimuli and 1.51 s for inconsistent stimuli) did not differ reliably, $F(1, 12) = 1.30$, $MSE = 1.29$, $p = .28$. Thus, as in Experiment 1, this result suggests that training a stimulus as a cue to perform a particular response can endow an associate of the trained stimulus with the same response-eliciting properties.

Experiment 3

The procedure used in Experiments 1 and 2 differed from that used in most previous studies of these effects in human subjects. Although it is possible (even likely) that the subjects applied a verbal label to the events used as the associates in the first stage of training, verbal labeling was neither encouraged nor required. Many earlier studies have been explicitly concerned with the effects produced by initial training in which verbal labels are learned for the target stimuli. In this experiment, we modified our procedure (using a version of the design of Experiment 1) to promote the likelihood that verbal labeling would occur during the first stage of training. We hoped thereby to extend the generality of the effects obtained in our earlier experiments and to establish a parallel with previously published studies.

As Table 4 shows, the design of the first two stages of this experiment was identical to that used in Experiment 1. Thus, all subjects received the same training with a set of four shapes as the stimuli in Stage 1; they were then divided into consistent and inconsistent groups given different versions of the Stage 2 discrimination task. The only novel feature was that the events presented as trial outcomes in Stage 1 were not colored rectangles but pronounceable nonsense syllables.

As before, Stage 3 was intended to assess the extent to which Stage 2 training had endowed events not actually present (the nonsense syllables in this case) with the power to elicit the motor responses trained to the shape stimuli in that stage. Although the Stage 3 procedure used in the preceding experiments yielded satisfactory results, we were unhappy about certain aspects of this procedure—in particular, some subjects in the inconsistent condition appeared puzzled and disconcerted by the demands of the task. Accordingly, we decided in this experiment to concentrate on the consistent group and to use a savings test for them in Stage 3. All subjects in the consistent group were required to learn a new discrimination in which each of the nonsense syllables signaled that one or other of the left–right response keys should be pressed. For half the subjects (the reversal condition of Table 4), the mapping of stimulus onto response was the opposite of the association putatively formed during Stage 2. For the remaining subjects (the nonreversal condition), the putative associations matched the Stage 3 response requirements. A superiority of the nonreversal over the reversal condition in this stage would suggest that associations had been formed in Stage 2 between some representation of the nonsense syllables and the response being trained in that stage.

Table 4
Design: Experiment 3

Stage 1	Stage 2	Stage 3
	Group consistent	
$A \rightarrow$ *wug*	$A \rightarrow$ left	Reversal: *wug* $\rightarrow$ right
$B \rightarrow$ *wug*	$B \rightarrow$ left	*zif* $\rightarrow$ left
$C \rightarrow$ *zif*	$C \rightarrow$ right	Nonreversal: *wug* $\rightarrow$ left
$D \rightarrow$ *zif*	$D \rightarrow$ right	*zif* $\rightarrow$ right
	Group inconsistent	
$A \rightarrow$ *wug*	$A \rightarrow$ left	
$B \rightarrow$ *wug*	$B \rightarrow$ right	
$C \rightarrow$ *zif*	$C \rightarrow$ left	
$D \rightarrow$ *zif*	$D \rightarrow$ right	

Note. Feedback was given after responses in Stage 2 and Stage 3. All subjects in a given group received all types of trial listed under a given stage of training. *A, B, C,* and *D* represent visual stimuli (see Figure 1); the nonsense syllables *wug* and *zif* were presented on a computer monitor. Left and right refer to keyboard responses required (left = *back slash;* right = *forward slash*).

Method

Twenty-four subjects, undergraduate students at the University of York, took part in the experiment. They were randomly allocated to two groups ($n = 12$), the consistent and inconsistent groups. The shapes used as stimuli *A–D* were set I of Figure 1. The nonsense syllables used in Stage 1, *wug* and *zif*, were presented centrally on the screen in 72-point Comic Sans MS font. The procedure for Stages 1 and 2 was identical to that described for Experiment 1, except that in Stage 1 the red and green rectangles used in the earlier experiment were replaced by the nonsense syllables.

Subjects in the consistent group received Stage 3 training. This consisted of 16 trials in which one of the nonsense syllables was presented on the screen. The subjects were required to categorize these labels by pressing either the *back slash* or *forward slash* key, as in Stage 2. Feedback was given; after the subject had pressed one of the keys, the stimulus disappeared and was replaced by a word—“Correct!” or “Wrong!”—during the 2-s ITI before the next syllable appeared. The subjects were again told that they should try to respond as quickly but as accurately as possible. For the 6 subjects in the nonreversal condition, the relationship between the required response and the stimulus was the same as that implied by the Stage 2 training they had received. The remaining 6 subjects (the reversal condition) experienced a reversed relationship between stimulus and required response.

Results and Discussion

Table 2 presents group mean percent correct scores for the Stage 2 task. Both groups learned the discrimination readily, but the small advantage shown by the consistent group proved to be statistically reliable, $F(1, 22) = 6.39$, $MSE = 25.45$, $\eta^2 = .23$. Mean latencies for correct responses were marginally shorter in group consistent (0.80 s) than in group inconsistent (0.83 s), but this difference was not statistically reliable ($F < 1$). These results are thus entirely in accord with those reported for the equivalent stage in Experiments 1 and 2 and indicate the operation of an acquired equivalence/distinctiveness effect in this training paradigm in which the mediating event is a verbal label.

The Stage 3 results confirmed that the verbal label had acquired the power to control the left–right keypress response. Table 2 shows group mean correct responses for the reversal and nonreversal groups of the subjects trained previously in the consistent condition. As in Stage 2, the task proved very easy and both groups performed well; but again, the small difference between the groups proved to be statistically reliable, $F(1, 10) = 5.98$, $MSE = 26.69$, $\eta^2 = .37$. The difference between the groups is what would be expected if some representation of the nonsense syllables had, during Stage 2, entered into an association with the motor response emitted in that stage. Such a preexisting association would produce a savings effect for group nonreversal and would interfere with learning for group reversal. The mean latencies for correct response in this stage of training were marginally longer for group nonreversal (0.70 s) than for group reversal (0.65 s), but the difference was not statistically reliable ($F < 1$).

Experiment 4

The results of the preceding experiments have been interpreted in terms of an associative account, which holds that when, in Stage 2, the participant learns to make a particular response to the stimulus presented on a given trial, the ability of an associate of that stimulus to control that same response will be enhanced. That is, for example, following *A*–red pairings in Stage 1 of Experiment 1, *A*–left pairings in Stage 2 appeared to result in the formation of an association between *A*'s associate, red, and the left response. Support for this interpretation was provided by a final test stage in each experiment that showed that the associate was indeed capable of eliciting the appropriate response. Thus, subjects trained with *A*–red in Stage 1 and *A*–left in Stage 2 responded with a left keypress when presented with the red rectangle in Stage 3.

It should be acknowledged, however, that another interpretation of the Stage 3 results of these experiments may be possible. When, for example, the subjects in the consistent group of Experiment 1 responded left to the red rectangle in Stage 3, they could have been using some linguistically based problem-solving mechanism ("because red goes with the circle, and the circle goes with left, then, perhaps, red goes with left"). Whether people do in fact reason in this way is open to debate. Indeed, there are grounds for thinking that the training regime used in Stages 1 and 2 might lead them to draw exactly the opposite inference—having been trained in Stage 1 that red follows the circle, they might conclude from Stage 2, when some other event follows the circle, that this other event actually prevents the occurrence of the red cue that would otherwise be forthcoming. They might conclude, therefore, that red and left explicitly do not "go together." That our subjects, during debriefing, did not reliably report making either of these inferences does not constitute decisive evidence one way or the other—our debriefing procedure was quite informal and may have been inadequate as a means of detecting what the subjects knew, or thought they knew, about the relationships among the various events. Accordingly, in the present experiment we included a final test designed to provide a formal assessment of their interpretation of the relationship between the relevant events.

A further important feature of the design of this experiment was the use of a Stage 1 training procedure intended to enhance the likelihood that subjects would make the inference that an antagonistic relation held between two events trained separately as outcomes of the presentation of a particular cue. In our previous experiments, subjects were trained in one stage with a given cue (such as cue *A* in the tables) followed by one event (such as the red rectangle), and then, in a second separate block of trials, they received training in which *A* was followed by a different event (e.g., response to the *back slash* key). In the present experiment the two associations involving cue *A* were trained concurrently; that is, the subjects experienced a mixture of trials, with *A* being followed by a color on some and by a nonsense syllable on others. We thought this procedure would make it especially apparent that the color and the syllable were alternative outcomes of the presentation of *A* and might lead subjects to form the conclusion that the occurrence of one prevented the occurrence of the other. We then tested for this preventive relationship directly in a final test that constituted Stage 3 of the experiment.

Whatever the Stage 3 test might reveal about the inferences drawn by the subjects, it should be noted that the procedure used in Stage 1 should still allow mediated conditioning to occur. Once training is under way and associations have begun to form, presentation of cue *A* will activate the representations of its associates so that, for instance, the representation of the syllable will be active on a trial on which the color follows *A*; similarly, the representation of the color will be activated on a trial on which the syllable occurs. In these circumstances, according to the associative account being developed here, excitatory associations will form between the color and the syllable. We tested for the existence of these associations in a second stage that used, as in the previous experiments, a discrimination task involving different motor responses. The cues used were those trained as associates in Stage 1. Subjects assigned to the consistent condition had to learn to make the same response to the color and the syllable; those assigned to the inconsistent condition had to make different responses to these two cues. Excitatory associations between the color and the syllable are predicted to enhance learning in the former condition and to retard it in the latter. The question of interest was whether such a difference could be obtained in subjects given a form of initial training that, we expected, might lead them to report an antagonistic relation between the color and the syllable.

The full experimental design is presented schematically in Table 5. All subjects received Stage 1 training with two geometrical figures (*A* and *B*), each of which was followed by a given color on some trials and by a given nonsense syllable on other trials. In Stage 2, the colors and the nonsense syllables served as the cues signaling which motor response should be made. For half the subjects (the consistent group), one response was required to the color and syllable that had been associated with *A*, and the other response was required to the color and syllable that had been associated with *B*. For the remaining subjects (the inconsistent group), different responses were required to the stimuli that had been associated with a common Stage 1 shape. In Stage 3, all subjects were presented with a color and asked to judge whether this color had previously prevented the occurrence of one or other of the nonsense syllables.

Method

The subjects were 30 psychology students from the University of New South Wales, Sydney, Australia, who volunteered for the experiment in return for course credit. They were randomly allocated to two equal-sized groups, the consistent and inconsistent groups.

Table 5
Design: Experiment 4

Stage 1	Stage 2	Stage 3
	Group consistent	
A → red	red → left	
A → *wug*	*wug* → left	red prevents *wug/zif?*
B → green	green → right	green prevents *wug/zif?*
B → *zif*	*zif* → right	
	Group inconsistent	
A → red	red → left	
A → *wug*	*wug* → right	red prevents *wug/zif?*
B → green	green → right	green prevents *wug/zif?*
B → *zif*	*zif* → left	

Note. Stimulus and response combinations were counterbalanced; those given above constitute one example. All subjects in a given group received all types of trial listed under a given stage of training. Feedback was given after responses in Stage 2. *A* and *B* represent geometrical shapes; the nonsense syllables *wug* and *zif* were presented on a computer monitor and spoken out loud. Left and right refer to keyboard responses required.

Apart from the exceptions specifically noted below, the procedure was the same as that described for Experiment 1. The stimuli were two geometrical shapes (a triangle and a snowflake), two colored patches (red and green), and two nonwords (*wug* and *zif*). The geometrical shapes were black and approximately 10 cm in height and width. The colored patches were presented as horizontal rectangles measuring 9 cm × 4 cm. The nonwords were in bold, 72-point Arial font, and all letters were lowercase. All the stimuli appeared in the center of the screen on a white background.

Stage 1 comprised 80 trials, each of which consisted of one of the two geometrical shapes followed by one of the two nonwords or one of the two colored patches. Subjects were initially informed that some visually distinct geometrical shapes would be presented on the screen and that each shape would be closely followed by either a nonword or a colored patch. They were asked to pay close attention to each shape and what followed it, and that when a nonword appeared, they should pronounce it out loud. The first 16 trials of Stage 1 consisted of nonword trials only, whereas the remaining 64 trials consisted both of nonword and colored-patch trials in equal numbers. The sequence of trials within these two substages was random. Throughout the stage, a given subject always experienced the same color and the same syllable after presentation of a particular shape, but all stimulus combinations were used and the specific combination assigned to each participant was determined at random.

At the end of Stage 1, instructions appeared on the screen telling the subjects that they would now be required to respond by pressing one of two keys on the computer keyboard. During Stage 2 there were 32 trials (eight of each type) in which one of the colored patches or one of the nonwords was presented on the screen. The subjects were required to categorize the stimuli by pressing either the *A* key, which was located on the left of the keyboard, or the *5* key (on the number pad), which was located on the right of the keyboard. Responses with reaction time greater than 2 s were scored as errors. As in previous experiments, feedback was given during the 1-s interval that separated trials. For subjects in the consistent group, the colored patch and nonword that had been preceded in Stage 1 by the same shape had to be categorized on the same side. For subjects in group inconsistent, the colored patch and nonword that had been preceded by opposite shapes had to be categorized on the same side. The assignment of the relevant stimuli to each side was determined at random.

Stage 3 consisted of just two trials, one on which the red patch appeared and one on which the green patch appeared. On both trials, a 5-point rating scale was present at the bottom of the screen. Point 1 was labeled *wug*, Point 5 was labeled *zif*, and Point 3 was labeled *don't know*. The five points on the scale were evenly spaced. At the beginning of Stage 3, instructions appeared informing the subjects that in the first (training) stage of the experiment, the appearance of a particular colored patch had actually prevented the appearance of a particular nonword. Subjects were required to use the mouse to rate which nonword was prevented by the colored patch on the screen. They were instructed to click on the number that best reflected their level of certainty about which nonword was correct.

Results and Discussion

No data were recorded during Stage 1. Table 2 presents group mean percent correct scores for the Stage 2 task in which the subjects had to learn to respond left or right to the colors and the syllables. The table shows that the inconsistent group (required to make different responses to events that had shared an antecedent in Stage 1) performed less well than did the consistent group (required to make the same response to events that had shared an antecedent). The difference between the groups was statistically reliable, $F(1, 28) = 5.79$, $MSE = 68.82$, $\eta^2 = .17$. Group inconsistent also responded more slowly than group consistent; the mean response time was 1.02 s for the former group and .75 s for the latter, $F(1, 28) = 7.08$, $MSE = 79.52$, $\eta^2 = .20$. This pattern of results is what would be expected if mediated conditioning had occurred during Stage 1, establishing excitatory associative links between the color and the syllable that had been trained as associates of a common cue in that stage. It is worth noting that an essentially equivalent effect has previously been demonstrated for nonhuman subjects by Hall, Ray, and Bonardi (1993). In their study, rats received initial training in which two auditory cues (e.g., a tone and white noise) were each preceded by the same event (the delivery of a food pellet). The noise was then paired with shock, and it was found that the conditioned response established by this treatment generalized readily to the tone. This result confirms the conclusion prompted by the present experiment that stimuli that have shared a common antecedent will come to be treated as equivalent.

The results reported by Hall et al. (1993) are enough to establish that this version of the acquired equivalence effect can be obtained in the absence of linguistically based inferential processes. The results from Stage 3 of the present experiment (see Figure 2) indicate that such processes were not at work here. The rating scale used in Stage 3 ran from 1 (*prevents wug*) to 5 (*prevents zif*). For each of the subjects, one of the colors (C1 in Figure 2) had shared a common antecedent with *zif* in Stage 1; the other color (C2 in Figure 2) had shared an antecedent with *wug*. Thus, if the participant thought that the color prevented the occurrence of the syllable with which it had shared an antecedent, rating scores should be high for C1 and low for C2. Figure 2 shows just this pattern, especially in subjects in the consistent group. An analysis of variance conducted on the data summarized in the figure showed there to be a significant difference between the scores for C1 and C2, $F(1, 28) = 5.74$, $MSE = 5.62$, $\eta^2 = .17$. There was no

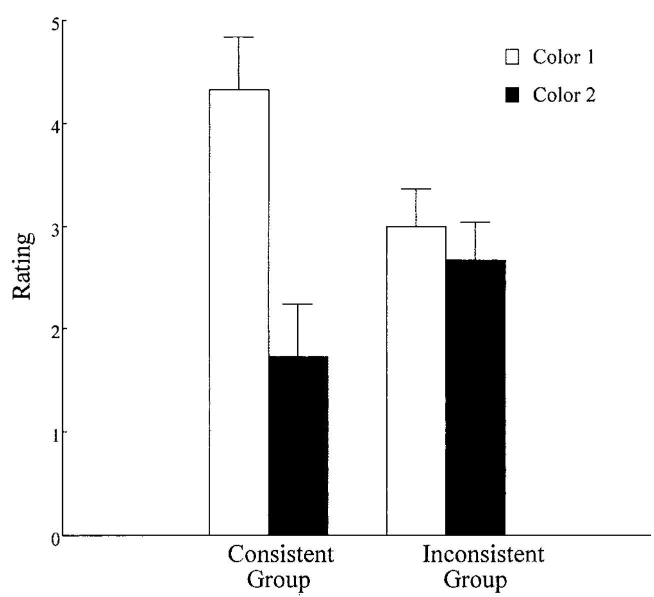

Figure 2. Mean ratings in Stage 3 of Experiment 4 in response to presentation of a color. A rating of 5 indicated *prevents zif*; a rating of 1 indicated *prevents wug*; a rating of 3 indicated *don't know*. For all subjects, Color 1 had shared an antecedent with *zif* in Stage 1 and Color 2 had shared an antecedent with *wug*. Vertical bars represent standard errors of the mean.

significant effect of group, $F(1, 28) = 2.07$, $MSE = 0.29$, $p = .16$, and the interaction between the variables also fell short of significance, $F(1, 28) = 3.43$, $MSE = 5.62$, $p = .08$. It may be noted that the difference between the scores for C1 and C2 was substantially less in the inconsistent than in the consistent group, perhaps implying that the training regime experienced by the former group during Stage 2 may have disrupted such inferences as had been formed in Stage 1. Because, however, statistical analysis did not confirm the reliability of this difference between the groups, we will speculate no further on his matter.

The results of the Stage 3 test demonstrate that the subjects judged there to be an antagonistic relation between the two events that had shared an antecedent in Stage 1, and this, in spite of the fact that their discrimination performance in Stage 2, was consistent with the interpretation that excitatory associations had been established between these events. The immediate conclusion justified by this dissociation is that the performance shown on the motor task of Stage 2 is not to be explained in terms of the inferential processes that generated the Stage 3 results.

Further implications of the dissociation between the two measures will be taken up in the General Discussion, but it is of interest to note at this point the parallel between our results and effects obtained in studies of explicit and implicit attitudes in social psychology. For instance, it has been suggested that attitudes (e.g., racist attitudes) that are explicitly denied, can sometimes be revealed using implicit (indirect) measurement techniques (see Wilson, Lindsey, & Schooler, 2000, for a review). One such indirect measure is the Implicit Attitude Test (IAT) developed by Greenwald, McGhee, and Schwartz (1998), a measure that has much in common with the tests used in the present experiments. In the IAT, subjects' preferences are inferred from the speed and accuracy with which they can respond to the target while at the same time responding to other items of known affective valence. For example, subjects might perform well at a task (record a lower response time and make fewer errors) in which the same response is assigned to the target object and a group of known pleasant items. In the IAT, the relative ease (against appropriate controls) with which the target object and the pleasant items are classified together is taken as a measure of liking for that object. Dual attitudes have been revealed using this technique by the finding that people who deny having a preference for White Americans over Black Americans may nevertheless demonstrate such a preference when tested using the IAT. The parallel with our own results is almost exact. When explicitly asked about the relationship between two events, our subjects assert that this is negative; the implicit measure provided by the classification task of Stage 2 of our experimental procedure, however, shows that they have formed a positive association between these events.

General Discussion

The experiments reported here, using designs previously used in experiments with animal subjects, have successfully demonstrated acquired equivalence/distinctiveness effects in human discrimination learning. Performance on a discrimination task involving a set of familiar stimuli was critically influenced by the prior training that these stimuli had received, even when the response requirements of the training and test tasks were quite different. When, as in Experiments 1–3, the test required subjects to perform the same motor response to stimuli that had shared a common consequence in Stage 1, performance was enhanced, in comparison with a condition in which different motor responses were required to these stimuli. This effect was found when, in Stage 1, shape stimuli were followed either by the same color patch (Experiments 1 and 2) or the same nonsense word (Experiment 3). Experiment 4 demonstrated a similar effect when the critical cues had shared a common antecedent in Stage 1—again, performance was better in the condition in which the same response was required to each of these cues than in the condition in which different responses were required.

As was briefly outlined in the discussion of Experiment 1, the results of these experiments are readily explained in terms of the principles of associative learning theory, provided it is allowed that the associatively activated representation of an event is capable of being learned about just as can the event itself. This assumption implies that when the subject learns to make a given response to a stimulus that is physically present, the ability of an event associated with that stimulus to govern that response will be increased. For example, if a triangle and a square are separately paired with the red color patch in Stage 1, a representation of that red color patch will be activated when either shape is later presented. Therefore, in Stage 2, if a left response is made to the triangle, an excitatory link will form between the associatively activated red color patch and that left response. As a consequence, when the square is presented in Stage 2, it will activate a representation of red, and, through this representation, a tendency to respond left. Such transfer of control over responses between the two shapes will aid performance in the consistent condition but hinder it in the inconsistent condition. To support this interpretation, it is necessary to show that associatively activated representations can indeed acquire associative strength. Experiments 1–3 included a final test stage designed to do this; and in each experiment it was found that the associate (not directly trained) was capable of eliciting the response that had been directly trained to the cues that had been used to signal it in a first stage of training. Thus, the red color patch in the example cited above proved to be capable of eliciting the left response in Stage 3, following triangle-left and square-left pairings in Stage 2.

The associative analysis of Experiment 4 is only slightly different. In this case, associatively mediated learning occurred in Stage 1, rather than between Stages 1 and 2. Thus, when the triangle was paired on different trials in Stage 1 with both the syllable *wug* and the red color patch, it was hypothesized that presentations of the triangle on later trials would activate representations of both *wug* and red. As a result, an excitatory link was expected to form on these trials between the representation activated associatively through presentation of the triangle (*wug* or a red color patch), and the stimulus actually presented following the triangle (the red color patch or *wug*, respectively). This association between *wug* and red can be expected to aid performance in Stage 2 for subjects required to make the same response to *wug* and to red (the consistent group) but not for those required to make different responses (the inconsistent condition). This is what was found in Stage 2 of Experiment 4.

In all the experiments reported here, the critical manipulation involved training that may be symbolized as *A–X* and *A–Y,* in which a given cue was followed, on different trials, by different events (the two trial types occurred in separate stages in Experiments 1–3, and concurrently in Experiment 4). Our analysis of the transfer effects obtained after such training has been in terms of the associations presumed to be formed among the cues present on a given trial and the associatively activated representations of the absent cues. We acknowledged the possibility, however, that a language-based problem-solving mechanism might be engaged by training of this sort—that subjects might infer, for instance, that *X* and *Y* go together, in some sense, on the basis of their experience with the *A*–*X* and *A*–*Y* pairings. The direct test supplied by Experiment 4, gave no support for this proposal. When asked specifically for their views about the relationship between *X* and *Y*, subjects reported that one tended to prevent the occurrence of the other—this in spite of the fact that their performance on a discrimination task involving these stimuli suggested the existence of excitatory associations between them. We conclude that the processes that govern the formation and operation of associative links between events may be governed by principles different from those that control a person's ability to make a verbal report about the relationship between events. This dissociation has relevance to the issue to be discussed next.

As we have already noted, the associative interpretation of the transfer effects obtained in the present experiments depends on the proposal that the associatively activated representation of an event will form an excitatory association with the (directly activated) representation of some other event with which it is paired. Indeed, the experimental results reported here can be seen as providing evidence in favor of the validity of this assumption, adding to that available from a number of studies of animals (e.g., Holland, 1990; Ward-Robinson & Hall, 1996, 1999; see Hall, 1996, for a review). What remains to be resolved, however, is why other training procedures, different in many details but conceptually similar to those used here, have generated what appear to be opposite effects. For example, Dickinson and Burke (1996), in a study of human causal judgment, gave their subjects initial training in which a compound cue, *AB*, was followed by a particular outcome. This should allow the formation of excitatory associations between *A* and *B*, and between each element of the compound and the outcome. In a second stage, *A* was presented alone, followed by the outcome, allowing the possibility that the representation of *B* would be activated (by way of the *A*–*B* association) concurrently with presentation of the outcome. According to the associative principles adopted here, this procedure should result in mediated conditioning with the *B*-outcome association becoming strengthened during Stage 2. What Dickinson and Burke found, however, in a final test, was that this treatment led to a *reduced* likelihood that the subjects would attribute the outcome to the occurrence of cue *B*.

One possible explanation for this apparent discrepancy is suggested by the results of the present Experiment 4. These results showed that the inferences a subject draws about the relationship between events in experiments of this type need not match (indeed, might seem to contradict) information encoded in associations that (on the basis of other measures) we have reason to believe have been formed. To draw the parallel between the present studies and social psychological studies of dual attitude (Wilson et al., 2000), it might be argued that our discrimination task (Experiments 1–4) is an "implicit" measure of subjects' view of the causal relationship between the cues, whereas our Stage 3 task in Experiment 4, and Dickinson and Burke's (1996) causal judgment task, is an "explicit" measure of that relationship. What follows is the possibility that in the Dickinson and Burke study, the second stage of training did indeed strengthen the *B*-outcome association, but that the nature of the test (which required a judgment about causality) tapped not associative strength, but the operation of inferential processes. Clearly, this speculation is in need of support from further experimental work. In particular, it would be worthwhile to devise some behavioral test of the strength of the *B*-outcome association generated by the Dickinson and Burke procedure, in order to test the implication that there may be a discrepancy between this strength and the result produced in an explicit judgment of causality. It is worth noting however, that the effect demonstrated by Dickinson and Burke (sometimes referred to as "backward blocking"), although widely obtained in studies of causal judgment (e.g., Chapman, 1991; Wasserman & Berglan, 1998), has proved to be very elusive in studies using conditioning procedures with nonhuman animals, usually failing to appear at all (e.g., R. R. Miller, Hallam, & Grahame, 1990), or doing so only under rather special conditions of training (R. R. Miller & Matute, 1996).

Finally, although the associative mediation account can provide a complete explanation for the results reported here, we cannot conclude that it is the sole source of acquired equivalence/distinctiveness effects. Other experiments, both with animal subjects (e.g., Delamater, 1998) and with humans (e.g., Goldstone, 1994), have generated these effects in experimental designs that seem to defy explanation in associative terms. Both these investigators have offered attentional explanations for their results, and it is possible that attentional processes of the sort they have postulated are operating in our training situation too. Our results demonstrate the reality of associative mediation, but do not rule out the possibility that attentional changes are also occurring. For the time being, the strongest conclusion we can reach is that an appeal to attentional mechanisms is unnecessary to explain the effects obtained in these experiments.

References

Battig, W. F. (1956). Transfer from verbal pretraining to motor performance as a function of motor task complexity. *Journal of Experimental Psychology, 51,* 371–378.

Bonardi, C., Rey, V., Richmond, M., & Hall, G. (1993). Acquired equivalence of cues on pigeon autoshaping: Effects of training with common consequences and with common antecedents. *Animal Learning & Behavior, 21,* 369–376.

Cantor, J. H. (1965). Transfer of stimulus pretraining in motor paired-associate and discrimination learning tasks. In L. P. Lipsitt & C. C. Spiker (Eds.), *Advances in child development and behavior* (Vol. 2, pp. 19–58). New York and London: Academic Press.

Chapman, G. B. (1991). Trial order affects cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 837–854.

Delamater, A. R. (1998). Associative mediational processes in the acquired equivalence and distinctiveness of cues. *Journal of Experimental Psychology: Animal Behavior Processes, 24,* 467–482.

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

Gagné, R. M., & Baker, K. E. (1950). Stimulus predifferentiation as a factor in transfer of training. *Journal of Experimental Psychology, 40,* 439–451.

Gibson, E. J. (1969). *Principles of perceptual learning and development.* New York: Appleton-Century-Crofts.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General, 123,* 178–200.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49,* 585–610.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74,* 1464–1480.

Grice, G. R. (1965). Investigations of response-mediated generalization. In D. I. Mostofsky (Ed.), *Stimulus generalization* (pp. 373–382). Stanford, CA: Stanford University Press.

Hall, G. (1991). *Perceptual and associative learning.* Oxford, England: Clarendon Press.

Hall, G. (1996). Learning about associatively activated representations: Implications for acquired equivalence and perceptual learning. *Animal Learning & Behavior, 24,* 233–255.

Hall, G., Ray, E., & Bonardi, C. (1993). Acquired equivalence between cues trained with a common antecedent. *Journal of Experimental Psychology: Animal Behavior Processes, 19,* 391–399.

Holland, P. C. (1990). Event representation in Pavlovian conditioning: Image and action. *Cognition, 37,* 105–131.

Holton, R. B., & Goss, A. E. (1956). Transfer to a discriminative motor task as a function of amount and type of preliminary verbalization. *Journal of General Psychology, 55,* 117–126.

Hull, C. L. (1939). The problem of stimulus equivalence in behavior theory. *Psychological Review, 46,* 9–30.

James, W. (1890). *The principles of psychology.* New York: Holt.

Kaiser, D. H., Sherburne, L. M., Steirn, J., & Zentall, T. R. (1997). Perceptual learning in pigeons: Decreased ability to discriminate samples mapped onto the same comparison in many-to-one matching. *Psychonomic Bulletin & Review, 4,* 378–381.

Kurtz, K. H. (1955). Discrimination of complex stimuli: The relationship of training and test stimuli in transfer of discrimination. *Journal of Experimental Psychology, 50,* 283–292.

Lawrence, D. H. (1949). Acquired distinctiveness of cues: I. Transfer between discriminations on the basis of familiarity with the stimulus. *Journal of Experimental Psychology, 39,* 770–784.

Miller, N. E., & Dollard, J. (1941). *Social learning and imitation.* New Haven, CT: Yale University Press.

Miller, R. R., Hallam, S. C., & Grahame, N. J. (1990). Inflation of comparator stimuli following CS training. *Animal Learning & Behavior, 18,* 434–443.

Miller, R. R., & Matute, H. (1996). Biological significance in forward and backward blocking: Resolution of a discrepancy between animal conditioning and human causal judgment. *Journal of Experimental Psychology: General, 125,* 370–386.

Norcross, K. J. (1958). Effects on discrimination performance of similarity of previously acquired stimulus names. *Journal of Experimental Psychology, 56,* 305–309.

Norcross, K. J., & Spiker, C. C. (1957). The effects of type of stimulus pretraining on discrimination performance in preschool children. *Child Development, 28,* 79–84.

Reese, H. W. (1972). Acquired equivalence and distinctiveness of cues in young children. *Journal of Experimental Child Psychology, 13,* 171–182.

Sutherland, N. S., & Mackintosh, N. J. (1971). *Mechanisms of animal discrimination learning.* New York: Academic Press.

Ward-Robinson, J., & Hall, G. (1996). Backward sensory preconditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 22,* 395–404.

Ward-Robinson, J., & Hall, G. (1999). The role of mediated conditioning in acquired equivalence. *The Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 51B,* 335–350.

Wasserman, E. A., & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgement: The role of within-compound associations. *The Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 51B,* 121–138.

Wilson, T. D., Lindsey, S., & Schooler, T. Y. (2000). A model of dual attitudes. *Psychological Review, 107,* 101–126.

Received January 22, 2002
Revision received December 12, 2002
Accepted December 16, 2002 ■