Published in *Journal of Experimental Psychology: Animal Behavior Processes* *2011, vol 37 (4), pp 446-465.*
http://www.apa.org/pubs/journals/xan/index.aspx

Comparing learned predictiveness effects within and across compound discriminations

Evan J. Livesey, Anna Thorwart, Nicole L. De Fina & Justin A. Harris

University of Sydney

Abstract

In four human learning experiments, we examined the extent to which learned predictiveness depends upon direct comparison between relatively good and poor predictors. Participants initially solved (1) linear *compound* discriminations in which one or both of the stimuli in each compound were predictive of the correct outcome, (2) biconditional discriminations where only the configurations of the stimuli were predictive of the correct outcome, or (3) pseudo-discriminations in which no stimulus features were predictive. In each experiment, subsequent learning and test stages were used to assay changes in the associability of each stimulus brought about by its role in the initial discriminations. Although learned predictiveness effects were observed in all experiments (i.e. previously predictive cues were more readily associated with a new outcome than previously non-predictive cues), the same changes in associability were observed regardless of whether the stimulus was initially learned about in the presence of an equally predictive, more predictive, or less predictive stimulus. The results suggest that learned associability is not controlled by competitive allocation of attention, but rather by the absolute predictiveness of each individual cue.

Address correspondence to:

Evan Livesey
School of Psychology
University of Sydney
Sydney 2006
Australia
Email: evan.livesey@sydney.edu.au

Running head: Learned predictiveness in compound discriminations

Keywords: Learned predictiveness, stimulus associability, biconditional discrimination.

This article may not exactly replicate the final version published in the APA journal. It is not the copy of record.

Many theories of associative learning rely upon some form of selective attention to explain evidence that simultaneously presented stimuli compete for learning. Attention-based models of conditioning (e.g. Mackintosh, 1975; Pearce & Hall, 1980) assume that selective attention to a conditioned stimulus (CS) is labile and governed by both the associative history of the CS – its past involvement in signalling a salient event – and also the associative histories of other stimuli with which it is presented. For instance, in Mackintosh's model, which is critical to the current study, attention to a stimulus increases if it predicts an observed outcome more strongly than others present at the same time but attention decreases if the stimulus is a poorer predictor of the outcome than others with which it is presented. Formally, the strength of the association (V) between a stimulus A and an outcome is modified according to the following equation:

Equation 1

$$\Delta V_A = S\,\alpha_A\,(\lambda - V_A)$$

In this equation, $(\lambda\text{-}V_A)$ represents the mismatch between the actual outcome on that trial (λ) and what was anticipated as a consequence of A. If $(\lambda\text{-}V_A)$ is low, it suggests that A is a good predictor of the outcome. S and α_A are parameters that govern the rate of learning. Most importantly α_A, which is unique to stimulus A and governs its associability (the rate at which associations with the stimulus are modified), is adjusted on each learning trial according to a comparison with other stimuli:

Equation 2

$$\Delta\alpha_A > 0 \qquad \text{if } |\lambda - V_A| < |\lambda - V_X|$$
$$\Delta\alpha_A < 0 \qquad \text{if } |\lambda - V_A| \geq |\lambda - V_X|$$

Here, α_A increases if A is a better predictor than other stimuli present at the same time (represented as V_X) but decreases if it is a relatively poor predictor. Thus stimulus associability changes according to a competitive stimulus comparison process. The Mackintosh model uses this comparison of the predictiveness of simultaneously presented cues, which we will refer to as relative predictiveness, to explain a range of cue competition effects. For instance, Mackintosh (1975) appealed to relative predictiveness to explain Kamin's (1968) well-known blocking effect. When a pretrained and novel cue are conditioned in compound, the pretrained cue will block learning about the novel cue because it has already developed a strong association with the unconditioned stimulus (US), rendering the novel cue a weak predictor by comparison. As a consequence of this comparison, attention to the novel cue diminishes, limiting further learning.

Several recent human learning studies confirm the influence of associative history over the rate of *de novo* learning by investigating how quickly an association is learned between the CS and a new outcome. These studies generally indicate that stimulus associability increases for strong predictors relative to weaker ones, in line with the general predictions made by the Mackintosh (1975) model and several others (e.g. Kruschke, 2001; Le Pelley, 2004; Pearce & Mackintosh, 2010). One of the clearest and most relevant examples is the learned predictiveness effect reported by Le Pelley & McLaren (2003; see also Lochman & Wills, 2003). Participants in their study were given a task in which they assumed the role of an allergist trying to discover the causes of a patient's allergic reactions. In the first stage of the experiment, participants were given compounds of two foods eaten by the patient (e.g. ham and tomato) and predicted which of two reactions (e.g. sweating or nausea) would occur as a result. Participants were given repeated trials of several combinations of foods where, in each compound, one food perfectly predicted the correct outcome and the other was irrelevant. A simplified version of this design is shown in Table 1. In Stage 1, some stimuli (A-D) are consistently reinforced, followed each time by the same outcome. These stimuli can therefore be considered strongly predictive of the outcome, whereas others (W-Z) are inconsistently reinforced and can be considered non-predictive. In Stage 2, participants were instructed to perform the same task for a new patient. This time new compounds were presented, each comprising one previously predictive and one previously non-predictive stimulus, with each compound predicting one of two new allergic reactions (e.g. itch or dizziness). Note that in Stage 2, the predictive and non-predictive stimuli are equally diagnostic of the outcome (the terms predictive and non-predictive indicate their role in Stage 1 only). The test stage measured learning of these new allergy outcomes by recombining the two predictive stimuli that were paired with the same Stage 2 outcome (AD, BC) and recombining the two non-predictive stimuli that were paired with the same Stage 2 outcome (XY, WZ). Test ratings for the predictive compounds favoured the correct outcome more strongly than the non-predictive compounds, indicating better learning.

This result demonstrates that the associability of the predictive stimuli increased relative to the non-predictive stimuli. Further studies have demonstrated that these learned predictiveness effects are dependent on the trial-and-error nature of the task. For instance, Le Pelley et al. (2010a) failed to find the same result when they simply displayed the relevant information of the two training stages in text form. This suggests that test rating biases reflect more than just inferences made at the time of testing, but rather involve differential selective attention devoted to the putative causes during learning in Stage 2.

Typically, the learned predictiveness effect has been interpreted using Mackintosh’s (1975) conception of selective associability (Le Pelley & McLaren, 2003; Le Pelley, 2004). That is, the learning bias towards previously predictive cues is attributed to a competitive attention process in which the individual compares the relative utility of all stimuli present on a given trial for predicting the outcome that actually occurred and shifts

Table 1. Learned predictiveness design, adapted from Le Pelley & McLaren (2003).

Stage 1	Stage 2	Test	Prediction
AW – O1	AY – O3	AD?	*(O3 – O4)*
AX – O1	BZ – O4	XY?	AD > XY
BW – O2	CW – O4	BC?	
BX – O2	DX – O3	WZ?	*(O4 – O3)*
CY – O1			BC > WZ
CZ – O1		*(rate for O3 & O4)*	
DY – O2			
DZ – O2			

Note: Letters refer to individual stimuli, A-D: predictive components, W-Z: non-predictive components. O1-O4 refer to four outcomes.

attention towards relatively good predictors and away from relatively poor predictors. The Mackintosh (1975) model elegantly describes this process, which has also been incorporated in more recent models of associative learning (Le Pelley, 2004, Pearce & Mackintosh, 2010). Critically, it is the *relative* predictiveness of the items occurring at a specific time that dictate the changes in associability. The associability of good predictors is assumed to increase through comparison with poorer predictors occurring on the same trial, while the associability of the poorer predictors diminishes as a consequence of the same comparison with better signals of the outcome. In this sense, the absolute predictiveness of stimulus A (e.g. gauged by $|\lambda\text{-}V_A|$ on each trial) has no direct influence on α_A, which only changes through comparison with other stimuli. In contrast, other explanations for the learned predictiveness effect do not place such an emphasis on relative predictiveness. Alternative accounts, based on either the general relevance of the stimulus to solving a discrimination or the individual predictiveness of the stimulus measured in an absolute fashion, do not require comparison with other accompanying stimuli.

The current experiments tested the relative predictiveness hypothesis by comparing learned predictiveness effects generated by different compound discriminations, some of which contain useful within-compound differences in predictive validity and some of which do not. Thus the aim was to test whether learned associability is enhanced (or in fact completely driven) by direct stimulus comparison, as predicted by the Mackintosh model. In the current study, we compared the contingencies used in a standard learned predictiveness design (Le Pelley & McLaren, 2003) with several alternative discriminations in which the various pairings of cues in compound remained the same but the cues in each compound were always equally valid or invalid predictors of the correct outcome. Any observed difference between the associabilities of the stimuli comprising each discrimination must be attributed to their differential associative histories because all factors to do with stimulus presentation, including frequency of occurrence and within-compound associations were identical across the discriminations.

Across different experiments, we compared the standard “component” discrimination that has typically been used to demonstrate the learned predictiveness effect with three alternative discriminations. All four discriminations are shown in Table 2 for ease of comparison. The first was the biconditional discrimination (Saavedra, 1975) in which four stimuli are paired in four different compounds, two pairs predicting one outcome and two predicting another (or no) outcome, in such a fashion that no single stimulus is predictive in and of itself (e.g. JN – O1, JO – O2, KN – O2, KO – O1). Humans and other animals are capable of solving biconditional discriminations. Indeed, Saavedra (1975) compared acquisition of a biconditional discrimination in rabbits to a component discrimination that is formally very similar to Le Pelley and McLaren’s (2003). She found that acquisition of the biconditional discrimination was slow compared to the component discrimination but was nevertheless eventually achieved. Biconditional discriminations are also relatively difficult compared to other similar nonlinear discriminations such as positive and negative patterning (e.g. Harris & Livesey, 2008; Harris, Livesey, Ghareai, & Westbrook, 2008) because, although attention to all stimuli is critical for solving the discrimination, the individual stimuli are all (equally) uncorrelated with the possible outcomes. On any given compound trial, neither stimulus is predictive in its own right, but the answer is nevertheless completely solvable using the configuration of the two cues. Thus, although both cues are technically non-predictive, they are both highly relevant for solving the discrimination because they are completely predictive as part of a compound.

Another discrimination used in the present experiments was much easier than the biconditional. Again four stimuli were presented in four different pairs (JN, JO, KN, KO) but all four predicted the same outcome. As shown in Table 2, a second set of four stimuli were presented in the same fashion and all predicted the alternative outcome. Although these contingencies were very easy to learn – easier still than the component discrimination with which learned predictiveness effects are usually induced – the individual stimuli are all equally predictive. Consequently, each component of these ‘simple’ discriminations was never accompanied by a more predictive or less predictive stimulus.

Finally, we also used an unsolvable pseudo-discrimination in which no cue or combination of cues was predictive. The same stimuli were presented in the same pairings, but each pair was equally often followed by O1 and O2. Accuracy across training for this impossible pseudo-discrimination should remain at chance. In this case, all cues are equally non-predictive and none are ever presented simultaneously with a stimulus that is predictive of the correct outcome.

All four experiments used a human predictive learning task similar to those employed in previous studies. Participants took part in a hypothetical scenario in which they have been given a magical deck of cards that controlled the weather. Using this scenario, the cues presented on each trial were different picture cards comprising simple line drawings of familiar objects and the outcomes were four different weather events, two presented as potential options in Stage 1 and two presented in Stage 2. Participants were required to predict which weather outcome would occur on each trial given the pair of

Table 2. Stage 1 training contingencies used in all experiments.

Component	Biconditional	Simple		Impossible
Exp. 1,2,& 4	*Exp. 1&2*	*Exp. 3*	*Exp. 4 grpS*	*Exp. 3 & 4 grpI*
AW – O1	JN – O1	AW – O1	JN – O1	JN – O1/O2
AX – O1	JO – O2	AX – O1	JO – O1	JO – O1/O2
BW – O2	KN – O2	BW – O1	KN – O1	KN – O1/O2
BX – O2	KO – O1	BX – O1	KO – O1	KO – O1/O2
CY – O1	LP – O1	CY – O2	LP – O2	LP – O1/O2
CZ – O1	LQ – O2	CZ – O2	LQ – O2	LQ – O1/O2
DY – O2	MP – O2	DY – O2	MP – O2	MP – O1/O2
DZ – O2	MQ – O1	DZ – O2	MQ – O2	MQ – O1/O2

Note: Letters refer to individual cues, A-D: predictive components, W-Z: non-predictive components (excepting Experiment 3, where A-D & W-Z were equally predictive), J-Q: equally predictive (or non-predictive) stimuli. O1 and O2 refer to two neutral outcomes.

cards that had been dealt, thereby learning which card or combination of cards led to which outcome. The final stage then tested the associations learned in Stage 2. Participants rated the likelihood of each of the Stage 2 outcomes occurring for a range of different pairs of cards.

The first two experiments used component and biconditional discriminations in order to replicate the standard learned predictiveness effect in this more complex design (Experiment 1) and to directly compare the associability of stimuli involved in these two discriminations (Experiment 2). Experiment 3 used a simple discrimination and impossible pseudo-discrimination to compare associability changes brought about as a result of absolute differences in the predictiveness of compounds comprising two equally valid (or invalid) cues. Finally, Experiment 4 directly pitted both of these discriminations against the component discrimination used in conventional learned predictiveness designs to test whether direct comparison with a more predictive or less predictive stimulus facilitates associability change. To preview the results, clear evidence of learned associability effects were observed in all four experiments. However, changes in associability for the predictive and non-predictive stimuli of the component discrimination were equivalent to changes in associability in the other discriminations where the two stimuli in each compound were equally valid or equally invalid predictors of the outcome.

Experiment 1

Experiment 1 aimed to replicate the standard learned predictiveness effect induced by learning a component discrimination using a design in which two biconditional discriminations were also learned concurrently. After Stage 1 training with two component and two biconditional discriminations (see Table 2), participants were given Stage 2 training with the contingencies shown in Table 3. In this stage, four compounds were made by recombining the predictive and non-predictive components from Stage 1, two predicting Outcome 3 and two predicting Outcome 4. Four additional compounds were made by recombining components of the biconditional discriminations. As in previous learned predictiveness experiments (Le Pelley & McLaren, 2003), Stage 2 learning was subsequently tested using new 'summation' compounds comprising two components that were predictive of the same outcome (e.g. AY and DX both predicted O3 in Stage 2, and in test AD and XY were presented). We also tested all of the compounds actually trained in Stage 2 as well as a series of 'negation' compounds comprising two stimuli predictive of different outcomes. The purpose of these negation tests was twofold. First, they provided an additional measure of learned associability effects. For instance, since AY led to O3 and BZ led to O4 in Stage 2, then biases towards learning about the predictive components should result in AZ being more strongly associated with O3 and BY more strongly associated with O4. Second, their inclusion meant that both components of a given test trial did not always predict the same outcome, forcing participants to consider both components while making their ratings during test.

If learned associability changes are not adversely affected by the presence of the biconditional discriminations in Stage 1 then performance on the predictive component summation tests (AD and BC) should be substantially better than performance on the non-predictive summation tests (XY and WZ).

Table 3. Design of Stage 2 training and Test in Experiment 1 (see Table 2 for Stage 1).

	Training	Test (all trials intermixed)		
	Stage 2	*Trained*	*Summation*	*Negation*
(component stimuli)	AY – O3	AY	AD	AZ
	BZ – O4	BZ	BC	BY
	CW – O4	CW	XY	CX
	DX – O3	DX	WZ	DW
(biconditional) stimuli)	JP – O3	JP	JM	JQ
	KQ – O4	KQ	KL	KP
	LN – O4	LN	OP	LO
	MO – O3	MO	NQ	MN

Note: Letters refer to stimuli, all trained previously in Stage 1, A-D: predictive components, W-Z: non-predictive components, J-Q: biconditional stimuli. O3 and O4 refer to two neutral outcomes. All test compounds were rated for the likelihood of being followed by O3 and by O4.

Method

Participants and Apparatus. Twenty-nine undergraduate psychology students (23 female, mean age = 21 years) participated in Experiment 1. Participants were group tested in a large classroom. The experiment was run on Apple iMac computers with 17 inch LCD screens, running software developed in Visual Basic. Computers were spaced approximately 1m apart.

Stimuli. Sixteen line drawings of familiar objects (Snodgrass & Vanderwart, 1980) were chosen to serve as the predictive cues in the experiment and were displayed as images appearing on a set of cards. The cards, measuring 8.5° x 12.5° of visual angle (viewing distance of approximately 57cm) were always presented in pairs appearing at the top of the screen, spaced 9.5° apart. On each trial the cards initially appeared as if they were face-down with the generic image of a card back, and were then ‘flipped over’ in a simple animation, revealing the line drawings. The outcomes used throughout the experiment were four weather phenomena, ‘rain’, ‘snow’, ‘hail’ and ‘fog’, each presented as a word accompanied by a schematic image of the weather phenomenon. For each participant, the 16 images were randomly allocated to serve as the predictive, non-predictive, and biconditional cues (A-D, W-Z, J-Q, respectively), as were the four weather phenomena to each of the outcomes (O1-O4).

Procedure. Participants were instructed to take part in a fictitious scenario in which they have been given a magical deck of cards that can control the weather. They were told that every morning they deal two cards and then observe what the weather is like on that day. They were also told that they would be required to predict which weather event would occur before being shown the outcome for that day.

Stage 1 consisted of 16 blocks of trials, with each block containing one of each of the 16 trial types shown in the “Component” and “Biconditional” columns of Table 2. On each trial, participants viewed the two card cues, and then used the mouse to click on one of two possible weather outcomes. Choices were self-paced and were met with immediate feedback indicating whether they were correct or incorrect, accompanied after 1 second by the actual weather outcome for that day. The order of trials within each block was randomised and the spatial layout of cards within each pair was counterbalanced across blocks.

At the beginning of Stage 2, participants were informed that now they had been given a new deck of cards, that this deck would also control the weather but not in the same way as the first deck. Stage 2 ran in a very similar way to Stage 1, except that the cards were shown in novel compounds, each predicting one of two new outcomes O3 and O4 (as shown in Table 3). Eight blocks of trials were presented, with one presentation of each of the eight compounds per block. Again, order was randomised within blocks.

At the beginning of the test stage participants were instructed that they would be shown pairs of cards from the second deck and that they would have to rate the likelihood of each of the two weather outcomes (O3 and O4) for each pair. For each test trial, two linear analogue scales appeared, one labelled for O3 and one for O4. Both scales were labelled “Very unlikely to occur” at the left extreme and “Very likely to occur” at the right extreme. Participants completed each trial by making a rating on both scales and then pressing the space bar. Twenty-four test trials were presented in the test stage, as shown in Table 3. These included the eight compounds actually trained in Stage 2, as well as summation compounds comprising two stimuli that predicted the same outcome in Stage 2 and negation compounds of two stimuli that predicted opposite outcomes in Stage 2. These were presented in a randomised order.

Results

Statistical tests reported for all four experiments were considered significant at *p*=.05. Greenhouse-Geisser corrections to *p* values were used where necessary for all repeated measures analyses with more than two levels of the same factor (uncorrected degrees of freedom are reported). Associability changes are thought to occur as a consequence of learning the component discrimination and are only expected if the discrimination is mastered to some extent. To ensure that the component discrimination was actually learned we excluded participants from further analysis if they failed to make correct predictions in at least 60% of the trials of the component discriminations in the final quarter of Stage 1. This is in line with past literature on learned predictiveness (Le Pelley & McLaren, 2003) and the same criterion was used across experiments. In Experiment 1, six participants failed to reach this criterion. Further analyses were conducted with the remaining 23 (16 female, mean age = 21.3 years).

Stage 1. Acquisition was initially gauged using the predictions made on each trial during the training stages. Mean accuracy in each block was averaged across the two component discriminations and across the two biconditional discriminations (8 trials per block for each) as shown in Figure 1. Consistent with previous results, accuracy was generally higher for the component discriminations than the biconditional. A repeated measures ANOVA with training block (1-16) and discrimination (component vs biconditional) as factors revealed significant main effects of both block ($F(15,330)=22.83$, $p<.001$, $\eta_p^2=.509$) and discrimination ($F(1,22)=42.22$, $p<.001$, $\eta_p^2=.657$). Although, the interaction between discrimination and block did not reach significance (F(15,330)=1.33, p=.237, η_p^2=.057), a significant interaction between discrimination and the quadratic trend over block (F(1,22)=7.82, p=.011, η_p^2=.262) suggests that the component discrimination was learned faster than the biconditional.

Stage 2. Figure 2 shows accuracy for the recombined compounds used in Stage 2, again averaged according to their roles in Stage 1 as stimuli in the component or biconditional discriminations. A repeated measures ANOVA with training block (1-8) and Stage 1 discrimination (component vs biconditional) as factors revealed a significant main effect of block (F(7,154)=9.01, p<.001, η_p^2=.290) but no significant effect of Stage 1 discrimination (F(1,22)=1.99, p=.172, η_p^2=.083). There was also no significant interaction between Stage 1 discrimination and block, nor interactions with linear and quadratic trends across blocks (Fs < 1). Although accuracy clearly improved over the course of Stage 2, there was no observable difference between the compounds composed of stimuli used in the component discrimination in Stage 1 and compounds composed of stimuli used in the biconditional discrimination in Stage 1.

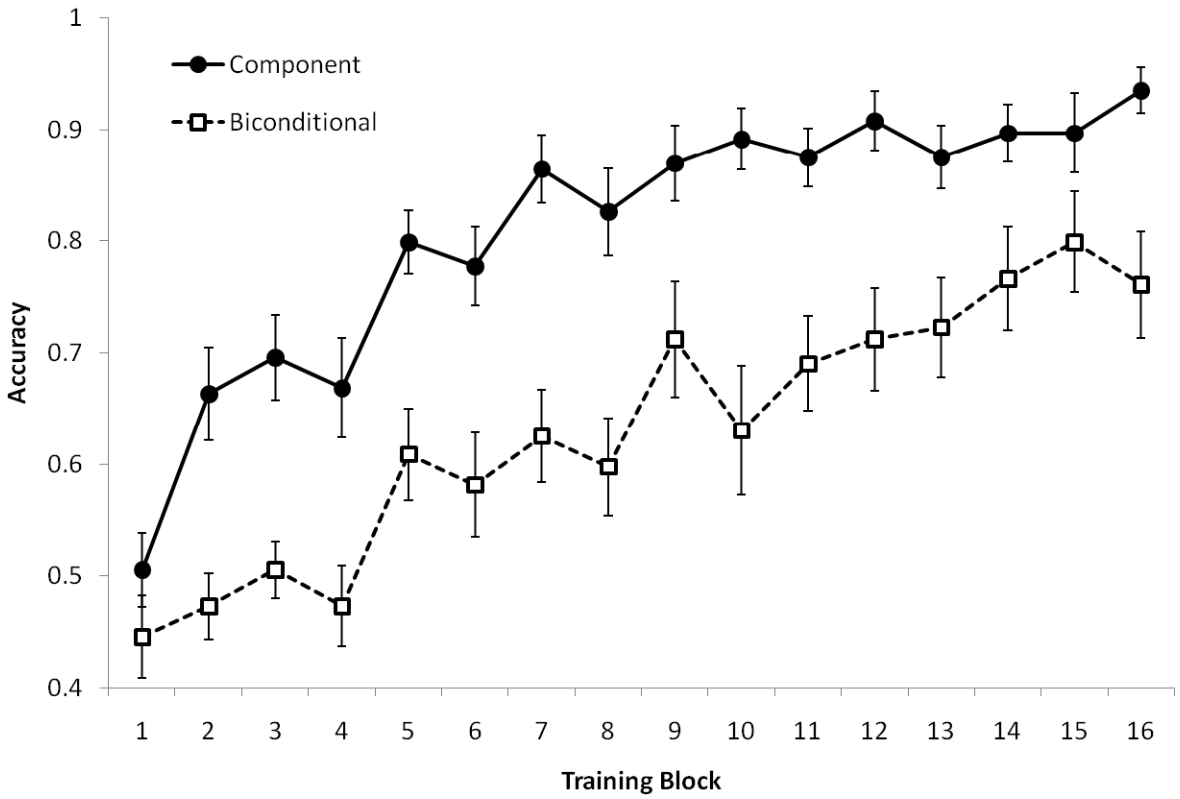

Figure 1. Experiment 1, mean accuracy for the component and biconditional discriminations across the 16 blocks of Stage 1 training. Error bars indicate standard error of the mean (SEM).

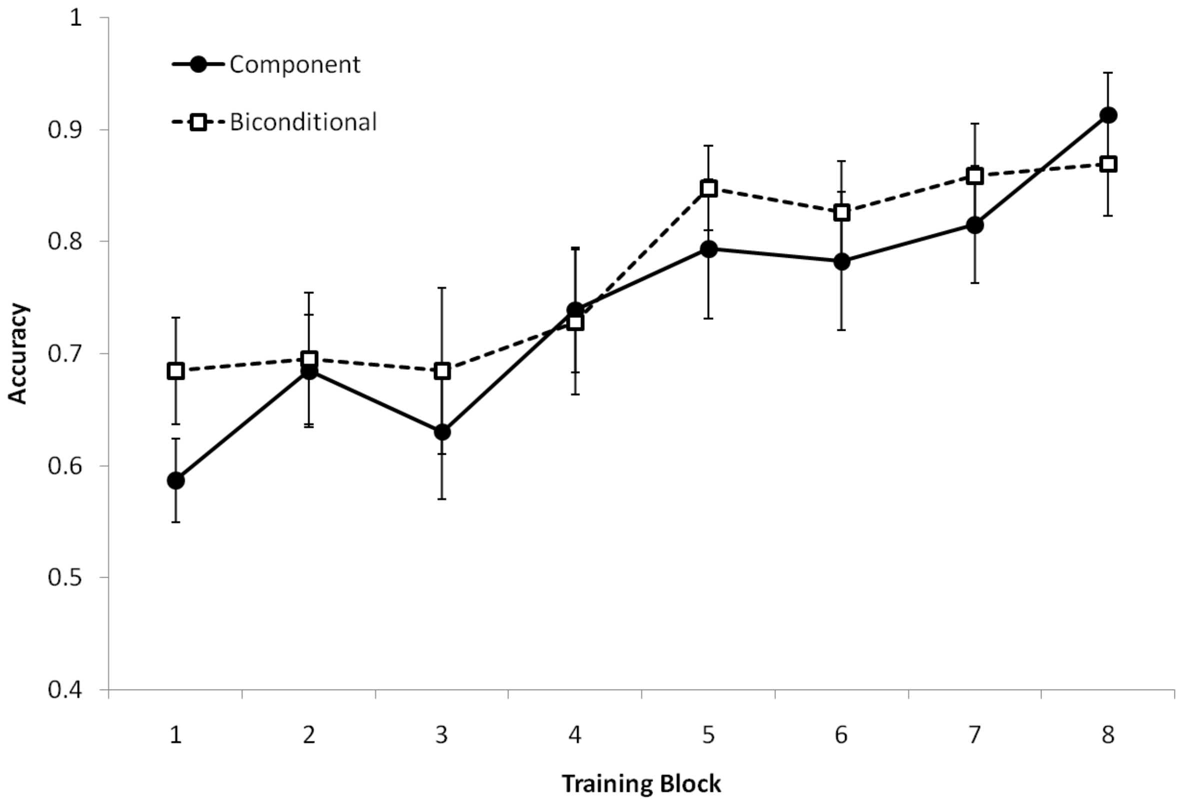

Figure 2. Experiment 1, mean accuracy across the eight blocks of Training Stage 2. Data are averaged across compounds of stimuli previously involved in Component and Biconditional discriminations. Error bars indicate SEM.

Test Stage. Mean ratings of the likelihood of O3 and O4 occurring for each of the 24 test compounds are shown in Figure 3A. Consistent with previous learned predictiveness research, we take the best indication of Stage 2 learning to be the difference between the ratings for the two outcomes. As shown in Table 3, the test compounds can be categorized as trained, summation, and negation compounds. Within each of these categories, the ratings were combined and collapsed into test scores reflecting the role of the stimuli in Stage 1. Thus for trained stimuli, a mean component score was obtained by taking the difference in rating for the predicted outcome and the alternate outcome for each of the compounds previously involved in the component discrimination (O3 – O4 for AY & DX, O4 – O3 for BZ & CW) and a mean biconditional score was obtained in the same way (O3 – O4 for JP & MO, O4 – O3 for KQ & LN). Likewise, summation test trials were used to calculate mean summation scores for 'predictive' trials (O3 – O4 for AD, O4 – O3 for BC), 'non-predictive' trials (O3 – O4 for XY, O4 – O3 for WZ), and 'biconditional' trials (O3 – O4 for JM & OP, O4 – O3 for KL & NQ). For each of these trained and summation scores, higher values indicate more accurate performance. On negation trials, the two components of each compound predict opposite outcomes. A 'component' negation score was calculated so that a positive score indicates a bias towards the previously predictive cues and a negative score indicates a bias towards the previously non-predictive cues (O3 – O4 for AZ & DW, O4 – O3 for BY & CX). A 'biconditional' negation score was calculated, matched to the component score (O3 – O4 for JQ & MN, O4 – O3 for KP & LO), although here there is no natural division between predictive and non-predictive components and consequently the expected mean score is zero. Figure 3B shows each of these mean scores.

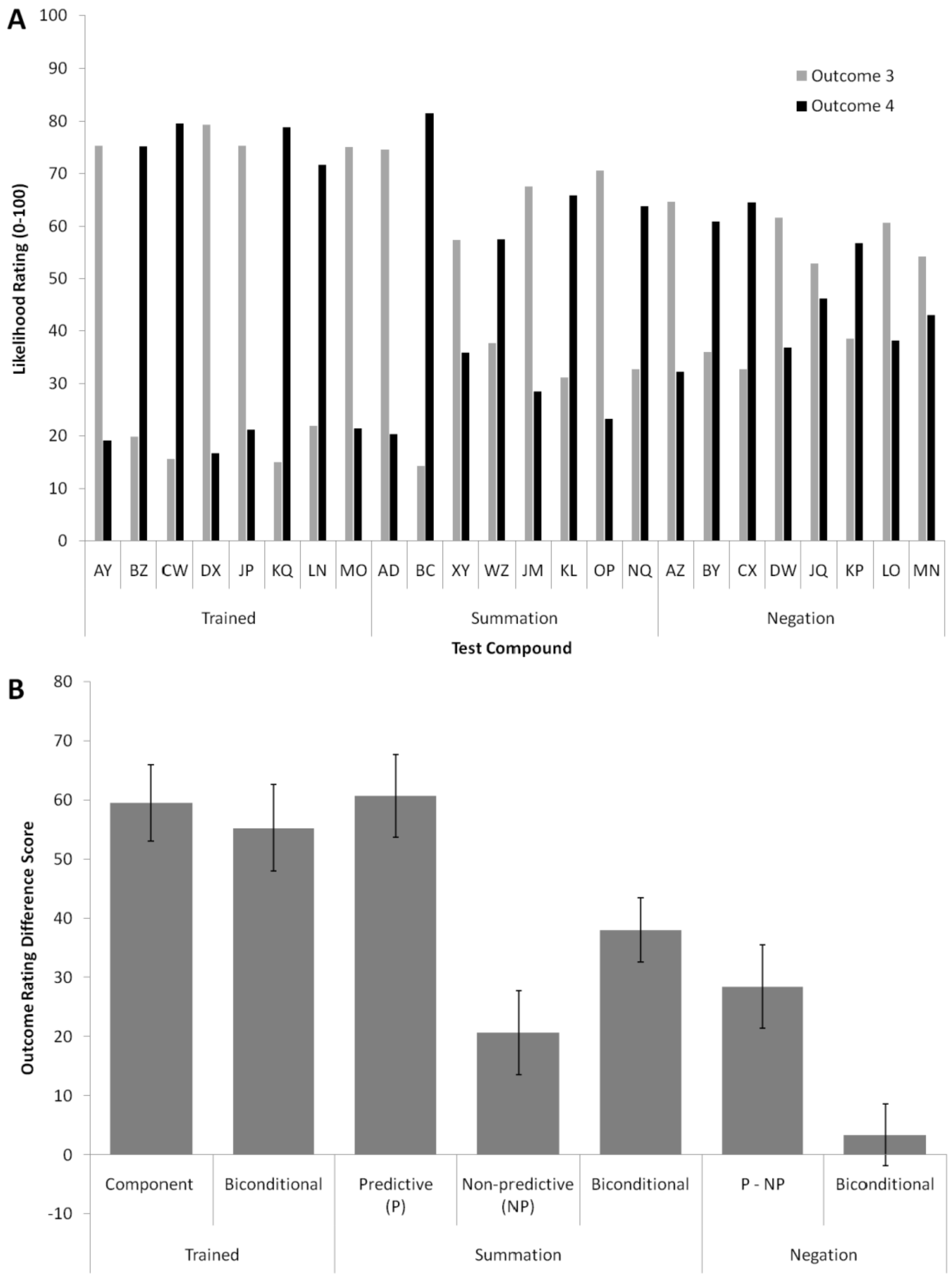

Figure 3. Experiment 1, ratings test data. A) Mean ratings of the likelihood of the Stage 2 outcomes (O3 & O4) for each of the 24 test trials. B) Test scores calculated by averaging the difference between the rating for the predicted outcome and the opposite outcome. Error bars indicate SEM.

For simplicity, all statistical analyses of the test stage results were run on the outcome rating difference scores as shown in Figure 3B. For the trained compounds, there was no significant difference between the trained component compounds and trained biconditional compounds (F<1). For the summation compounds, scores for the predictive compounds were significantly higher than scores for the non-predictive compounds ($F(1,22)=17.06$, $p<.001$, $\eta_p^2=.437$), replicating the standard learned predictiveness effect. Moreover, scores for the predictive compounds were significantly higher than scores for the biconditional compounds ($F(1,22)=10.25$, $p=.004$, $\eta_p^2=.318$) and scores for the non-predictive compounds were significantly lower than scores for the biconditional compounds ($F(1,22)=5.43$, $p=.029$, $\eta_p^2=.198$). For the negation test compounds, the component score (P-NP) was significantly higher than the biconditional score ($F(1,22)=17.81$, $p<.001$, $\eta_p^2=.447$). Perhaps more importantly, the component negation score differed significantly from zero ($t(22)=4.03$, $p=.001$), whereas the biconditional did not ($t<1$).

Discussion

Experiment 1 demonstrates that stimulus associability changes in the expected manner, favouring previously predictive stimuli, even when participants are also engaged in other discriminations in which the solution cannot be derived from attending more to one particular stimulus over another. Performance on the biconditional discrimination was worse than the component discrimination as expected, but learning the biconditional discrimination did not prevent associability change manifesting in the later stages of the experiment. This establishes a useful paradigm for the comparisons made in Experiments 2-4.

The conventional interpretation of the learned predictiveness effect is an elemental one, where associability is assumed to increase and decrease amongst the elements that make up the representation of the stimuli (Le Pelley & McLaren, 2003; Livesey & McLaren, 2007; Suret & McLaren, 2005). Several researchers have recently considered the possibility that experience with nonlinear problems like the biconditional discrimination diminishes elemental encoding of stimuli (e.g. Melchers, Shanks & Lachnit, 2008, Livesey & Boakes, 2004; Thorwart & Lachnit, 2010; Urcelay & Miller, 2009; Williams & Braker, 1999). Although the strength of the evidence in favour of this view is disputed (e.g. Liljeholm & Balleine, 2008; Livesey & Harris, 2008), the idea itself suggests that a concurrent biconditional discrimination may cause a general shift away from elemental encoding of the stimuli, which in turn might reduce or abolish learned associability effects if they rely upon elemental encoding. Experiment 1 served to verify that learned predictiveness effects still operate in the presence of nonlinear discriminations.

On the basis of the summation test results, it is tempting to conclude that the biconditional stimuli maintained a higher associability than non-predictive stimuli. However, this cannot be ascertained with the current design. Even if associability only increased for the predictive components and did not change for either the biconditional or non-predictive stimuli, one would predict greater overshadowing of the non-predictive components than the biconditional stimuli. In order to gauge the stimulus associability of one stimulus relative to that of another stimulus, both stimuli must be trained in compound in Stage 2. Experiment 2 was designed to do just this for stimuli of the biconditional and component discriminations.

Experiment 2

In Experiment 2, Stage 2 learning directly pitted the stimuli from the component discriminations against stimuli from the biconditional discriminations. This is shown in Table 4, which describes the compound training and test design used in all of the remaining experiments. In Stage 2, stimuli are presented in novel compounds comprising one cue from each of the two types of Stage 1 discrimination. As with the original learned predictiveness design, the new outcomes predicted by each of the stimuli are balanced to avoid an overall congruency between the outcomes predicted in Stage 1 and the outcomes predicted in Stage 2. For instance, in Stage 1 cues A and C predict O1, B and D predict O2 but in Stage 2 cues A and D predict O3, B and C predict O4.

The purpose of this design was to test the learned associability of stimuli involved in the biconditional discrimination relative to predictive and non-predictive stimuli in the component discrimination. Each biconditional stimulus has the unique status of being entirely necessary for solving the discrimination but also non-predictive of any particular outcome in its own right. In terms of their absolute predictiveness, no single stimulus within the biconditional discrimination is predictive of one outcome over another, and thus any effect of associability based solely on the associative strength of an individual cue should be low. In terms of relative predictiveness, each biconditional stimulus is equivalent in its validity to the other cue with which it is paired. However, the individual stimuli in the biconditional discrimination are less predictive than the configuration as a whole.

.

Table 4. Stage 2 training and Test design of Experiments 2-4 (see Table 2 for Stage 1).

	Training	Test Stage (all intermixed)		
(role in Exp. 2&4)	Stage 2	*Trained*	*Summation*	*Negation*
(predictive	AJ – O3	AJ	AD	AL
components	BK – O4	BK	BC	BM
vs other)	CL – O4	CL	JM	CJ
	DM – O3	DM	KL	DK
(non-predictive	WN – O3	WN	WZ	WP
components	XO – O4	XO	XY	XQ
vs other)	YP – O4	YP	OP	YN
	ZQ – O3	ZQ	NQ	ZO

Note: Letters refer to stimuli, all trained previously in Stage 1. In Experiment 3, A-D & W-Z: simple discrimination, J-Q: impossible pseudo-discrimination. In Experiments 2 and 4, A-D: predictive components, W-Z: non-predictive components, J-Q: stimuli from the “other” discrimination; biconditional (Exp. 2), simple (Exp. 4 Group S), or impossible (Exp. 4 Group I). O3 and O4 refer to two neutral outcomes. All test compounds were rated for the likelihood of being followed by O3 and by O4.

To the extent that the configuration of two cues can be considered a functional stimulus in and of itself (e.g. Whitlow & Wagner, 1972; Spence, 1952) the individual cues are poor predictors relative to another physically intangible stimulus that is present at the same time. Consequently, their associability may decrease, equivalent to the non-predictive stimuli in the component discrimination. However, if learned associability effects are more strongly driven by comparison with a physically isolable stimulus with higher or lower validity (as is possible for the component discrimination) then the biconditional stimuli should display a greater associability than the non-predictive stimuli because only the non-predictive stimuli have been compared to a more relevant physical stimulus.

The above points notwithstanding, the cues in the biconditional discrimination have high stimulus relevance. Participants must attend to both cues in each compound in order to solve the discrimination. If learned predictiveness is driven by attention to cues that were previously relevant to discrimination then associability for the biconditional stimuli should be high, equivalent to the associability for the predictive cues. There is some evidence for this hypothesis from animal learning. George and Pearce (1999) trained pigeons on a biconditional discrimination (AP+ AQ– BP– BQ+; + refers to reinforcement, – to nonreinforcement) where every trial included a third but completely irrelevant cue (W or X) which had no impact on the schedule of reinforcement. In a second phase, the pigeons received a new biconditional discrimination comprising stimuli from the same stimulus dimensions as those used in the first phase. Pigeons showed impaired acquisition of the second discrimination if the previously irrelevant stimuli formed an integral part of the biconditional contingencies (e.g. CW+ CX- DW- DX+; P & Q also present but irrelevant) relative to a condition where the relevance of the stimuli remained the same (e.g. CP+ CQ- DP- DQ+; W & X also present but irrelevant). This demonstrated that pigeons learn to attend to the stimuli that are necessary for solving the discrimination even though no single stimulus was uniquely predictive of reinforcement or nonreinforcement. George and Pearce’s (1999) design was very different to the human learning paradigm used here. The stimuli were actually variations along three dimensional properties (location, color and orientation) of a single integral stimulus. Even under an elemental theoretical framework, it is debatable whether one should expect the representations of each stimulus value to operate as discrete and independent components (e.g. see Harris & Livesey, 2010). Nevertheless, it is clear that learning a biconditional discrimination could potentially lead to enhanced attention for the biconditional components because of their utility in solving the discrimination. This outcome would suggest that associability changes (including the conventional learned predictiveness effect) are the consequence of the general relevance of a stimulus to solving a discrimination and not stimulus predictiveness, relative or otherwise.

Method

Participants and Apparatus. Thirty-one undergraduate psychology students (19 female, mean age = 22 years) participated in Experiment 2. Participants were group tested using the same apparatus as in Experiment 1.

Stimuli and Procedure. All other aspects of the experimental method were identical to Experiment 1 except for the contingencies shown in Stage 2 (which paired the predictive and non-predictive components with stimuli from the biconditional discriminations, rather than with each other) and the test stimuli, as shown in Table 4.

Results

As in Experiment 1, participants were excluded from further analysis if they failed to achieve greater than 60% accuracy across the last 4 blocks of training. Using this criterion, three participants were excluded and their data discarded. All analyses were performed on the remaining 28 participants.

Stage 1. As with Experiment 1, mean accuracy in each block was averaged across the two component discriminations and across the two biconditional discriminations (shown in Figure 4). Again, accuracy was higher for the component discriminations than the biconditional. A repeated measures ANOVA with training block (1-16) and discrimination (component vs biconditional) as factors revealed significant main effects of both block ($F(15,405)=35.64$, $p<.001$, $\eta_p^2=.569$) and discrimination ($F(1,27)=34.69$, $p<.001$, $\eta_p^2=.562$). The interaction between discrimination and block was not significant ($F(1,27)=1.20$, $p=.300$, $\eta_p^2=.043$), nor were the interactions between discrimination and the linear and quadratic trends across block (larger $F(1,27) = 2.80$, $p=.106$, $\eta_p^2=.094$).

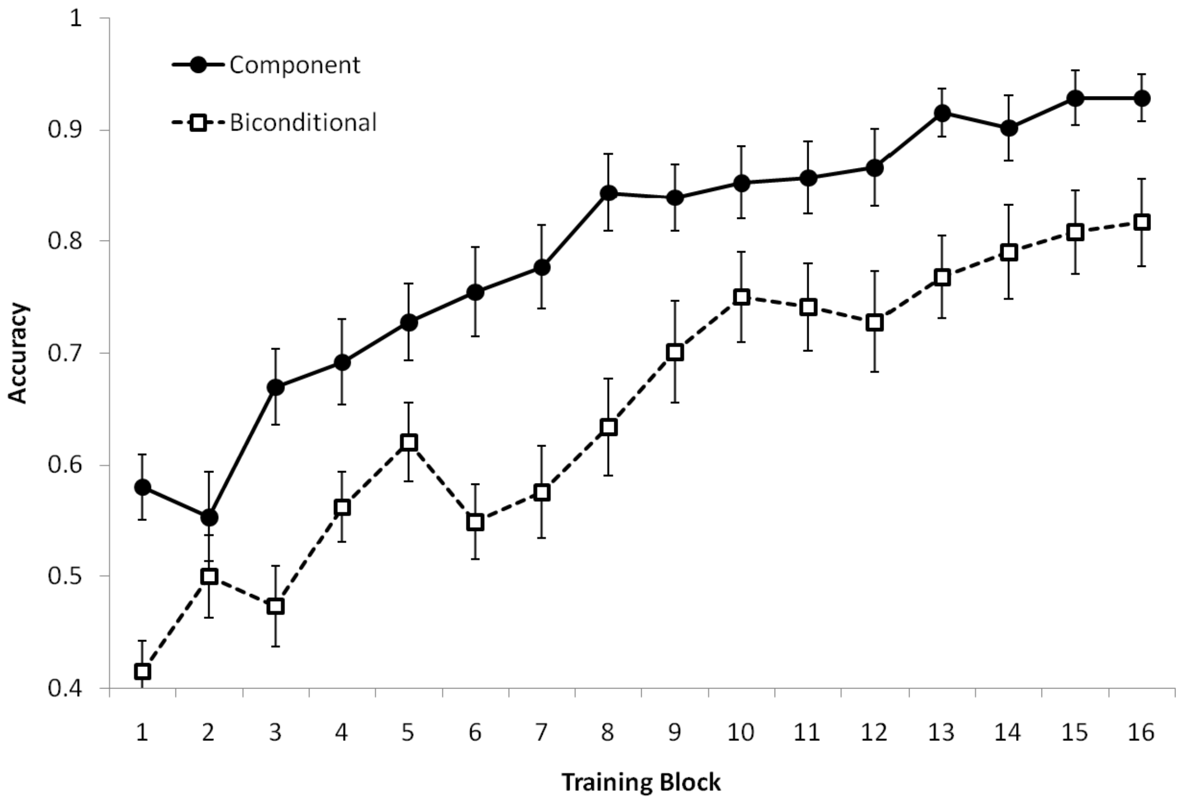

Figure 4. Experiment 2, mean accuracy for the component and biconditional discriminations across the 16 blocks of Training Stage 1. Error bars indicate SEM.

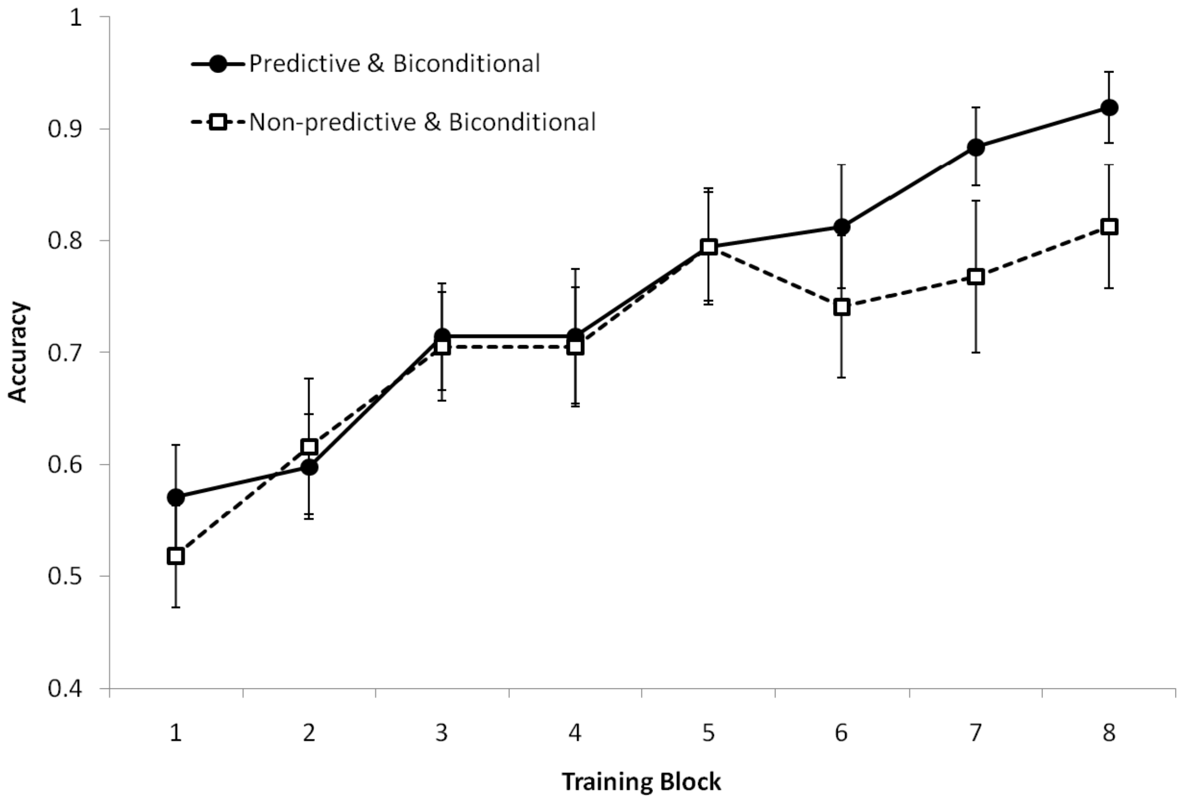

Figure 5. Experiment 2, mean accuracy across the eight blocks of Training Stage 2. All compounds were composed of one stimulus previously involved in a biconditional discrimination and one stimulus previously involved in a component discrimination. Data are averaged across compounds with previously predictive components and previously non-predictive components. Error bars indicate SEM.

Stage 2. Figure 5 shows accuracy for the recombined compounds used in Stage 2, averaged separately for the biconditional plus predictive components (Bic+P) and biconditional plus non-predictive components (Bic+NP). A repeated measures ANOVA with training block (1-8) and Stage 1 discrimination (Bic+P vs Bic+NP) as factors revealed a significant main effect of block ($F(7,189)=12.22$, $p<.001$, $\eta_p^2=.312$) and a significant effect of Stage 1 discrimination ($F(1,27)= 5.08$, $p<.033$, $\eta_p^2=.158$), indicating significantly better performance for the Bic+P components. There was no significant interaction between Stage 1 discrimination and block ($F<1$), nor interactions with linear trend ($F(1,27)= 2.64$, $p=.116$, $\eta_p^2=.089$) and quadratic trend ($F<1$) across blocks.

Test stage. Data from the outcome ratings for all test compounds are shown in Figure 6, both as mean likelihood ratings and collapsed into the outcome difference scores for each classification of test trial. As with Experiment 1, we used the outcome ratings difference scores for all statistical analyses. For the compounds trained in Stage 2, there was no statistical difference between the Bic+P compounds and the Bic+NP compounds ($F<1$). For the summation test trials, it is now valid to pit the stimulus associability of the biconditional stimuli against that of the stimuli trained in the component discriminations by comparing predictive components to the biconditional stimuli with which they were paired, and the non-predictive components to the biconditional

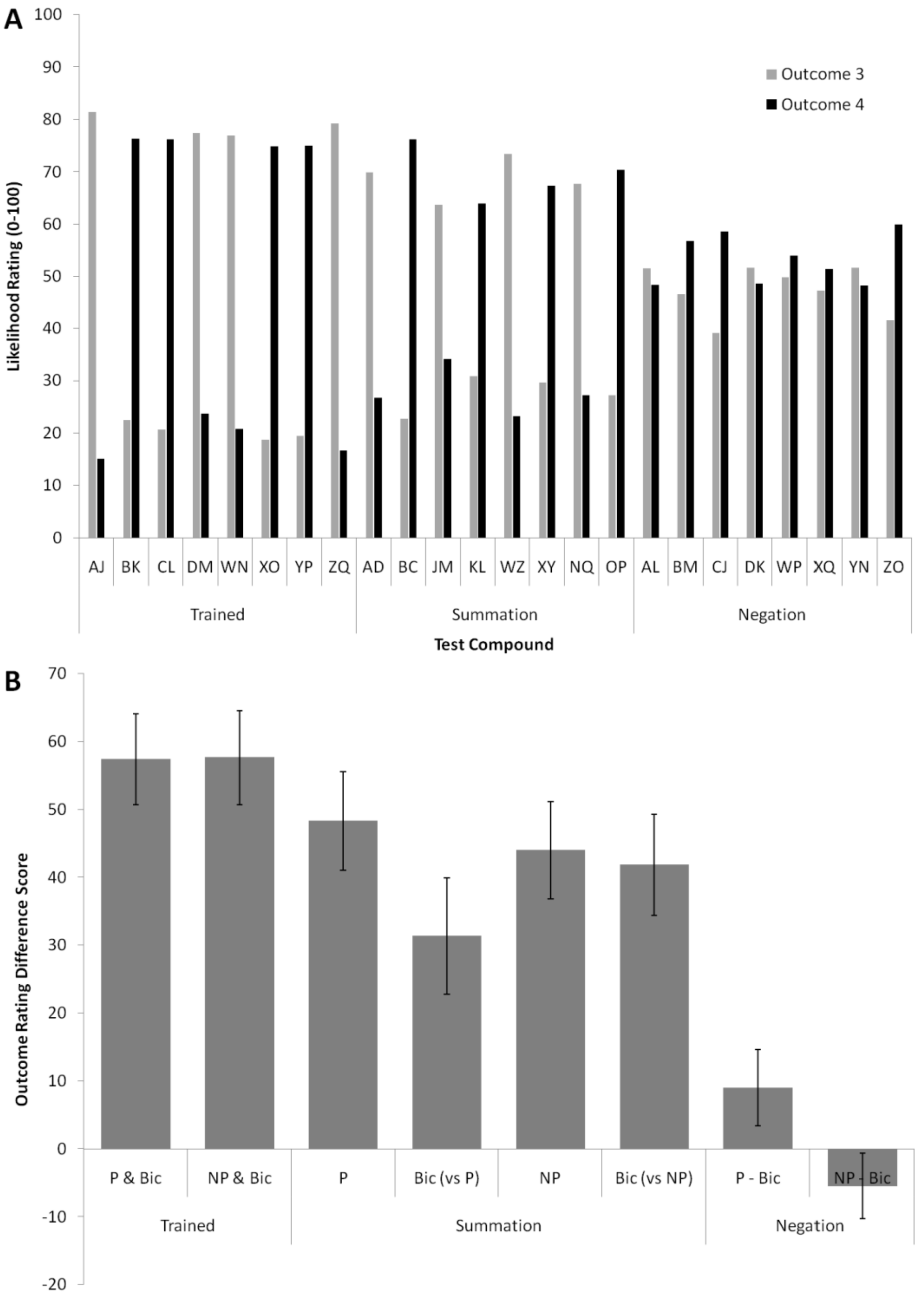

Figure 6. Experiment 2, ratings test data. A) Mean ratings of the likelihood of the Stage 2 outcomes (O3 & O4) for each of the 24 test trials. B) Test scores calculated by averaging the difference between the rating for the predicted outcome and the opposite outcome. Bic, P and NP refer to the roles of the stimuli in Stage 1 (P = Predictive components, NP = non-predictive components, Bic = Biconditional). Error bars indicate SEM.

stimuli with which those were paired. (it is worth noting that the downside is that the design of Experiment 2 is no longer optimal for comparing predictive vs non-predictive components.) With this in mind, the predictive stimuli were significantly higher than the biconditional stimuli paired with predictive components in Stage 2 (F(1,27)=5.04, p=.033, η_p^2=.157) but the non-predictive stimuli were not significantly different from the biconditional stimuli paired with non-predictive components in Stage 2 (F<1). For the negation test trials, the 'Predictive – Biconditional' scores were significantly higher than the 'Non-predictive – Biconditional' scores (F(1,27)=4.27, p=.049, η_p^2=.136), however neither of the individual scores differed significantly from zero (larger t(27)=1.60, p=.121).

Discussion

The results of Experiment 2 clearly indicate that the predictive component stimuli maintained a higher associability than the biconditional stimuli. Biconditional and non-predictive stimuli were equivalent on test ratings, suggesting that learning the biconditional discrimination has the same consequence for stimulus associability as learning about the non-predictive stimuli in the component discrimination.

The results are at odds with the stimulus relevance hypothesis that participants will attend to stimuli that were previously relevant to discrimination learning regardless of their individual associative histories. Learning the biconditional discrimination requires that attention be paid to both stimuli simultaneously. Since distributing attention across both stimuli is the most effective strategy, one might expect that each biconditional stimulus would command greater attention on subsequent tasks than a stimulus that was completely irrelevant to the discrimination (i.e. one which should be actively ignored). Clearly, this account was not supported by the data. Instead, the pattern of results suggests that the single stimuli from the biconditional discrimination received the same attention as the non-predictive stimuli. The results are therefore consistent with a stimulus predictiveness approach in which elements lose (or at least fail to gain) associability as a consequence of being paired with multiple and conflicting outcomes. The low associability of the biconditional stimuli is also consistent with the *relative* predictiveness account, if one assumes that information about the configuration of cues serves as a functional (and relatively predictive) stimulus.

It is clear in Experiment 2 that the non-predictive stimuli, each of which is routinely presented with better predictors of the outcome in Stage 1, end up with the same stimulus associability as the biconditional stimuli, which are only shown with stimuli of equivalent predictive validity. It seems that comparison with a physically isolable stimulus of higher predictive validity results in no greater loss of associability than comparison with the hypothesised configural stimulus that controls biconditional discrimination. Alternatively, a relative predictiveness account might explain the results without appealing to configural information by instead assuming that each stimulus started out with a very low associability, which only increased through comparison with a worse predictor. Hence, only the associabilities of the predictive components increase, while the associabilities of the nonpredictive and biconditional stimuli remain unchanged.

Experiment 3

Rather than using biconditional and component discriminations, Experiment 3 compared the associability of stimuli drawn from a simple discrimination, in which all stimuli within Stage 1 compounds were good predictors of the outcome, to the associability of stimuli drawn from an impossible pseudo-discrimination in which no stimulus within the Stage 1 compounds was predictive. In comparison to either the biconditional or component discriminations, the simple discrimination should be learned very rapidly because both stimuli in each compound are consistent predictors the relevant outcome.

In some respects, this experiment is similar to the design used by Le Pelley, Turnbull, Reimers, Knipe, & Murphy (2010b). In their experiments, cues were trained individually and were either consistently or inconsistently paired with outcomes in each of several training phases. Cues that led consistently to the same outcome were learned about faster in subsequent stages than cues that were initially non-predictive. This shows that, at least to some degree, single cue training results in good predictors having higher associability than poor predictors even though there is no direct competition for attention. Experiment 3 served to replicate this effect with compounds in which the two components were equally good predictors (simple discrimination) or equally non-predictive (impossible discrimination). Thus, in the present design, direct comparison between stimuli within a compound was possible but also unhelpful. The results of Le Pelley et al (2010b) suggest that stimuli that are consistently paired with only one outcome should have higher associability, and therefore their absolute predictiveness will be sufficient to provide an associability advantage over the stimuli previously involved in the impossible pseudo-discrimination.

Method

Participants and Apparatus. Twelve undergraduate psychology students (7 female, mean age = 20 years) participated in Experiment 3. Participants were group tested in a quiet laboratory using Dell Optiplex desktop computers. In all other respects, the apparatus was the same as Experiments 1 and 2.

Stimuli and Procedure. The stimuli used in the experiment were identical to Experiments 1 and 2. All aspects of the procedure were also identical to Experiment 2 except for the contingencies presented in training Stage 1, in which the participants received two simple discriminations and two impossible pseudo-discriminations. These are shown in Table 2 (Note that we use the letters A-D and W-Z for the simple stimuli as a matter of convenience: even though in this experiment they are all equally valid predictors in Stage 1, using the same letters means that the description of Stage 2 and the Test stage are identical for Experiments 2, 3, and 4). Thus the simple contingencies were AW-O1, AX-O1, BW-O1, BX-O1 and CY-O2, CZ-O2, DY-O2, DZ-O2, whereas the impossible contingencies involved the presentation of JN, JO, KN, KO, LP, LQ, MP, MQ, all followed 50% of the time by O1 and 50% of the time by O2. Training Stage 2 compounds and contingencies, and Test stage compounds were the same as in Experiment 2 (see Table 4).

Results

The data of all participants were included in the following analyses.

Stage 1. Accuracy across the course of stage 1 is shown in Figure 7. As expected, the simple discrimination was learned very quickly, reaching >90% accuracy by the sixth block. In contrast, performance on the pseudo-discriminations was never consistently above chance. The differences in performance between the two discriminations are too stark to warrant statistical comparison, and hardly suprising given that one discrimination is unsolvable.

Stage 2. The compounds trained in Stage 2 were all composed of one stimulus from the simple discrimination and one from the impossible pseudo-discrimination. Mean accuracy across the eight blocks rose from initially chance level in the first block (mean = 52.3%, SEM = 4.7%) to a reach a steady asymptote well above chance over the last four blocks (mean = 79.8%, SEM = 5.2%).

Test stage. Mean outcome likelihood ratings and outcome difference scores for Experiment 3 are shown in Figure 8. There were fewer test trial classifications in this experiment because the only relevant distinction was between stimuli used in the simple discrimination in Stage 1 and stimuli used in the impossible pseudo-discrimination in Stage 1. Looking at the summation test trials, the simple compounds were rated significantly higher than the impossible compounds ($F(1,11)=6.79$, $p=.024$, $\eta_p^2=.382$). For the negation test, the 'Simple – Impossible' compounds were significantly greater than zero ($t(11)=3.09$, $p=.010$), indicating a bias towards the outcome predicted by the simple components and away from the outcome predicted by the impossible components.

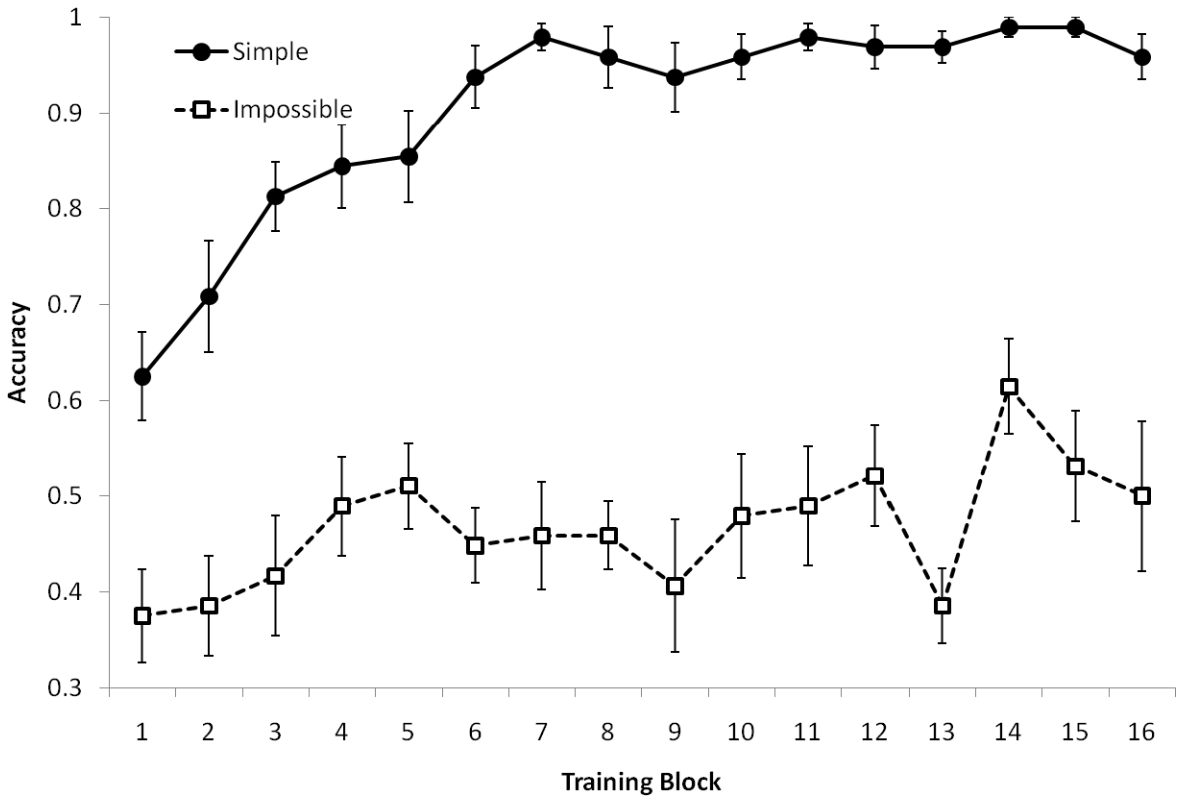

Figure 7. Experiment 3, mean accuracy for the simple discrimination and impossible pseudo-discrimination across the 16 blocks of Stage 1 training. Error bars SEM.

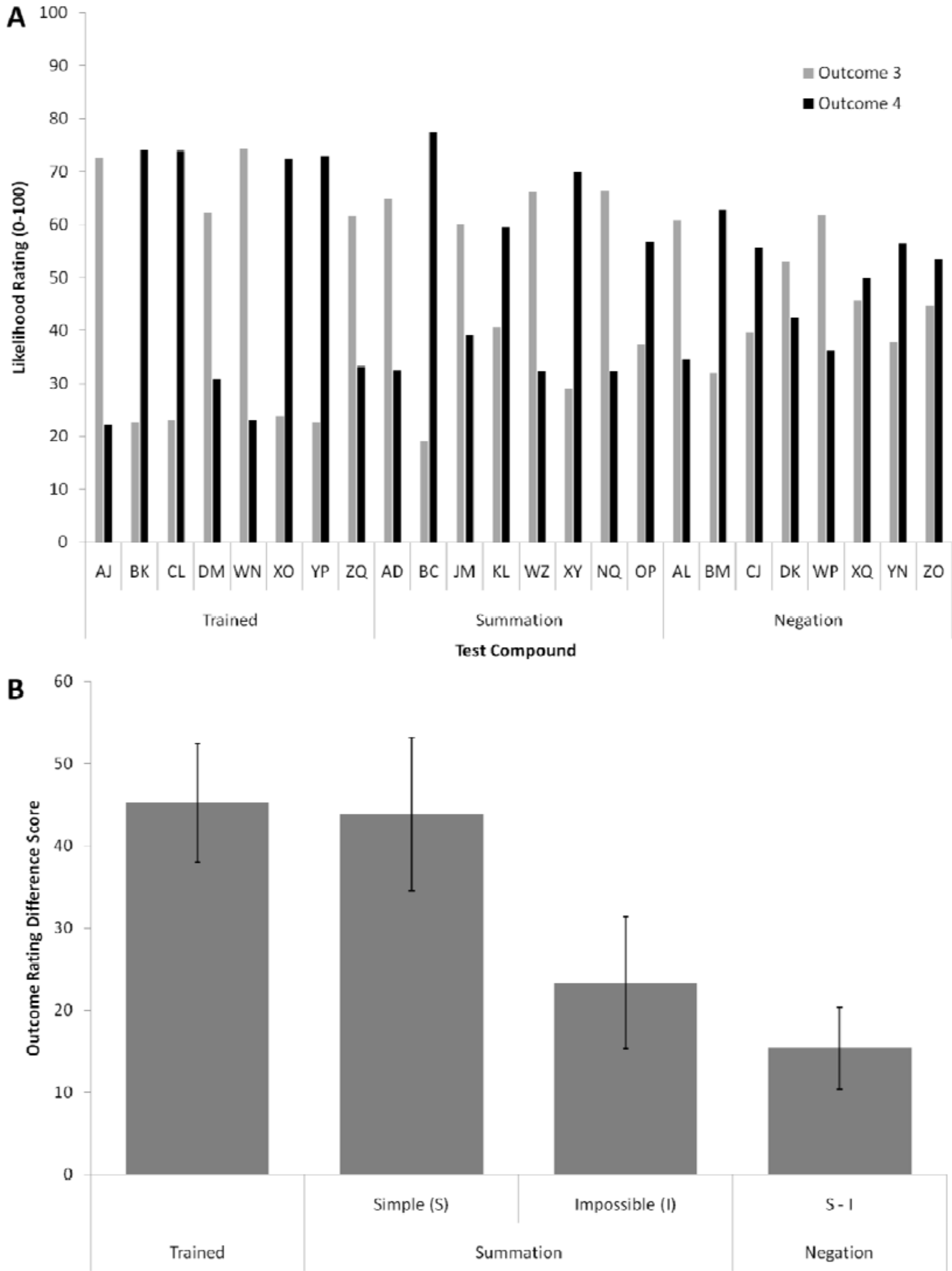

Figure 8. Experiment 3, ratings test data. A) Mean ratings of the likelihood of the Stage 2 outcomes (O3 & O4) for each of the 24 test trials. B) Test scores calculated by averaging the difference between the rating for the predicted outcome and the opposite outcome. T and I refer to the roles of the stimuli in Stage 1 (S = simple discrimination, I = impossible pseudo-discrimination). Error bars indicate SEM.

Discussion

In Experiment 3, the absolute predictiveness of the cues in Stage 1 was sufficient to drive a robust learned predictiveness effect. Cues that were originally trained in compounds of equally predictive stimuli (simple discrimination) and equally non-predictive stimuli (impossible pseudo-discrimination) showed marked differences in stimulus associability when trained in mixed compounds in Stage 2. Pitting the impossible stimuli against the simple stimuli reveals a strong learning bias for the cues from the simple discrimination. This replicates similar findings with single stimulus presentations by Le Pelley et al. (2010b).

An alternative, and very successful approach to learned attention in animal learning has been to assume that the associability of a CS is maintained or increased to the extent that its consequences are surprising and decreased to the extent that its consequences are well predicted. This assumption, described formally by Pearce and Hall (1980), seems to contradict the Mackintosh approach because it suggests that a CS will only command attention when its consequences are poorly predicted. Nevertheless, efforts to formally combine the two processes (Le Pelley, 2004; Pearce & Mackintosh, 2010) demonstrate that they can be complementary. For instance, in the model proposed by Le Pelley (2004) associability is assumed to increase as a consequence of high *relative* predictiveness – better predictors win out over worse predictors – but decline with high *absolute* predictiveness. That is, a well-predicted outcome generally fails to maintain the associability of the CSs with which it has been paired (i.e., absolute predictiveness), but a highly predictive CS can nevertheless maintain high associability through comparison with other less predictive stimuli with which it co-occurs (i.e., relative predictiveness). However, as with Le Pelley et al. (2010b), the results of Experiment 3 do not seem to support this dichotomy because differences in the absolute predictiveness across discriminations appear to yield similar changes in associability to differences in relative predictiveness within trials. That is, both favour stimuli that clearly signal an outcome. Nevertheless, to demonstrate this more conclusively, each of the two discriminations used in Experiment 3 must be compared with the component discrimination, where within-compound differences in relative predictiveness potentially contribute to learned associability effects.

Experiment 4

Experiment 4 compared the associability of predictive and non-predictive stimuli from a component discrimination against stimuli from either the simple discrimination or impossible pseudo-discrimination used in Experiment 3. The aim of the experiment was to test whether within-compound differences in relative predictive validity contribute to the learned predictiveness effect over and above what is observed between compounds of different absolute predictiveness.

The Mackintosh (1975) model assumes that relative predictiveness is critical for changes in associability. The non-predictive stimuli in the component discrimination should lose out when pitted against stimuli from either the simple or impossible discrimination because, during Stage 1, the non-predictive stimuli are consistently paired with good predictors and are therefore always the worse of the two potential signals of the outcome. Consequently the non-predictive components should lose associability faster than stimuli in either of the other two discriminations. Similarly, the predictive stimuli in the component discrimination should win out over stimuli from both the simple and impossible discriminations because, in Stage 1, they are always paired with a non-predictive stimulus and therefore always benefit from comparison with a relatively weak signal.

Experiment 4 used a two group design. Group I received the component discriminations used in Experiments 1 and 2 along with the impossible pseudo-discriminations used in Experiment 3. Group S received the component discrimination along with the simple discriminations used in Experiment 3. Both groups then received the same Stage 2 training and test as were used in Experiments 2 and 3.

Method

Participants and Apparatus. Seventy-two undergraduate psychology students (50 female, mean age = 20 years) participated in Experiment 4, each assigned to group according to time of arrival. Apparatus and testing conditions were identical to Experiment 3.

Stimuli and Procedure. The stimuli used were identical to the previous three experiments. The procedure was identical to Experiments 2 and 3 in all respects except for the Stage 1 contingencies. Both groups received two component discriminations as shown in Table 2. Group I also received two impossible pseudo-discriminations. Group S received two simple discriminations.

Results

As in Experiments 1 and 2, participants were excluded from analysis if they failed to achieve greater than 60% accuracy across the last 4 blocks of training on the component discrimination. Using this criterion, 10 participants were excluded and their data discarded for analyses (remaining n=62 participants, 31 in each group).

Stage 1. Prediction accuracy data from Stage 1 training are shown in Figure 9. Because the simple discrimination and the impossible discrimination give rise to such different levels of performance, the full 2 x (2) x (16) ANOVA of Stage 1 accuracy (with group, discrimination, and block as factors) is not very informative. Instead, we performed three separate analyses, examining each group individually for performance on the two discriminations across blocks and then comparing their performance on the component discrimination.

Analysing the data of Group I, there were significant main effects of discrimination ($F(1,30)=100.61$, $p<.001$, $\eta_p^2=.770$) and block ($F(15,450)=9.60$, $p<.001$, $\eta_p^2=.242$) and significant interactions between discrimination and block ($F(15,450)=7.13$, $p<.001$, $\eta_p^2=.192$) and between discrimination and the linear trend across block ($F(1,30)= 73.06$, $p<.001$, $\eta_p^2=.709$). As expected, performance on the component discrimination was clearly better and improved faster than the impossible pseudo-discrimination. The same analyses applied to Group S yielded significant main effects of discrimination ($F(1,30)=38.24$, $p<.001$, $\eta_p^2=.560$) and block ($F(15,450)=56.91$, $p<.001$, $\eta_p^2=.655$). The interaction between discrimination and block approached significance ($F(15,450)=2.02$, $p=.067$, $\eta_p^2=.063$) and there was a significant interaction between discrimination and the linear trend across block ($F(1,30)=13.22$, $p=.001$, $\eta_p^2=.306$). Hence, performance on the simple discriminations was better and improved significantly faster than the component discrimination.

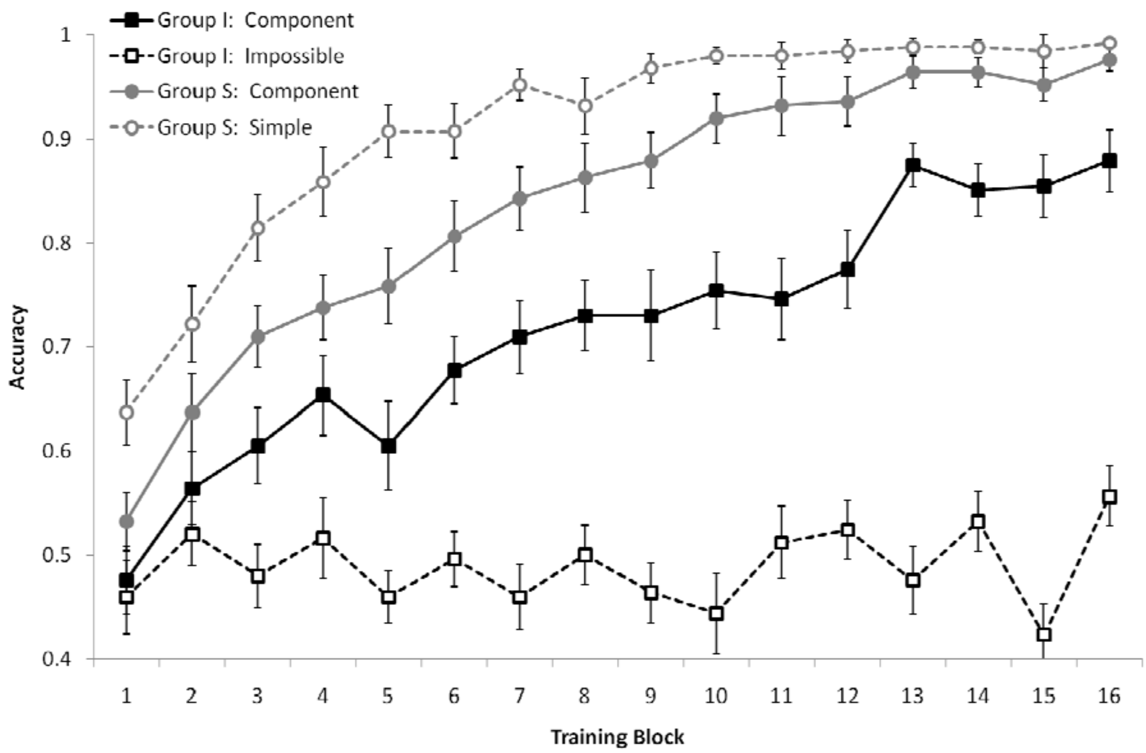

Figure 9. Experiment 4, mean accuracy during Stage 1 , as a function of group, discrimination and training block. Error bars indicate SEM.

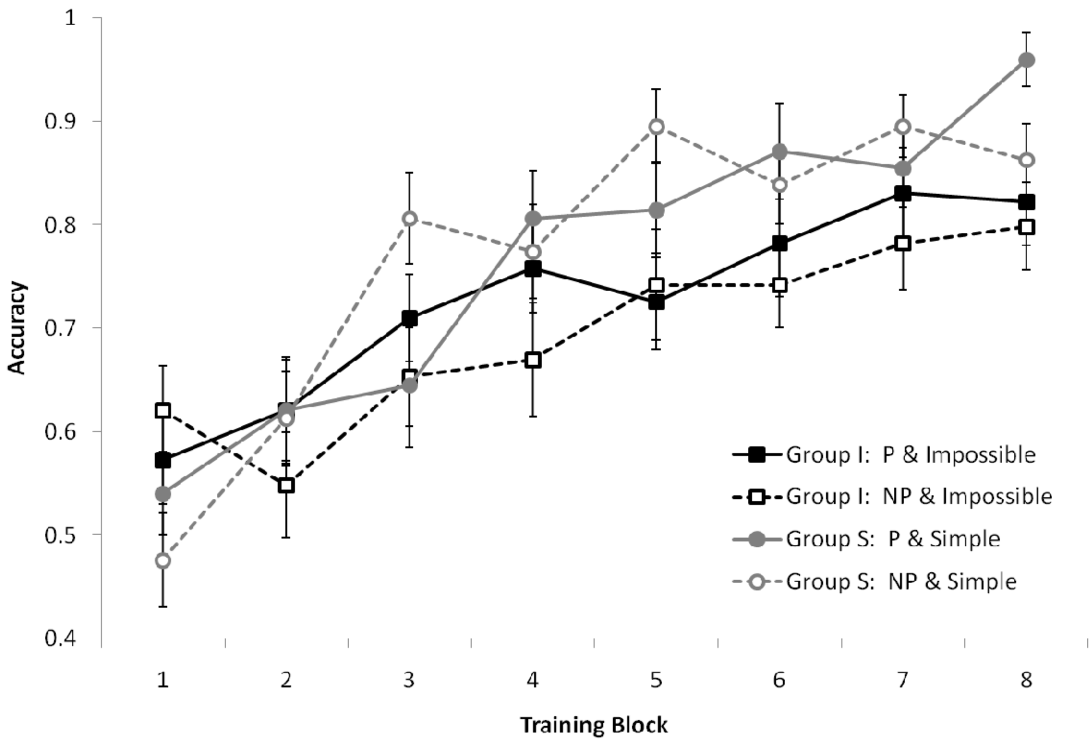

Figure 10. Experiment 4, mean accuracy in Stage 2 as a function of group and type of compound. P and NP refer to stimuli from the component discrimination in Stage 1 (P = predictive; NP = non-predictive). Error bars indicate SEM.

Finally, comparing the two groups for their performance on the component discrimination, there were significant main effects of group ($F(1,60)=20.61$, $p<.001$, $\eta_p^2=.256$) and block ($F(15,900)=45.12$, $p<.001$, $\eta_p^2=.429$). The interaction between group and block was not significant ($F(15,900)=1.01$, $p=.432$, $\eta_p^2=.017$) but the interaction between group and the quadratic trend across block did reach significance ($F(1,60)=4.75$, $p=.033$, $\eta_p^2=.073$). This suggests that Group I performed significantly worse on the component discriminations than Group S.

Stage 2. All compounds in Stage 2 contained one stimulus from the component discriminations and one from the other discriminations (either impossible or simple depending on group), with the critical within-subjects variable being whether the compound contained a predictive or non-predictive component, as shown in Figure 10. Therefore, Stage 2 was analysed with group, Stage 1 component (predictive vs non-predictive), and training block as factors. This revealed a significant main effect of training block ($F(7,420)=33.12$, $p<.001$, $\eta_p^2=.356$) but no significant effect of either group ($F(1,60)=2.21$, $p=.142$, $\eta_p^2=.036$) or Stage 1 component ($F<1$). There were significant interactions between group and training block ($F(7,420)=2.79$, $p=.014$, $\eta_p^2=.044$), between group and the linear trend across blocks ($F(1,60)=8.54$, $p=.005$, $\eta_p^2=.125$) and between group and the quadratic trend across blocks ($F(1,60)=4.40$, $p=.040$, $\eta_p^2=.068$), suggesting that Group I may have improved more slowly in Stage 2. Stage 1 component did not interact significantly with group ($F(1,60)=2.05$, $p=.158$, $\eta_p^2=.033$), with block ($F(7,420)=1.66$, $p=.143$, $\eta_p^2=.027$) or with linear or quadratic trends across block (larger $F(1,60)=1.52$, $p=.223$, $\eta_p^2=.025$). The three-way interaction between Stage 1 discrimination, group and block approached significance ($F(7,420)=2.15$, $p=.057$, $\eta_p^2=.035$), although this cannot be interpreted as a difference in the rate of learning across conditions because the three-way interaction with linear trend across blocks did not approach significance ($F<1$).

Test stage. Test ratings in full for the two groups are shown in Figure 11. As with the previous experiments, all statistical analyses are based upon ratings difference scores, as shown in Figure 12. Statistical analyses were performed separately on the trained, summation and negation test trials.

For the trained test trials, an analysis with Group as between-subjects factor and Stage 1 component (P vs NP) as within subjects factor revealed that the difference between groups fell just short of the conventional level of significance ($F(1,60)= 3.39$, $p=.071$, $\eta_p^2=.053$). As is evident in Figure 12, Group I performed slightly worse than Group S, though this difference was not significant. There was no effect or interaction with Stage 1 component (Fs < 1).

The summation test scores were initially analysed with a 2 x (2) x (2) ANOVA, with Group, Stage 1 discrimination (Component vs Other), and component (P vs NP) as factors. This revealed a significant main effect of group ($F(1,60)= 9.85$, $p=.003$, $\eta_p^2=.141$), again indicating that Group I performed worse overall, but no main effects of discrimination or component (Fs < 1). There was a significant interaction between Stage 1 discrimination and component ($F(1,60) =9.08$, $p=.004$, $\eta_p^2=.131$), indicating that the difference in performance between 'component' stimuli and 'other' stimuli depended on whether the component was predictive or non-predictive. This can be seen as replicating an overall learned predictiveness effect. There was also a significant interaction between group and Stage 1 discrimination ($F(1,60) =4.04$, $p=.049$,

η_p^2=.063), indicating that performance for the component stimuli was relatively good against 'impossible' stimuli and relatively poor against the 'simple' stimuli. No other interactions were significant (highest $F(1,60)$ =1.06, p=.306, η_p^2=.017).

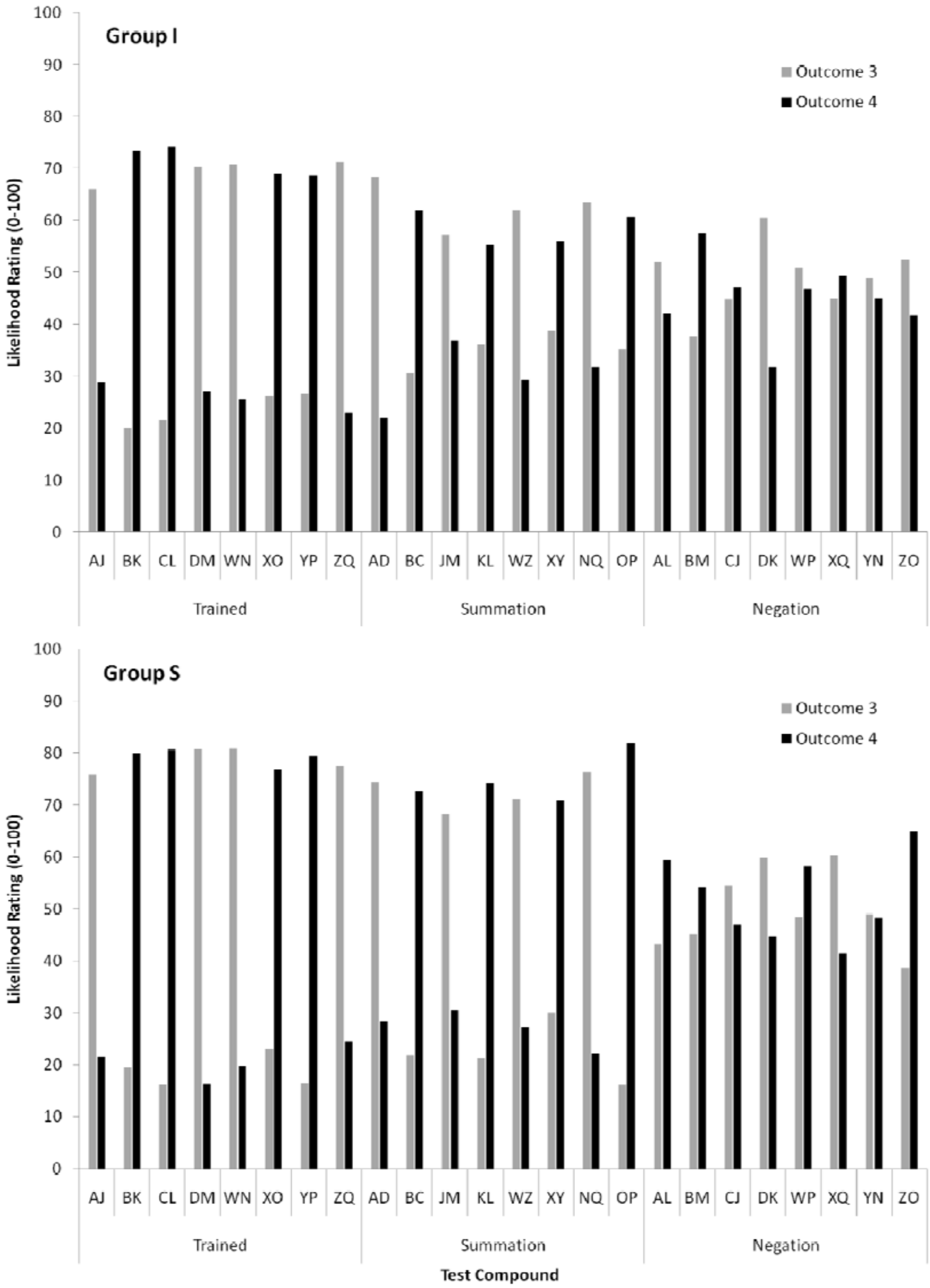

Figure 11. Experiment 4, mean test ratings of the likelihood of the Stage 2 outcomes (O3 & O4) for each of the 24 test trials. Top panel: ratings for Group I; Bottom panel: ratings for Group S.

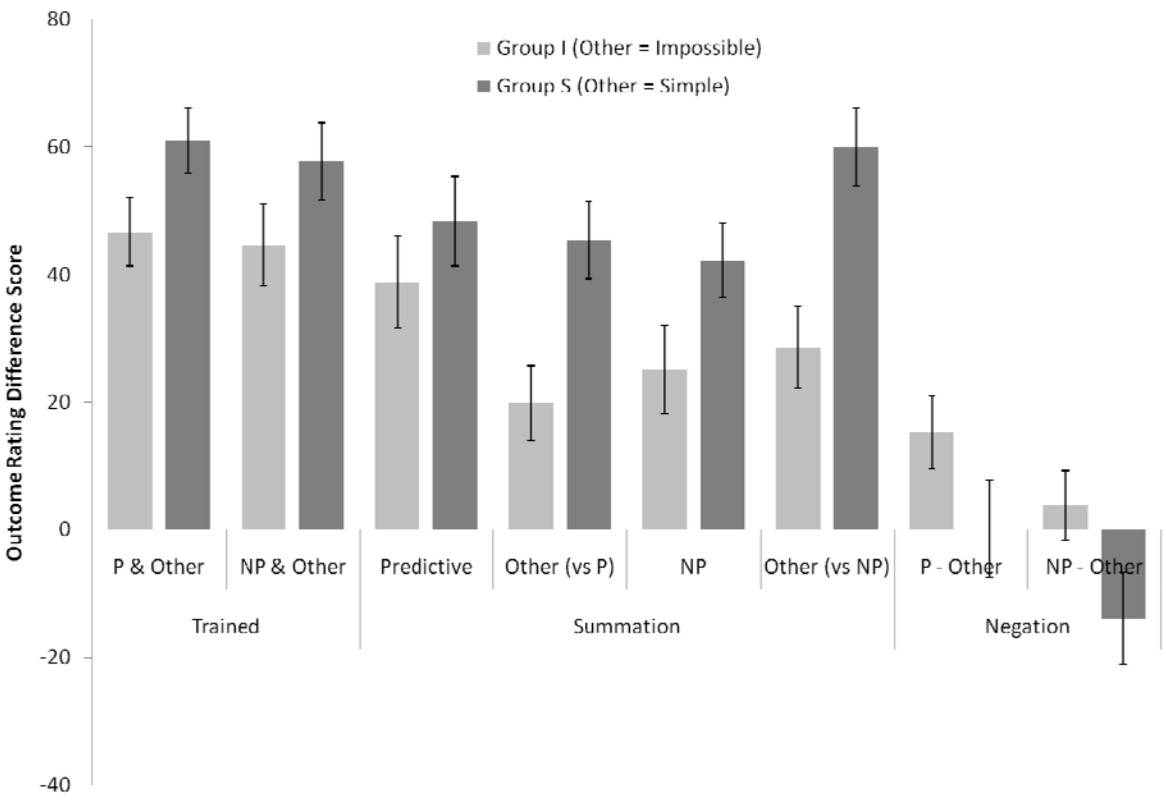

Figure 12. Experiment 4 test phase, outcome rating difference scores. P, NP, and Other refer to the roles of the stimuli in Stage 1 (P = Predictive components, NP = non-predictive components, Other = Impossible pseudo-discrimination or Simple discrimination, depending on Group). Error bars indicate SEM.

Learned predictiveness effects in each group were examined for the predictive components and non-predictive components individually (in each case, compared to the ‘other’ stimuli with which they were trained in Stage 2). For Group I, difference scores were significantly higher for the predictive stimuli than for the impossible stimuli ($F(1,30)=6.34$, $p=.017$, $\eta_p^2=.175$), but non-predictive stimuli did not differ from the impossible stimuli with which they were trained ($F<1$). For Group S, predictive stimuli did not differ significantly from the simple stimuli with which they were trained ($F<1$) but difference scores for the non-predictive stimuli were significantly lower than for the simple stimuli ($F(1,30)=6.40$, $p=.017$, $\eta_p^2=.176$).

Negation test difference scores are coded so that a positive score indicates a bias towards the outcome predicted by the component discrimination stimuli and a negative score indicates a bias towards the outcome predicted by stimuli from the other discrimination (Impossible or Simple). Analysis of the negation test trials was conducted using an ANOVA with group and component (predictive vs non-predictive) as factors. The effect of component fell just short of conventional significance ($F(1,60)=3.77$, $p=.057$, $\eta_p^2=.059$), indicating a trend towards the negation trials with previously predictive components being more positive than the negation trials with previously non-predictive trials. This again reflects a standard learned predictiveness effect. There was also a significant effect of group ($F(1,60)=6.15$, $p=.016$, $\eta_p^2=.093$), indicating that negation scores were more positive in Group I than Group S. This indicates a learned predictiveness effect for absolute predictive validity as ratings were more biased towards the Stage 2 outcome signalled by the component discrimination stimuli when the other discrimination was impossible than when it was trivially easy. The interaction between the two factors did not approach significance ($F<0.1$). Each of the negation scores was the examined individually for each group. For Group I, the 'predictive - impossible' negation score was significantly greater than zero ($t(30)=2.68$, $p=0.012$), indicating a bias towards the predictive components, whereas the 'non-predictive - impossible' negation score did not differ significantly from zero ($t<1$), suggesting no consistent bias towards either the non-predictive components or the impossible discrimination stimuli. For Group S, the 'predictive - simple' negation score did not differ significantly from zero ($t<1$), suggesting no systematic bias towards either the predictive or simple stimuli. The 'non-predictive - simple' negation score fell just short of being significantly less than zero ($t(30)=1.91$, $p=0.065$), suggesting a non-significant bias towards the simple stimuli .

Discussion

The summation and negation tests in Experiment 4 reveal learned predictiveness effects based on the relative validity of the predictive vs non-predictive components and learned predictiveness effects based on the absolute validity of the simple vs impossible discrimination stimuli. However, there was no evidence that predictive components sustained higher associability than the simple stimuli, or that the non-predictive components sustained lower associability than the impossible discrimination stimuli. Therefore, the opportunity to compare a more predictive stimulus against a less predictive one did not contribute to learned changes in the associability of those stimuli.

The two groups in Experiment 4 display an overall difference in performance which can be seen in learning the component discrimination during Stage 1 as well as both the trained and summation test trials. In Stage 1, half of the trials received by Group I are impossible to solve, whereas Group S instead received a very easy task. This could well have an effect on the motivation of participants to attend to the stimuli or task in general or their confidence or certainty in making predictions about which outcome is likely to occur next. Importantly, learned predictiveness effects are still evident in both groups. In Group I, predictive stimuli display higher associability than impossible stimuli (which did not differ from non-predictive stimuli). In Group S, non-predictive stimuli displayed lower associability than the simple stimuli (which did not differ from the predictive stimuli). This suggests that the learned predictiveness effects cut across factors which affect performance overall.

In models of associability, the associative strength of the stimulus typically dictates changes in associability, either by affecting the relative or the absolute predictiveness of the stimulus in question. In this sense, it is predictiveness derived through learning rather than the objective predictiveness defined by the experimenter that counts. Although the predictive components and simple stimuli were both highly predictive, they may not have developed associations of the same strength. Differences in Stage 1 performance on the two discriminations in Group S might be taken to mean that those participants learned more about the simple stimuli than about the component stimuli. On the other hand, learning about any single stimulus from the simple discrimination may well have been limited to some degree by overshadowing, such that the competition between two predictive cues prevented any single cue from attaining maximum associative strength. Rather than being matched for associative strength, what the simple and predictive stimuli share in common is that, in Stage 1, they were unambiguously associated with only one outcome. In this sense, their signal validity is high in an absolute sense, sufficiently so to maintain comparably high associability during Stage 2 learning. Regardless of their associative strength, in Stage 1 the simple stimuli were presented with a stimulus of equal predictiveness whereas the predictive components were presented with much less predictive stimuli. Thus, where one would expect to see a difference in associability between simple and predictive stimuli on the basis of *relative* predictiveness during learning, there is none.

General Discussion

In four experiments, participants displayed more learning for cues that were strongly predictive of one outcome in an earlier discrimination than non-predictive cues that were paired equally often with two outcomes in an earlier discrimination. This simple learned predictiveness effect was unaffected by the opportunity to compare within-compound stimuli of different predictive validities. Instead, the absolute predictiveness of the cues seemed to affect associability to a consistent degree across a variety of Stage 1 discriminations that manipulated the relative and absolute predictiveness of the cues.

In Experiment 1, we replicated the standard learning advantage for previously predictive cues over previously non-predictive cues within a design where participants also concurrently learned two biconditional discriminations. In Experiment 2, stimuli from a biconditional discrimination were subsequently trained in compound with either previously predictive components or previously non-predictive components. Previously predictive components displayed an associability advantage over the biconditional stimuli, whereas biconditional and non-predictive stimuli displayed equivalent associability. In Experiment 3, we used a very simple and an impossible (unsolvable) training schedule in which there were no differences in the relative predictiveness of within-compound stimuli but the absolute predictiveness of the stimuli differed greatly between discriminations. Cues from the simple discrimination (that had been predictive, but no more so than the other stimuli with which they were paired) displayed an associability advantage over cues from the impossible discrimination (that had been non-predictive, but no less so than the cues with which they were paired). In Experiment 4, each of the Experiment 3 training schedules was pitted against the component discrimination so that the relative validities of predictive and non-predictive stimuli can be compared within-compound on each trial. Although previously predictive cues from both the simple and component discriminations showed an associability advantage over previously non-predictive cues from the component and impossible discriminations, there were no differences between the simple cues and the predictive cues, and there were no differences between the non-predictive and impossible cues. Across the experiments, learned predictiveness effects occurred whether each cue was trained with a more predictive, less predictive or equivalent stimulus. Therefore it appears that these effects were not influenced by the opportunity to directly compare the relative predictiveness of the cues.

The present findings cast some doubt on the mechanisms responsible for learned predictiveness, specifically whether stimulus associability is in fact a consequence of direct stimulus comparison. Relative predictive validity – derived solely by comparing simultaneously presented stimuli – was a critical variable in the Mackintosh (1975) model. Comparison of relative prediction error was responsible for generating blocking and overshadowing, as well as positive and negative transfer effects. More recent models that incorporate attention or associability tend not to rely on stimulus comparison so heavily. For instance, the Le Pelley (2004) model anticipates blocking as a consequence of a summed error term which restricts the associability of the outcome, *as well as* reduction in the associability of the blocked cue as a consequence of a stimulus comparison mechanism. Recent evidence suggests that the associability of a blocked cue is reduced in much the same way as an explicitly non-predictive cue. For instance, Le Pelley, Beesley and Suret (2007) used a compound testing procedure very similar to the one employed here to demonstrate that blocked cues are less readily associated with a new outcome than overshadowing control cues. However, if blocking is achieved by a non-attentional process (e.g. Rescorla & Wagner, 1972) then this result is also consistent with an account of associability based on absolute associative strength, since the blocked cues acquire weaker associations with the original outcomes than do the control cues. Many of the results that are well accounted for by the Mackintosh model could still be explained by appealing to a mechanism based on absolute predictiveness (a point that Le Pelley et al., 2010b have also noted in relation to the learned predictiveness effect), provided a summed error term is also incorporated. However, perhaps the biggest challenge for this approach is to explain why high absolute predictiveness drives associability up while, at least in some circumstances, absolute predictiveness appears to drive associability down, as formalised in the Pearce-Hall (1980) model.

So far, our discussion of learned predictiveness and the predictions of the Mackintosh (1975) model has not taken into consideration any influence of the context on associability. Contextual cues can be viewed as an additional stimulus that might compete for associative control and, at least according to the relative predictiveness hypothesis, may influence stimulus associability. In Experiment 4, the fact that Group I performed worse than Group S on the component discrimination might indicate stronger associations with the context (and hence greater interference) in Group I. This is an important factor to consider if the predictions made by the Mackintosh (1975) model are altered by taking the context into account when calculating V_X. Take the impossible pseudo-discrimination as one example. On a JN – O1 trial, if we assume that $V_J = V_N$ but also that the context has acquired some associative strength ($V_{context}$) with O1, then $| \lambda - V_J | > | \lambda - (V_N + V_{context}) |$. When the stimuli do not predict an outcome consistently, their associations are likely to be relatively weak and the association between context and outcomes relatively strong, meaning the difference between $| \lambda - V_J |$ and $| \lambda - (V_N + V_{context}) |$ would be particularly pronounced. Consequently, one might conclude that the Mackintosh model amply predicts the lower associability of the impossible stimuli relative to the simple stimuli in Experiment 3 and associability equivalent to the non-predictive components in Experiment 4.

There are, however, two problems with the account described above. First, the application of context in this fashion also predicts that the associability for the simple stimuli will decrease. As long as the context retains at least some associative strength, then the associative strength of one predictive stimulus will be less than the summed associative strength of the other predictive stimulus and the context. This clearly does not match the observed pattern in Experiment 4, where the simple stimuli maintain the same associability as the predictive components. Second, this analysis ignores the influence of the context on learning about the omitted outcome.[1] Consider again for a moment a JN – O1 trial in the impossible discrimination. For O1, $\lambda = 1$ and $| \lambda - V_J | > | \lambda - (V_N + V_{context}) |$ where V_X is assumed to be the sum of $V_N + V_{context}$. However, on the very same trial, participants learn about the omission of O2, which should be expected to the same extent as O1. In relation to O2, $\lambda = 0$ and $| \lambda - V_J | < | \lambda - (V_N + V_{context}) |$ because the context is also associated with O2. Hence, to the extent that J is a worse predictor of O1, it is also a better predictor of the omission of O2. Therefore, it is far from clear that taking context into consideration will substantially change the predictions of the model in any circumstance where the context predicts two outcomes equally.

In addition to investigating how relative and absolute predictiveness influence associability, Experiment 2 also investigated whether it is enough that a stimulus be relevant for solving a discrimination in order to observe learned predictiveness effects. However, the associability of biconditional stimuli was demonstrably lower than predictive stimuli and equivalent to non-predictive stimuli. This result is inconsistent with an account of learned predictiveness which assumes that participants maintain attention to those aspects of a discrimination which were useful or relevant in the past, regardless of their individual associative history (e.g. George & Pearce, 1999). If this were the case, then the associabilities of the biconditional stimuli should have been higher than the non-predictive components. From the perspective of elemental learning, it also suggests that the individual stimuli lose associability because they do not predict any particular outcome in their own right. It is tempting to assume that other aspects of the representation of the biconditional compounds benefit from commensurate increases in associability as they gradually acquire discriminative control. However, this may prove challenging to test because of their poorly specified nature. For instance, information about a particular configuration of stimuli changes as soon as the configuration changes, much more so than an isolable stimulus presented in different compounds.

In conclusion, these results suggest that direct comparison between relatively good and relatively poor predictors of an outcome is not only unnecessary for the production of learned predictiveness effects but may contribute nothing to the effect at all. This casts doubt, at least with respect to human associative learning, on the comparison of relative predictiveness as the source of learned associability effects, as is assumed by the Mackintosh (1975) model and several more recent attention-based learning models. The results are consistent with a general notion that individual stimuli which have previously predicted an outcome maintain higher associability than stimuli which, when considered in isolation, predict no particular outcome. However, the fact that stimuli previously involved in a biconditional discrimination appeared to have low associability in Experiment 2 suggests that

[1] Mackintosh (1975) was quite explicit in stating that associability could be governed by learning about non-reinforcement, or the omission of an anticipated outcome, and therefore it seems appropriate that it should be modelled in this fashion in the more complex situation with two competing outcomes.

learned predictiveness is not just a consequence of maintaining attention to stimuli that were relevant for solving the initial discrimination. Rather it suggests that associability partly reflects an evaluation of the stimulus in terms of its individual predictiveness.

References

George, D. N., & Pearce, J. M. (1999). Acquired Distinctiveness is Controlled by Stimulus Relevance Not Correlation with Reward. *Journal of Experimental Psychology: Animal Behavior Processes, 25*, 363-373.

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

Harris, J. A., & Livesey, E. J. (2010). An attention-modulated associative network. *Learning & Behavior, 38*, 1-26. doi: Doi 10.3758/Lb.38.1.1

Harris, J. A., Livesey, E. J., Ghareai, S., & Westbrook, R. F. (2008). Negative patterning is easier than a biconditional discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 34*, 494-500.

Kruschke, J. K. (2001). Toward a Unified Model of Attention in Associative Learning. *Journal of Mathematical Psychology, 45*, 812-863.

Le Pelley, M. E. (2004). The role of associative history in models of associative learning: A selective review and a hybrid model. *Quarterly Journal of Experimental Psychology, 57B*, 193-243.

Le Pelley, M. E., Beesley, T., & Suret, M. B. (2007). Blocking of human causal learning involves learned changes in stimulus processing. *Quarterly Journal of Experimental Psychology, 60*, 1468-1476.

Le Pelley, M. E., & McLaren, I. P. L. (2003). Learned associability and associative change in human causal learning. *Quarterly Journal of Experimental Psychology, 56B*, 68-79.

Le Pelley, M. E., Reimers, S. J., Calvini, G., Spears, R., Beesley, T., & Murphy, R. A. (2010a). Stereotype Formation: Biased by Association. *Journal of Experimental Psychology: General, 139*, 138-161.

Le Pelley, M. E., Turnbull, M. N., Reimers, S. J., Knipe, R. L., & Murphy, R. A. (2010b). Learned predictiveness effects following single-cue training in humans. *Learning & Behavior, 38*.

Liljeholm, M., & Balleine, B. W. (2008). It's elemental my dear Watson. *Behavioural Processes, 77*, 434-436.

Livesey, E. J., & Boakes, R. A. (2004). Outcome additivity, elemental processing and blocking in human causality judgements. *Quarterly Journal of Experimental Psychology, 57B*, 361-379.

Livesey, E. J., & Harris, J. A. (2008). What are flexible representations? Commentary on Melchers, Shanks and Lachnit. *Behavioural Processes, 77*, 437-439.

Livesey, E. J., & McLaren, I. P. L. (2007). Elemental Associability Changes in Human Discrimination Learning. *Journal of Experimental Psychology: Animal Behavior Processes, 33*, 148-159.

Lochmann, T., & Wills, A. J. (2003). Predictive history in an allergy prediction task Proceedings of EuroCogSci 03: *The European Conference of the Cognitive Science Society* (pp. 217-222). Mahwah, NJ: Lawrence Erlbaum.

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

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: DOI 10.1016/j.beproc.2007.09.013

Pearce, J. M., & Hall, G. (1980). A model of Pavlovian conditioning: Variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review, 87*, 332-352.

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

Saavedra, M. A. (1975). Pavlovian compound conditioning in the rabbit. *Learning and Motivation, 6*, 314–326.

Shanks, D. R., Lachnit, H., & Melchers, K. G. (2008). Representational flexibility and the challenge to elemental theories of learning: Response to commentaries. *Behavioural Processes, 77*, 451-453. doi: DOI 10.1016/j.beproc.2007.09.005

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 174-215.

Suret, M., & McLaren, I. P. L. (2005). Elemental representation and associability: An integrated model. In A. J. Wills (Ed.), *New directions in human associative learning* (pp. 155-187). Mahwah, NJ: Lawrence Erlbaum Associates.

Thorwart, A. & Lachnit, H. (2010). Generalization decrements: Further support for flexibility in stimulus processing. *Learning & Behavior, 38*, 367-373. doi: doi:10.3758/LB.38.4.36

Urcelay, G. P., & Miller, R. R. (2009). Potentiation and Overshadowing in Pavlovian Fear Conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 35*, 340-356. doi: DOI: 10.1037/a0014350

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 complex stimuli. *Journal of Experimental Psychology: Animal Behavior Processes 25*, 461 – 474.