What determines the learned predictiveness effect? Separating cue-outcome correlation from choice relevance

Jessica C. Lee (jessica.lee@unsw.edu.au)
School of Psychology, University of New South Wales Sydney
Kensington, NSW 2052 Australia

Justine K. Greenaway (jgre7001@uni.sydney.edu.au)
Evan J. Livesey (evan.livesey@sydney.edu.au)
School of Psychology, University of Sydney
Sydney, NSW 2006 Australia

Abstract

Evidence from a variety of learning tasks suggests that cues that are more predictive of an outcome attract greater attention and are learned about more effectively in subsequent tasks. We tested whether this learned predictiveness effect is due to the objective strength of the cue-outcome association (cue-outcome correlation), or the degree to which the cue is informative for making the correct choice on each trial (choice relevance), by manipulating the possible outcome choices available on each trial. Experiment 1 compared two sets of cues that were equally (and imperfectly) correlated with outcomes and showed learning biases in favor of the set of cues that had initially been more relevant for choices made on each trial. Experiment 2 used a more conventional learned predictiveness design in which the cue-outcome correlation was stronger for one set of cues (perfect predictors) than the other set (imperfect predictors). However, here we manipulated whether or not the imperfect predictors could be used to make a correct choice, and thus whether the imperfect predictors possessed choice relevance that was equal to or less than the perfect predictors. In this case, we found no evidence that the relevance manipulation made any difference; learning biases towards the perfect predictor were evident regardless. The results suggest that both cue-outcome correlation and choice relevance can lead to changes in associability and learning biases; both were individually sufficient but neither were necessary.

Keywords: attention; associative learning; learned predictiveness; associability; choice relevance

Introduction

The *learned predictiveness* effect describes the tendency for cues that are good predictors of outcomes to attract more attention than cues that are poor predictors of outcomes. This attentional bias towards predictive stimuli persists, generating biases in new learning, even in situations where it is no longer relevant or warranted. Formal theories of learning account for this phenomenon as a change in the attention paid to cues (Kruschke, 1996; 2001) or their *associability* (Mackintosh, 1975), which acts as a boost to the learning rate, typically combining multiplicatively with prediction error (see Don, Beesley & Livesey, 2019).

Le Pelley and McLaren (2003) first reported the learned predictiveness effect using the design shown in Table 1 (see also Lochman & Wills, 2003). In this design, participants completed a causal learning task where they played the role of an allergist determining the cause of a fictitious patient's allergic reactions. On each trial, participants were presented with compounds of two cues, and asked to predict which of two possible outcomes they thought would occur (O1 and O2). In this design, stage 1 serves to establish the learned predictiveness bias, and stage 2 serves as a test to determine whether this bias transfers to a new set of contingencies.

Table 1: Stage 1 training, stage 2 training, and test trials used by Le Pelley & McLaren (2003).

Stage 1	Stage 2	Test	LP effect
AW – O1	AY – O3	AD	Prediction of O3:
AX – O1	BZ – O4	BC	AD > XY
BW – O2	CW – O4	XY	
BX – O2	DX – O3	WZ	
CY – O1	EF – O3	EH	Prediction of O4:
CZ – O1	GH – O4	FG	BC > WZ
DY – O2	IJ – O3	IJ	
DZ – O2	KL – O4	KL	

In the typical design, half of the cues in stage 1 are perfect predictors (A-D; paired with the same outcome 100% of the time), and the other half are imperfect predictors (W-Z; paired with both O1 and O2, each 50% of the time), arranged so that one cue in each compound is a perfect predictor and the other an imperfect predictor. For example, on some trials apple and watermelon (AW) are presented and led to a fever (O1). On other trials, banana and watermelon (BW) are presented and led to a rash (O2). In this example, watermelon (W) is an imperfect predictor as it sometimes leads to fever and sometimes to rash (paired with O1 and O2). On the other hand, A is a perfect predictor of fever (O1) and B a perfect predictor of rash (O2).

Critically, the second stage of training involves new outcomes (O3 and O4) occurring in a new context (a new patient suffering different allergic reactions), rendering any learning about cues in stage 1 irrelevant. Each trial type in stage 2 consists of a perfect predictor and an imperfect predictor from stage 1, but now *both* of these are perfectly predictive of a new outcome. If participants are able to disregard what they have learned in stage 1, then there should be no difference in what participants learn about each set of cues (A-D vs. W-Z) in stage 2.

Instead, Le Pelley and McLaren (2003) found that a learning bias persisted in stage 2. In a subsequent test phase, participants were presented with compounds consisting of 2 cues that were previously predictive in stage 1 (e.g., AD, see Table 1), or 2 cues that were previously non-predictive in stage 1 (e.g., XY, see Table 1) and asked to rate to what extent these compounds predicted the outcomes presented in stage 2 training (O3 and O4). Participants showed better learning for the compounds consisting of cues that were previously predictive (AD, BC) than for compounds of cues that were previously non-predictive of O1 and O2 (WZ, XY). This bias has been replicated many times. It has been shown when cues are tested individually (e.g. Shone, Harris & Livesey, 2015). It coincides with a number of independent measures of overt and covert attention (e.g. Le Pelley, Beesley & Griffiths, 2011). It also displays at least some resistance to instruction (e.g. Don & Livesey, 2015). Evidence to date thus suggests the learned predictiveness effect is pervasive and relatively automatic.

Le Pelley et al. (2016) have recently proposed that many learned attentional changes (including the learned predictiveness effect) can be explained by assuming that attention to a cue is proportional to its strongest association with an outcome. However, evidence from other learning procedures suggests that attention might be directed to cues that are relevant for solving a discrimination, but are not necessarily correlated with an outcome. For instance, in a biconditional discrimination (AB+/BC-/CD+/DA-), the outcome can be predicted on the basis of the combination of two cues even though neither cue in isolation is strongly associated with a particular outcome. In some instances, features that are relevant for solving a biconditional discrimination appear to be attended more strongly in a subsequent learning task (Kruschke, 1996a; Uengoer & Lachnit, 2012). There are conflicting results, however, that suggest that discrete cues like those used in the original learned predictiveness paradigm may actually suffer the *opposite* fate, losing attention as a consequence of being involved in a biconditional discrimination (Livesey et al., 2011; 2019).

In the original learned predictiveness design (see Table 1), the predictive cues are perfectly correlated with their respective outcomes, but they are also the most relevant cues to attend to in order for participants to respond correctly on each trial. Since there are only two possible outcomes in stage 1 of the typical design (see Table 1), the cue-outcome correlation and relevance of the cues for responding accurately are confounded. Therefore, it is unclear whether the learned predictiveness effect in these tasks is determined by differences in cue-outcome correlation, differences in the relevance of the cue for making correct choices, or both.

The aim of the current study was to examine the locus of the learned predictiveness effect by separately manipulating the cue-outcome correlation and the relevance of the cues to making the correct response on each trial (*choice relevance*). We define choice relevance as the extent to which a cue is informative in identifying the correct response among the options available on a particular trial. Note that in these experiments, cues were either fully relevant (in principle, the choice could be made with 100% accuracy on the basis of the cue alone) or fully irrelevant (the cue in isolation predicted multiple outcome choices equally). This was achieved by increasing the number of possible outcomes in stage 1 training from two to four, and making only a subset of the outcome choices available on any one trial. For example, imagine that cue A leads to O1 50% of the time and O2 50% of the time but never O3 or O4 (i.e., A is an imperfect predictor of both O1 and O2). If O1 and O2 are presented as choices on a given trial, cue A is uninformative as it is equally correlated with both outcomes. However, if given a choice between O1 and O3, the partial correlation with O1 can be used to answer correctly 100% of the time. Thus, a cue that is an imperfect predictor of a given outcome can be highly relevant to making the correct outcome choice.

Experiment 1 tested whether choice relevance alone was sufficient to produce a learned predictiveness effect when all cues were imperfect predictors. Experiment 2 tested whether manipulating choice relevance had any impact on associability when one set of cues was more predictive than the other (emulating the typical conditions of the learned predictiveness effect). As in Le Pelley and McLaren (2003), we used competitive learning in stage 2 of training to test for changes in associability as a result of stage 1 training.

Experiment 1

The aim of Experiment 1 was to test whether choice relevance alone could drive changes in associability. Table 2 shows the design of Experiment 1 and Figure 1 shows a schematic of how cue-outcome correlation and choice relevance were separated in the design.

Table 2: Design of Experiment 1.

Stage 1		Stage 2		Test
Trials	Choices	Trials	Choices	
AW - 1	1 v 3	AZ - 6	5 v 6	All cues
BX - 2	2 v 4	BY - 5	5 v 6	individually
CY - 3	3 v 2	CW - 5	5 v 6	tested for
DZ - 4	4 v 1	DX - 6	5 v 6	stage 2
AY - 2	2 v 3			learning
BZ - 1	1 v 4			then for
CX - 4	4 v 2			stage 1
DW - 3	3 v 1			learning

In Table 2, all of the cues (A-D, W-Z) are imperfect predictors. For example, A is partially correlated with both O1 and O2, and W is partially correlated with both O1 and O3. However, cues A-D are rendered relevant by the outcome choices available on each trial, while W-Z are rendered irrelevant. For example, on AW trials, only O1 and O3 are available as choices. Since W is an imperfect predictor of O1 and O3, W is not relevant or useful for making the correct choice. A, on the other hand, is an imperfect predictor of O1

and O2, and thus enables participants to make the correct choice (O1), rending it relevant on AW trials (see Figure 1).

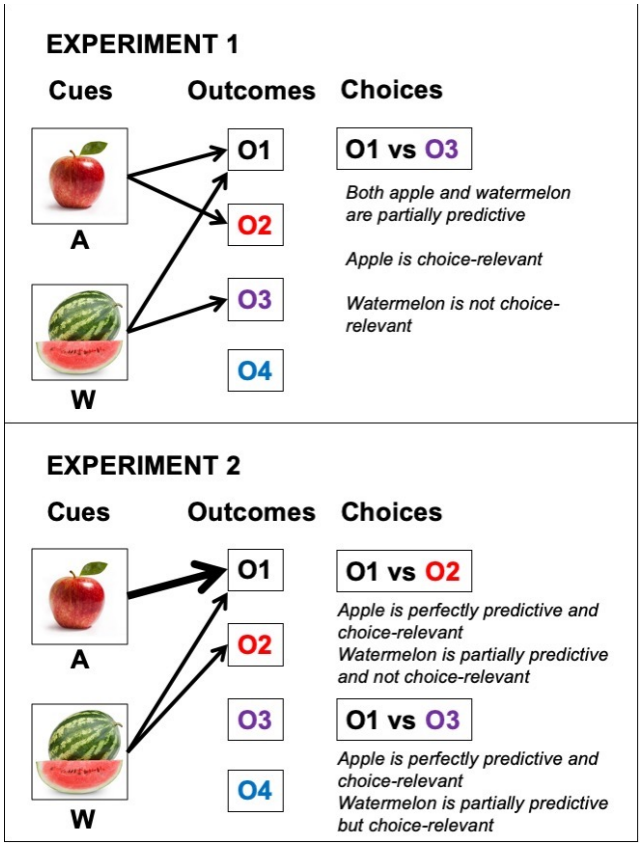

Figure 1: Schematic diagram of how cue-outcome correlation and choice relevance were manipulated in each experiment. The correct outcome for AW is O1 in both cases.

Method

Participants Experiments 1 and 2 were both run using Amazon's Mechanical Turk online platform. Participants were compensated USD5 for their participation. The task took around 30 minutes to complete. 105 participants completed Experiment 1 (32 female, mean age = 34.9 years).
Procedure The design of Experiment 1 is shown in Table 2. The experiment consisted of two training stages, and two test phases. Each participant received 12 trials of each of the 8 trial types in stage 1, and 8 trials of each of the 4 trial types in stage 2. Trials were randomized in blocks of 16 for stage 1 and blocks of 8 for stage 2 (2 of each trial type per block, with the spatial layout of the cues counterbalanced within blocks).

At the beginning of stage 1 training, participants were told that they would be playing the role of an allergist and their job was to determine what foods were causing different allergic reactions in a fictitious patient, "Mr. X". On each trial, participants were presented with a combination of two foods, and asked to make a prediction about the outcome by pressing keys corresponding to the possible outcomes on each trial. Prior to stage 1, participants were told that they would have to choose between two of the four possible outcomes on each trial. After participants made their choice, feedback about the correct outcome was presented for 2 seconds along with the food cues. After a 1s inter-trial-interval (ITI), the next trial began.

Critically, participants were told prior to the start of stage 2 training that they would now be diagnosing a new patient, "Ms. Y" and that this patient is suffering from different allergic reactions. Stage 2 training proceeded in a similar way to stage 1 training, with participants making choices between just two outcomes (O5 and O6).

After stage 2 training, participants were first tested on their stage 2 learning (with O5 and O6), and then on their stage 1 learning (with O1-O4) in a separate test phase. They were told that they would no longer receive feedback as to what the correct outcome was. Participants were presented with each of the 8 individual cues (A, B, C, D, W, X, Y, Z), and asked to rate the likelihood that the patient would suffer from each of the allergic reactions. Participants gave a rating for each outcome (i.e. O5 and O6 for the critical stage 2 test) on a visual analogue scale ranging from "Extremely unlikely" to "Extremely likely" with equally spaced intermediate labels "Somewhat unlikely" and "Somewhat likely". Each cue was presented once.

Finally, participants were asked to report how many patients they saw across the two training stages and whether they wrote anything down during the task.

Results and Discussion

We used a threshold of 60% accuracy in the second half of stage 1 to identify those who failed to learn (the choice of 60% is arbitrary but consistent with past studies from our lab on attention changes in causal learning, e.g. Don et al., 2019; Livesey et al., 2019; Shone et al., 2015). We excluded participants who did not meet this criterion, as well as participants who failed the writing check, since we explicitly told participants not to write anything down. After exclusions there were 78 participants remaining. Note that the pattern of critical test results in both experiments was the same when including all participants. We conducted our frequentist analyses using the afex package (Singmann, Bolker, Westfall, Aust, & Ben-Shachar, 2019) and Bayesian analyses using the BayesFactor package (Morey & Rouder, 2018) in R (R Core Team, 2019).

Figure 2 shows the data from the two training phases. It is clear that participants acquired the contingencies over both training phases. Table 3 reports mean predictive ratings for the correct and incorrect outcomes, for the choice-relevant cues (A-D) and the choice-irrelevant cues (W-Z), when tested first for stage 2 learning then subsequently stage 1 learning. The correct ratings for stage 1 learning were calculated by taking the mean rating for any outcomes paired with a particular cue (e.g., for A, the average of O1 and O2), and the incorrect ratings obtained by taking the mean rating for outcomes not paired with a particular cue (e.g., for A, O3 and O4). Note that critical learning scores are calculated from the stage 2 predictive ratings only (stage 1 ratings are presented

for completeness only). We calculated a learning score for each participant for each cue using the difference in ratings for the correct and incorrect outcomes in stage 2 (O5 and O6, positive scores indicate better learning). Figure 3 shows the distribution of the mean learning scores for the relevant (A-D) and irrelevant (W-Z) cues (a), and their difference (b).

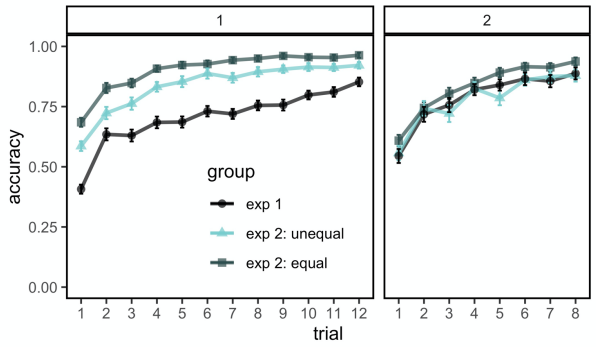

Figure 2: Mean accuracy over both stages of training in Experiments 1 and 2. Error bars represent standard error of the mean.

Table 3: Experiment 1 predictive ratings.

	Relevant (A-D)		Irrelevant (W-Z)	
	Correct	Incorrect	Correct	Incorrect
Stage 2	79.3	25.0	75.6	29.3
	(14.8)	(18.5)	(17.3)	(18.9)
Stage 1	65.0	33.0	61.6	35.8
	(14.4)	(19.3)	(13.3)	(16.9)

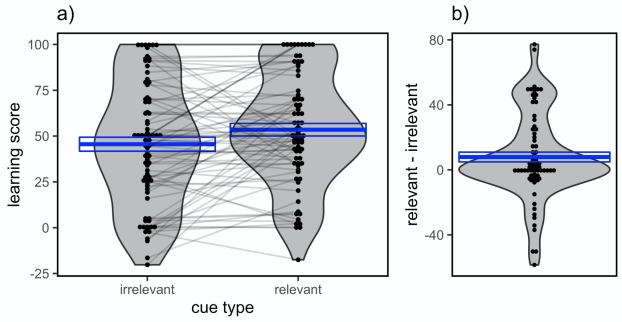

Figure 3: a) Violin plots showing the distribution of learning scores in Experiment 1 for each cue type with data points for each subject joined by lines, and b) violin plot showing the distribution of learning score differences between relevant and irrelevant cues. Blue boxes indicate +/- 1 standard error of the mean, solid blue line indicates the mean.

Using a one-way analysis of variance (ANOVA), we found a significant difference between learning scores for the relevant and irrelevant cues, $F(1,77) = 6.97$, $p = .010$, $\eta 2 = .083$. Consistent with this result, a Bayesian t-test produced a Bayes Factor (BF) in favor of the alternative, $BF_{10} = 3.15$. This means that the alternative hypothesis is 3 times as likely as the null hypothesis for the effect of relevance. Thus, despite all cues being imperfect predictors in stage 1 and perfect predictors in stage 2, learning in stage 2 was biased towards cues that were previously more useful in selecting the correct outcome on each trial. The results of Experiment 1 show that choice relevance alone can produce changes in associability, suggesting that it may contribute to, or be wholly responsible for, the learned predictiveness effect.

Experiment 2

The aim of Experiment 2 was to test whether choice relevance has an effect in a more typical learned predictiveness design where cues differ in their objective relationship with the outcomes in stage 1. The design of Experiment 2 is shown in Table 4. Unlike Experiment 1, in Experiment 2, one cue of each compound was a perfect predictor (cues A-D) and the other was an imperfect predictor of the outcome (cues W-Z).

We manipulated relevance in stage 1 training by allocating participants to one of two groups that differed only in the choice relevance of the imperfect predictor. The perfect predictor in the learned predictiveness design is always choice-relevant (for instance cue A informs the learner that they should choose O1, see Table 4 and Figure 1). However, here we manipulated whether each choice in stage 1 could be made on the basis of the imperfect predictor alone, and thus whether or not the imperfect predictor possessed choice-relevance that was equivalent to the perfect predictor.

In Group Unequal, participants completing stage 1 were given choices between the two outcomes that were associated with the imperfect predictor. For instance, W is paired with O1 and O2, which are both possible outcome choices on AW and BW trials (see Table 4 and Figure 1). Therefore, the imperfect predictor is not useful for making the correct choice (W possesses less choice relevance than A or B).

In contrast, in Group Equal, participants completing stage 1 were given choices between an outcome associated with both cues (the correct choice) and a foil outcome that was never paired with either cue. In this case, the choice can be made perfectly on the basis of either cue, and thus the relevance of the predictive and non-predictive cues is (in this sense) equal. For example, examine AW in Table 4. A is a perfect predictor of O1, while W is an imperfect predictor of O1 and O2. However, since O2 does not appear in the response choices for AW trials (only O1 and O3 appear), W can be seen as just as relevant as A in making the correct choice (O1, see Figure 1). If choice relevance plays a role, then we would expect greater biases in learning towards the perfectly predictive cues in Group Unequal.

Method

Participants A total of 199 Mechanical Turk participants (71 female, mean age = 34.6 years) completed Experiment 2, randomly allocated to Groups Equal ($n = 94$) and Unequal ($n = 105$).

Table 4: Design of Experiment 2.

Stage 1			Stage 2		Test
Trials	Choices: Group Unequal	Choices: Group Equal	Trials	Choices	
AW - 1	1 v 2	1 v 3	AY - 5	5 v 6	All cues
AX - 1	1 v 2	1 v 3	BZ - 6	5 v 6	individually tested
BW - 2	1 v 2	2 v 4	CW - 6	5 v 6	for stage 2 learning
BX - 2	1 v 2	2 v 4	DX - 5	5 v 6	then for stage 1
CY - 3	3 v 4	1 v 3			learning
CZ - 3	3 v 4	1 v 3			
DY - 4	3 v 4	2 v 4			
DZ - 4	3 v 4	2 v 4			

Procedure The procedure was identical to Experiment 1 except for the changes in design shown in Table 4. In particular, in this design, stage 1 compounds always contained one perfect predictor and one imperfect predictor. For group Equal, the choices presented in stage 1 were always between an outcome paired with both the perfect and imperfect predictor and an incorrect outcome choice that was never paired with either cue. For group Unequal, the choices presented in stage 1 were always between an outcome paired with both the perfect and imperfect predictor and an incorrect outcome choice associated with the imperfect predictor.

Results and Discussion

After applying the same exclusion criteria as Experiment 1, there were 70 participants remaining in Group 1 and 92 participants in Group 2.

Figure 2 shows the results from the training phase in Experiment 2. As in Experiment 1, participants learned the contingencies. There was a significant group difference in overall accuracy for both stage 1 training, $F(1,160) = 10.9$, $p = .001$, $\eta^2 = .064$, $BF_{10} = 23.9$, and stage 2 training, $F(1,160) = 6.86$, $p = .010$, $\eta^2 = .041$, $BF_{10} = 3.90$. Participants in Group Equal performed better, presumably because both cues on each trial were informative about the correct outcome, compared to Group Unequal who could only use the perfect predictor.

We analyzed the data in the same way as Experiment 1, calculating a learning score for each cue by taking the difference between ratings for the correct and incorrect cues for stage 2 training (O5 and O6). Table 5 reports mean ratings for the perfect and imperfect predictors, for the correct and incorrect outcomes, and Figure 5 shows the distribution of mean learning scores (a) and difference between perfect and imperfect predictors (b). From Figure 5, it appears that learning is better for the perfect predictors than the imperfect predictors, but the difference is similar across groups.

Table 5: Experiment 2 predictive ratings.

	Perfect predictors (A-D)		Imperfect predictors (W-Z)	
	Correct	Incorrect	Correct	Incorrect
Stage 2				
Unequal	74.9 (21.5)	26.8 (25.5)	62.0 (25.9)	28.8 (22.2)
Equal	79.8 (20.5)	17.4 (20.7)	71.4 (24.7)	21.1 (21.0)
Stage 1				
Unequal	80.6 (19.4)	24.2 (20.3)	41.4 (22.9)	26.6 (21.1)
Equal	85.0 (18.1)	18.5 (21.1)	38.4 (27.9)	23.1 (21.6)

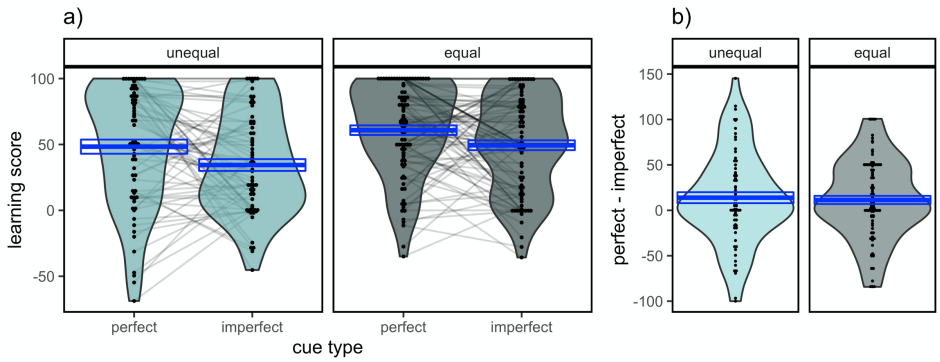

Figure 5: a) Violin plots showing the distribution of learning scores in Experiment 2 for each cue type with data points for each subject joined by lines, and b) violin plots showing the distribution of learning score differences between perfect and imperfect predictors. Blue boxes indicate +/- 1 standard error of the mean, solid blue line indicates the mean.

We analyzed the learning scores in a 2-way ANOVA with group as a between-groups factor, and predictiveness as a within-groups factor. The results were consistent between the frequentist and Bayesian ANOVAs. There was a significant effect of group, $F(1,160) = 8.3$, $p = .004$, $\eta_p^2 = .049$, $BF_{10} = 15.3$, a significant effect of predictiveness, $F(1,160) = 11.8$, $p < .001$, $\eta_p^2 = .049$, $BF_{10} = 7.08$, but most importantly, no significant interaction, $F < 1$, $BF_{10} = 0.17$. The Bayes Factor indicates that the null hypothesis is more than 5 times more likely than the alternative for the interaction effect. Participants in both groups showed a bias in learning for the previously predictive cues, replicating the typical learned predictiveness effect. Interestingly, learning scores were higher overall in Group Equal. This may indicate a motivational or metacognitive consequence of stage 1 training being easier for this group, even though stage 2 was identical in objective difficulty for the two groups.

Critically, the difference in learning scores between perfect predictors and imperfect predictors did not differ between groups, suggesting that the effects of differential cue-outcome correlation, and differential choice relevance are not additive. One explanation for these results is that the strength of the objective relationship between cues and outcomes comes to control the competitive allocation of attention when differences in this factor are present, whereas in its absence, subordinate factors such as choice relevance, have a stronger influence. Thus, when cues differ in their correlation with relevant outcomes, or when a single perfect predictor exists, manipulating the choice relevance of the imperfect predictor does not affect the relative associability of the two cues.

General Discussion

In two experiments, we tested whether the classic learned predictiveness effect (Le Pelley & McLaren, 2003) could be attributed to the relative strength of the correlation between a cue and outcome, or the relevance of a cue in making the correct choice on each trial. Using a modified learned predictiveness design in which only a subset of the possible outcomes was presented on each trial, Experiment 1 showed a *choice relevance* effect between sets of cues equated for their objective relationship with the outcome. Participants learned more about imperfect predictors that were informative about the correct outcome choice on each trial, compared to imperfect predictors that were uninformative. This result suggests that relative differences in choice relevance are sufficient to produce learned changes in attention, similar to the learned predictiveness effect.

Experiment 2 manipulated cue-outcome correlation within-subjects and choice relevance between-subjects. Group Unequal received training with compounds of cues where the perfect predictor (a cue that is perfectly correlated with a single outcome) was choice-relevant, while the imperfect predictor (a cue that was partially correlated with two outcomes) was not choice-relevant (see Figure 1). The other group, Group Equal, was similar except that the imperfect predictor *was* choice-relevant (see Figure 1). Both groups showed a learned predictiveness effect (more learning for the perfect than the imperfect predictors) but there was no difference between groups. This result suggests that differential choice-relevance (in the Unequal group) did not have any effect over and above differential cue-outcome correlation (present in *both* groups).

Taken together, the results from both experiments suggest that learned changes in associability may be driven by either choice relevance or by cue-outcome correlation. Although Experiment 1 showed a clear effect of choice relevance, it is unclear at this stage whether the typical learned predictiveness effect can be explained entirely by choice relevance. Experiment 2 showed that manipulating choice relevance did not affect the magnitude of the learned predictiveness effect, suggesting that participants' attention is more strongly biased by differences in objective cue-outcome relationships than by differences in the relevance of each cue to their current choice.

Our results are consistent with other studies showing that choice-relevant stimuli maintain attention in tasks where they are not uniquely or differentially correlated with the outcome, such as in the case of a biconditional discrimination (Kruschke, 1996; Uengoer & Lachnit, 2012). However, our demonstration is important because relevance effects in non-linear discriminations have been difficult to replicate using this discrete causal cue paradigm (see Livesey et al., 2011, 2019 for the *opposite* result).

The demonstration of a choice relevance effect in this study falls outside the scope of simple associative learning and categorization models that quantify attention changes simply on the basis of prediction error (Kruschke, 1996; Mackintosh, 1975) or on the basis of the strength of associations (Le Pelley et al., 2016). Additional assumptions would be necessary in order to accommodate choice relevance effects. For instance, cues might form inhibitory associations with incorrect choice options that are always presented with correct choice options (e.g., O3 is always incorrect when presented with A in Table 2), and attention may be sensitive to these inhibitory associations. Alternatively, associative models could be modified so that updating of associative strength differs depending on whether an outcome is presented as a choice on a particular trial. It is plausible that incorrect outcomes need to be present as a choice on a trial in order for updating to occur based on negative prediction error, while absent outcomes receive less, or even no updating. Accounting for these effects will be an important challenge for future model development in this area.

In conclusion, the current study demonstrates that we selectively attend to stimuli for different reasons. We have shown that participants attend more, and learn more, about stimuli that are more relevant to making the current choice, but that increasing choice relevance does not have an effect on attention when there is a strong competing predictor. In terms of the objective qualities of the learning task, differences in *either* cue-outcome correlation or choice-relevance are sufficient to produce biases resembling the classic learned predictiveness effect, while neither appear to be necessary.

Acknowledgments

This research was funded by two Discovery Project grants awarded by the Australian Research Council (DP190100410 awarded to EJL, and DP190103738 provided salary support for JCL).

References

Don, H. J. & Livesey, E. J. (2015). Resistance to instructed reversal of the learned predictiveness effect. *Quarterly Journal of Experimental Psychology, 68*, 1327-1347.

Don, H. J., Beesley, T., & Livesey, E. J. (2019). Learned predictiveness models predict opposite attention biases in the inverse base-rate effect. *Journal of Experimental Psychology: Animal Learning and Cognition, 45*, 143-162.

Kruschke, J. K. (1996). Dimensional relevance shifts in category learning. *Connection Science, 8*, 225–248.

Kruschke, J. K. (1996). Base rates in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 3–26.

Kruschke, J. K. (2001). Toward a unified model of attention in associative learning. *Journal of Mathematical Psychology, 45*, 812–863.

Le Pelley, M. E., Beesley, T., & Griffiths, O. (2011). Overt attention and predictiveness in human contingency learning. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 220–229.

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

Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N., & Wills, A. J. (2016). Attention and associative learning in humans: An integrative review. *Psychological Bulletin, 142*, 1111-1140.

Livesey, E. J., Thorwart, A., De Fina, N. L., & Harris, J. A. (2011). Comparing learned predictiveness effects within and across compound discriminations. *Journal of Experimental Psychology: Animal Behavior Processes, 37*, 446–465.

Livesey, E. J., Don, H. J., Uengoer, M., & Thorwart, A. (2019). Transfer of associability and relational structure in human associative learning. *Journal of Experimental Psychology: Animal Learning and Cognition, 45*, 125-142.

Lochmann, T. & Wills, A.J. (2003) Predictive history in an allergy prediction task. *Proceedings of European Cognitive Science Society, 3*, 217-222.

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

Morey, R. D. & Rouder, J. N. (2018). BayesFactor: Computation of Bayes Factors for Common Designs. R package version 0.9.12-4.2. https://CRAN.R-project.org/package=BayesFactor

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Shone, L. T., Harris, I. M., & Livesey, E. J. (2015). Automaticity and cognitive control in the learned predictiveness effect. *Journal of Experimental Psychology: Animal Learning and Cognition, 41*, 18–31.

Singmann, H., Bolker, B., Westfall, J., Aust F., & Ben-Shachar, M. S. (2019). afex: Analysis of Factorial Experiments. R package version 0.25-1. https://CRAN.R-project.org/package=afex

Uengoer, M., & Lachnit, H. (2012). Modulation of attention in discrimination learning: The roles of stimulus relevance and stimulus-outcome correlation. *Learning & Behavior, 40*, 117–127.