Variations in the Accuracy of Old-New Judgements Following Compound Learned-Predictiveness Training.

Sandra Lagator[1], Clara Muñiz-Diez[2], Tom Beesley[2], Mark Haselgrove[1]

[1]School of Psychology, The University of Nottingham, Nottingham, UK

[2]Department of Psychology, Lancaster University, Lancaster, UK

Author Note.

Please address correspondence to: Sandra Lagator, School of Psychology, University of Nottingham, University Park, Nottingham, NG7 2RD, UK. Email: sandra.lag3@gmail.com.

This research was supported by grant ES/W013215/1 from the Economic and Social Research Council of the UK.

Abstract

Connectionist theories of psychology propose that the accuracy of stimulus representations is a function of the extent to which its elements become interconnected, or associated, as a consequence of experience. To explore this idea, across three experiments, participants were required to solve a compound learned-predictiveness task, in which pairs of stimuli were established as either predictive or nonpredictive of a subsequent outcome. Outcome probability was further manipulated as either probabilistic or deterministic. It was predicted that the propensity for one stimulus to connect with another within the compound would be greater when (a) it possessed higher relative predictive validity for an outcome and (b) was followed by an uncertain outcome relative to a deterministic outcome. Consequently, participants accuracy in recognising the conjunction of the previous exposed stimulus pairs should vary as a function of these variables. Across three experiments, which used old/new judgements following compound learned-predictiveness training, a consistent effect of predictive validity was observed – participants' hit rate for detecting a change in the conjunction of predictive stimuli was superior to changes in the conjunction of nonpredictive stimuli. However, there was no effect of outcome uncertainty. The implications of these results for theories of learning and representation are discussed.

Keywords: Attention; Learning; Uncertainty; Recognition; Within-compound; Representation.

Introduction

Understanding the circumstances under which a stimulus representation will be more or less accurate is an issue that interests experimental psychologists across a variety of domains. For example, studies of "gist" have investigated which aspects of, often real-world, scenes are accurately perceived after short durations of stimulus exposure (e.g., Biederman, 1972; Fei-Fei et al., 2007; Greene & Oliva, 2009; Potter, 1976; Reger et al., 2025). At rather longer stimulus-exposure times, studies of inattentional or change blindness have explored the circumstances under which, sometimes dramatic, inclusions or changes in visual scenes fail to be detected or reported (e.g., Mack, 2003; Neisser, 1979; Rensink, et al., 1997; Simons & Chabris, 1999; Simons & Rensink, 2005). Together, these studies suggest that the accuracy of stimulus representations is often far from veridical.

Learning, however, can improve this situation. Experience has long been acknowledged to be a crucial part of accurate stimulus perception and representation. For example, for Hebb (1949), learning was necessary to accurately see a stimulus as simple as a triangle. In his discussion of the perception of patterns, Hebb noted that when a stimulus was presented briefly, then errors in its description were prominent, and that responding to the stimulus was not as if it was a whole, but instead as parts of separate entities. However, he noted a "clear effect of earlier experience" which fills in the gaps, to generate something familiar or a combination of familiar things – "a reconstruction on the basis of experience" (p.47). Hebb's solution to how this may be implemented physiologically was through association. He proposed *the cell assembly* (p. 70-74), a collection of afferent fibres that are interconnected and which correspond to basic stimulus features (e.g., edges). As the organisms scans the environment, or samples a stimulus, one assembly can become associated with another assembly, and so on, to eventually constitute a representation of the entire stimulus (see also Hall, 1991; p. 235-238). This idea was emphasised in more

psychological terms by McLaren and Mackintosh (2000) in their theoretical treatment of unitization. They proposed that one effect of exposure to a moderately complex stimulus will be to establish a more detailed and accurate representation of that stimulus. For them, like Hebb, this process depended on the formation of associations between any simultaneously-sampled elements of a stimulus. They suggested that if the changes in these associations were governed by an error-correcting rule (e.g., Rescorla & Wagner, 1972) then the initially variable and inaccurate representation of a stimulus would "settle down into a more stable, detailed and *accurate representation*" (p. 232-234). Similarly, McClelland & Rumelhart's (1985) description of their distributed model of memory noted that "The ability to retrieve *accurate completions* of similar patterns is a property of the model which depends on the use of the delta learning rule" (p. 173).

Clearly, the preceding theoretical positions emphasise the role of learning in the accuracy of stimulus representations: more accurate representations can be acquired to the extent that a stimulus's constituent elements are well connected or associated. This being the case, it follows that any manipulation which can, in principle, improve the associability of the elements of a stimulus should correspondingly improve the accuracy of its representation – its elements should be better connected, unitized, or associated. This idea was investigated recently by Lagator et al. (2025). They drew on the large body of both theoretical and empirical work which shows that stimuli which have either (1) high predictive validity or (2) are associated with outcome uncertainty come to capture relatively high levels of attention and/or associability (e.g., Beesley et al., 2015; Chao et al., 2021; Easdale et al., 2019; Esber & Haselgrove, 2018; Le Pelley, 2004; Le Pelley & McLaren, 2003; Le Pelley et al., 2016; Mackintosh, 1975; Pearce & Hall, 1980; Torrents-Rodas et al., 2021, 2025). On each trial in the training phase of Lagator et al.'s Experiment 1, participants were presented with four relatively complex images on screen (e.g., ABWX) and asked to predict which of two

outcomes (O1 or O2) would follow. After participants made their prediction, corrective feedback was given. At first, participants would have to guess which outcome followed each quartet of stimuli; however, the trial structure of the experimental design permitted participants to make more accurate predictions as training progressed (see Table 1). For Group Certain, within each quartet of stimuli, pairs of stimuli were established as either perfectly predictive or entirely non-predictive of the identity of the subsequently presented outcome. Thus, stimulus pair AB was always followed by O1, while stimulus pair CD was always followed by O2. In contrast, stimulus pairs WX and YZ were followed by O1 and O2 equally often and were thus non-predictive cues. On the basis of previous experiments and theory (e.g., Eatherington & Haselgrove, 2022; Le Pelley, Beesley & Griffiths, 2011; Le Pelley & McLaren, 2003; Lochmann & Wills, 2003; Mackintosh, 1975), Lagator et al. suggested that both the attention to and the associability of the predictive A, B, C and D stimuli would be higher than to the non-predictive W, X, Y and Z stimuli. Consequently, despite the predictive stimuli being paired together *just as frequently* as the non-predictive stimuli are paired together, the predictive pairs of stimuli (e.g., AB) should better associate with one another than the non-predictive stimuli (e.g., WX). This was precisely the result that was observed by Lagator et al. In a subsequent test phase (see Table 1) participants were presented, on each trial, with a target-stimulus (e.g., A) and two options (e.g., B and C) and were asked to select the option that accompanied the target. The results showed that participants made significantly more correct selections on test trials with the predictive stimuli than on test trials with the non-predictive stimuli. The implication of these results is that learning, beyond mere exposure, can differentially bias the accuracy of stimulus representations.

Table 1

Design of Lagator et al. (2025) Experiment 1.

Training phase				Test phase	
Group	Compound	Outcome Probability		Trials	
		O1	O2	Predictive	Non-predictive
Certain	ABWX	1	0	A: **B** vs C	W: **X** vs Y
	ABYZ	1	0	A: D vs **B**	W: Z vs **X**
	CDWX	0	1	B: **A** vs C	X: **W** vs Y
	CDYZ	0	1	B: D vs **A**	X: Z vs **W**
Uncertain	ABWX	0.8	0.2	C: **D** vs A	Y: **Z** vs W
	ABYZ	0.8	0.2	C: B vs **D**	Y: X vs **Z**
	CDWX	0.2	0.8	D: **C** vs A	Z: **Y** vs W
	CDYZ	0.2	0.8	D: B vs **C**	Z: X vs **Y**

Note. O1 and O2 refer to Outcome 1 and Outcome 2 respectively. Bold typeface indicates the correct option on each test trial.

A second group of participants, Group Uncertain, received the same exposure to the stimuli as Group Certain; however, for this group the relationship between the stimulus quartets and the outcomes was probabilistic rather than deterministic (see Table 1). For Group Certain the AB and CD pairs of stimuli were still better predictors of O1 and O2, respectively, than were the WX and YZ pairs; however, the probability of O1 following AB, for example, was now 0.8 rather than 1. On the remaining trials, the alternative outcome (O2) was presented. On the basis of previous studies, Lagator et al. suggested that the introduction of uncertainty in this group would be expected to increase the amount of attention paid to the stimuli, as a whole, in this group (e.g., Beesley et al., 2015; Easdale et al., 2019; Pearce &

Hall, 1980; Walker et al., 2019, 2022) and, also, potentially their associability (Chao et al., 2021). Thus, it was expected, that recognition memory would be higher in Group Uncertain than in Group Certain group. However, this result was not observed – there was evidence for equivalent memory performance in the final test in the two groups ($BF_{Inclusion}$ = 0.26).

Lagator et al.'s (2025) results are interesting for two reasons. First, they reveal that the degree to which simultaneously-presented stimuli become associated with one another, is influenced by the extent to which those stimuli are predictive of a subsequently-presented outcome. If we accept that a stimulus's representational accuracy can be, at least in part, defined as the extent to which its elements are correctly associated (a la. Hebb, 1949; McClelland & Rumelhart, 1985; McLaren & Mackintosh, 2001), then it provides a potential mechanism for how task-relevancy can influence the veridicality of a stimulus representation. Second, they add to a growing body of literature that points to a paradox in the relationship between learning, attention and uncertainty. Two strong predictions of the class of attentional learning theory exemplified by Pearce and Hall (1980; see also Le Pelley et al., 2012; Pearce et al., 1982; Schmajuk et al., 1996) are that stimuli which are followed by an inconsistent outcome (i.e., associated with uncertainty; Collins & Pearce, 1985) will (1) be more likely to capture attention and (2) have high associability. Studies of learning and attention have generally supported the first prediction: stimuli which are paired with inconsistent outcomes capture longer dwell times than stimuli which are paired with a consistent outcome (e.g., Beesley et al., 2015; Hogarth et al., 2008; Walker et al., 2019, 2022; see also effects of outcome uncertainty of attentional capture: e.g., Cho & Cho, 2021; Ju & Cho, 2023; Pearson et al., 2024). However, the second prediction, that uncertainty increases cue-associability, has been less well supported, at least in studies with human participants. Beesley et al., for example, found that stimuli which had received training in which a probabilistic relationship was established between stimuli and outcomes were, if anything, subsequently *less* able to

enter into novel associations - despite attracting high levels of overt attention (see also Kattner, 2015; Le Pelley et al., 2010; Livesey et al., 2011). To our knowledge, only Chao et al. (2021) have shown that stimuli paired with uncertainty are subsequently learned about *more* rapidly than stimuli that have been paired with a consistent outcome. Thus, we have the paradoxical situation in which a stimulus can command a relatively high level of attentional processing but that this high level of attention does not result in preferential learning about that stimulus in the future. Lagator et al.'s results added to this paradox: learning, in this case learning about the structure of the cue compound, was no better in Group Uncertain relative to Group Certain, despite the former group we presume, paying the stimuli more attention than the latter group.

Across three experiments, we sought to reproduce and expand the generality of the results reported by Lagator et al. (2025); in particular, by employing a more subtle test of the accuracy of the associations between the stimulus-pairs that made up the predictive and non-predictive elements of a compound. Across three experiments, we reliably observed an effect of predictive validity: test performance consistently indicated better associations between the stimuli from predictive than non-predictive pairs of stimuli. Curiously, however, the impact of uncertainty was more variable.

Experiment 1

Seminal animal-conditioning experiments by Rescorla (e.g., Rescorla, 1980; Rescorla, 1981; Rescorla & Cunningham, 1978; Rescorla & Durlach, 1981; Rescorla & Freberg, 1978) provided important demonstrations of the formation of associations between simultaneously presented stimuli (so called, within-compound associations). The typical method for detecting such learning in non-human animals is to present subjects with a stimulus compound (e.g., AB) during phase 1 of the experiment, before then changing the value of one of the stimuli in

phase 2; for example, by establishing an aversion to A. In a final test phase, responding to stimulus B is assessed, to provide a measure of the extent to which it too will elicit behaviour consistent with the changed value of A (for a review, see Rescorla & Durlach, 1981). Changes in responding evoked by B (compared to suitable control cues) provides a measure of the within-compound association that has developed between A and B. Experiments with human participants have employed either direct or indirect procedures for determining the presence of associations between simultaneously presented stimuli (Welham & Wills, 2011). Experiments using the direct procedure will, typically, first expose participants to compounds of stimuli, and then at test, present a target stimulus and ask participants to select, from two or more options, the stimulus that previously accompanied the target. As noted earlier, using this procedure, Lagator et al. (2025) detected associations between simultaneously presented stimuli, and, more importantly, revealed a bias in the strength of these associations based on the predictive validity of the stimuli (see also Melchers et al., 2004; 2006). However, the direct procedure does have some shortcomings. First, this type of test will typically instruct participants to identify the stimulus option that was paired with a target stimulus; and hence, to one degree or another, inform participants of the very psychological phenomenon under investigation – the simultaneous association. Second, by presenting the target stimulus alongside correct and incorrect choice options the conditions of stimulus presentation at test will be different to the conditions of stimulus presentation during training *on every single trial*, thus, potentially leading to an underestimate of the effect under investigation.

Employing an indirect procedure can overcome these shortcomings. Here participants are presented with a recognition test in which previously experienced and novel stimulus compounds are presented to participants, and an “old” or “new” response must be made, respectively (e.g., Larkin et al., 1998; Wasserman & Berglan, 1998). In this test, no reference is made to the presence (or absence) of a previous relationship between the stimuli that are

presented at test – instead the association is inferred by the extent to which a change in the compound disrupts reporting a "new" response. Furthermore, if appropriately designed, each trial of the test stage can be made to appear very similar to the training stage; thus, in principle, permitting better transfer of the representations formed during training to the test (see Figures 1 and 2). Experiment 1 employed the same training procedures used by Lagator et al. (2025) but now used an indirect test of within-compound associations with an "old-new" test (see Table 2). Thus, participants first received training trials with compounds of four stimuli, ABWX, ABYZ, CDWX, and CDYZ, and had to make predictions as to which of the two outcomes was correct. For Group Certain, O1 always followed the first two compounds (ABWX and ABYZ) and O2 always followed the second two compounds (CDWX and CDYZ). Thus, the AB and CD pairs of stimuli were predictive of the trial outcome, whereas the WX and YZ pairs were non-predictive. Group Uncertain received identical exposure to the compounds of four stimuli, however, O1 and O2 were probabilistically related to the compounds: trials with ABWX and ABYZ were followed by O1 on 80% of occasions and O2 on the remaining 20%, for trials with CDWX and CDYZ the reverse was true, these compounds were followed by O2 on 80% of occasions and O1 on the remaining 20%. Once the training phase was complete, participants moved to the test phase. Here participants were again presented with trials that contained compounds of four stimuli but now were asked to respond "old" or "new" depending upon whether participants had seen them in the preceding training phase. Half of the trials ("old trials") comprised presentations of the compounds previously presented during the training phase. Crucially, the remainder of the test phase comprised "new" trials, in which either the predictive or non-predictive pairs of stimuli were swapped with another stimulus of the same type to make novel compounds. Thus, the "New Predictive" trials comprised ADWX, BDWX, CBYZ and CAYZ, in which the non-predictive pairs were identical to training, but the predictive pairs of stimuli were

switched. In contrast, the "New Non-predictive" trials comprised ABWY, CDZX, ABWZ and CDYX, in which the predictive pairs were identical to training, but the non-predictive pairs of stimuli were switched.

Table 2

Design of Experiment 1

Training phase				**Test phase**		
Group	Compound	Outcome Probability		Trials		
		O1	O2	Old (x2)	New Predictive	New Non-predictive
Certain	ABWX	1	0	ABWX	**AD**WX	AB**WY**
	ABYZ	1	0			
	CDWX	0	1	CDWX	**BD**WX	CD**ZX**
	CDYZ	0	1			
Uncertain	ABWX	0.8	0.2	ABYZ	**CB**YZ	AB**WZ**
	ABYZ	0.8	0.2			
	CDWX	0.2	0.8	CDYZ	**CA**YZ	CD**YX**
	CDYZ	0.2	0.8			

Note. Bold typeface indicates the change in the compound.

Based on the results of Lagator et al. (2025) it was expected that the within-compound associations between the predictive pairs of the compound would be better established than the non-predictive pairs. Consequently, the change in the configuration of the compounds on the "New Predictive" test trials should be easier to detect than on the "New Non-predictive" test trials and, consequently, the proportion of correct responses should be higher during the

former trials than on the latter trials. From Lagator et al., we also anticipated that, contrary to theoretical predictions that can be derived from the Pearce and Hall (1980) theory, there would be no benefit provided by the inclusion of uncertainty during training – thus the proportion of correct responses should be equivalent in Group Certain and Group Uncertain.

Method

Transparency and Openness.

Stimulus presentation was controlled using PsychoPy (Pierce et al., 2019; version 2023.1.3) and all experiments were hosted on Pavlovia (https://pavlovia.org/). The analyses for Experiments 1, 2, and 3 were conducted in R version 4.5.0 (R core team, 2023); the exploratory analyses of Experiments 2 and 3 were conducted in JASP 0.17.1 (JASP Team, 2023). Experimental materials, data, and data analysis scripts are available at this link[1]. The design of the experiment and the analyses were based on Lagator et al. (2025) but were not pre-registered. The experiments reported in this paper received ethical approval by the Ethics Committee at the School of Psychology, University of Nottingham, UK.

Design

During the training phase, participants were presented with four different compounds, each of which was paired with one of two outcomes: ABWX – O1, ABYZ – O1, CDWX – O2, CDYZ – O2. Compounds AB and CD were predictive of outcomes O1 and O2 respectively, while compounds WX and YZ were non-predictive. Uncertainty was manipulated between two groups of participants: in Group Certain, compounds containing the predictive pairs AB and CD were always followed by O1 and O2 respectively, while the predictive compounds in Group Uncertain had a probabilistic association with the outcomes. In Group Uncertain, 20% of trials of each type resulted in the alternative outcome (ABWX –

[1] Full link: https://osf.io/s56wj/overview?view_only=5f2528f099f747619f1b35d6023e8e80

O2, ABYZ – O2, CDWX – O1, CDYZ – O1). At test, participants were presented with further trials that showed different cue compounds, some of which they had seen before (e.g., ABWX) and some of which involved novel combinations of the cues presented during the training phase. On the new trials, participants saw a set of four cues that involved a change in one of the pairs of stimuli, such that two cues that had not been paired before were presented together (e.g., **AD**WX where A and D formed a novel pair within the compound). This change could occur in either a predictive (e.g., **CA**YZ) or in a non-predictive pair (e.g., CD**YX**). On every trial, participants were asked to indicate whether the trial was old or new. See Table 2 for the list of cue combinations shown during the test phase.

Participants.

As Lagator et al., (2025) note, studies of differences in the associability of predictive/non predictive stimuli and certain/uncertain stimuli have heterogeneous effect sizes (varying from small to medium, to large effect sizes). Erring on the side of caution, we aimed to recruit approximately 50 participants per group. G*Power 3.1 (Faul et al., 2009) revealed that this would provide a power of >.95 to detect an interaction with an effect size of η_p^2 = .05, which would reflect a difference in within-compound knowledge about predictive and nonpredictive stimuli that varies as a function of experimental group. There were 120 participants recruited through Prolific from the UK, the US, and Australia, with English as their primary language, and within the age range of 18-30. Participants were instructed that they could complete the task using their desktop computer but not using their phone or tablet. Participants received £6 as an inconvenience allowance. Some participants made double attempts at starting the experiment and were excluded from further analyses (10 exclusions). Nine participants were further excluded because they did not finish the task. Following the exclusions, there were 101 participants in total: seven participants did not report their age and gender; among the remaining participants, there were 49 men, 44 women, and 1 non-binary

person. The age range of these participants was 19-30 years (M = 25.3, SD = 3). There were 48 participants in Group Certain and 53 in Group Uncertain.

Stimuli and Apparatus.

The images used to represent the cues were 16 different images of cluttered bedrooms (also used in Lagator et al., 2025). All the images were generated using DALL·E 1, an artificial intelligence system developed by OpenAI, which generates images based on a given text description (e.g., 'a cluttered bedroom'). Eight of these images were randomly selected to represent cues A-Z for each participant. The stimuli are available at the OSF link provided above. PsychoPy (Peirce et al., 2019) was used to present stimuli and control the experimental events.

Procedure.

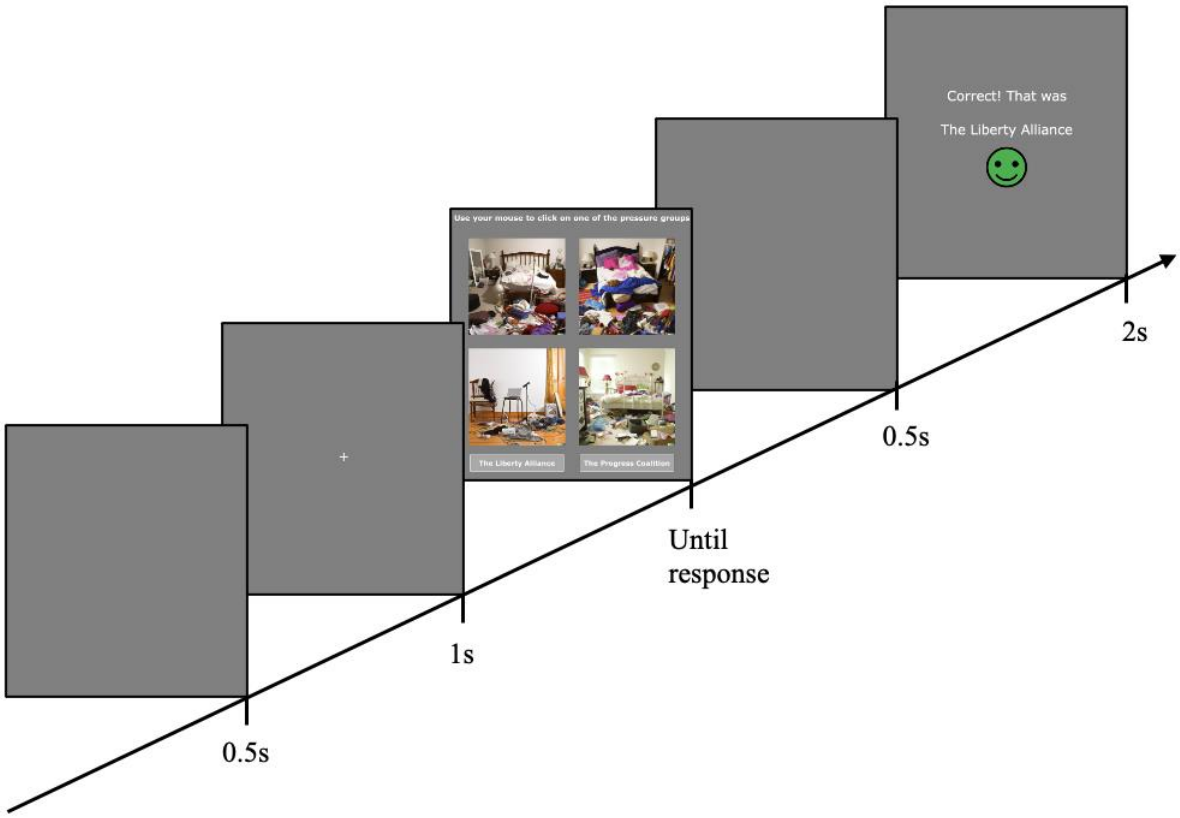
Training Phase. Participants were told that they will engage in a spy training activity. They were told that two pressure groups, 'the Liberty Alliance' and 'the Progress Coalition' (representing outcomes O1 and O2), had started collecting images of different rooms, and that each email contained four images of rooms (representing the four cues on every trial). They were told that their task was to work for "Spymaster M" and learn to classify each collection of images as belonging to one of the two pressure groups. See the OSF link above for verbatim instructions.

Each training trial consisted of five sequentially presented displays (see Figure 1 for an example training trial). The first display was a grey screen (0.5s), followed by a fixation cross presented at the centre of the screen (1s). Participants were then presented with four images of rooms, representing the four cues. Each image had a size of 0.35 x 0.35 (w x h) in PsychoPy height units, where the size of images is scaled relative to the height of the monitor

window. The centres of each image had the following positions, left to right, top to bottom: (-0.2, 0.2), (0.2, 0.2), (-0.2, -0.2), (0.2, -0.2), where the bottom left corner of the screen window had the location (-0.8, -0.5), the top right corner had the position (+0.8, +0.5). The allocation of the four images to the four locations was randomly determined on each trial. There were two buttons below the images, one for each of the pressure groups representing outcomes O1 and O2 (the Liberty Alliance & the Progress Coalition). Each button had a size of 0.34 x 0.06; the centres of the left and right outcome buttons were positioned at (-0.2, -0.44) and (0.2, -0.44) respectively. A statement at the top of the display instructed participants to use their mouse to select the correct pressure group and this display remained on the screen until participants clicked one of the buttons. A grey screen was then presented (0.5s), followed by a feedback display (2s) with the following statement: 'Correct/Incorrect! That was the Liberty Alliance/the Progress Coalition'. There was a green smiling face on correct trials and a red sad face on incorrect trials.

Figure 1

An example of a Training Trial in Experiment 1



Note. The timings represent the duration of each display.

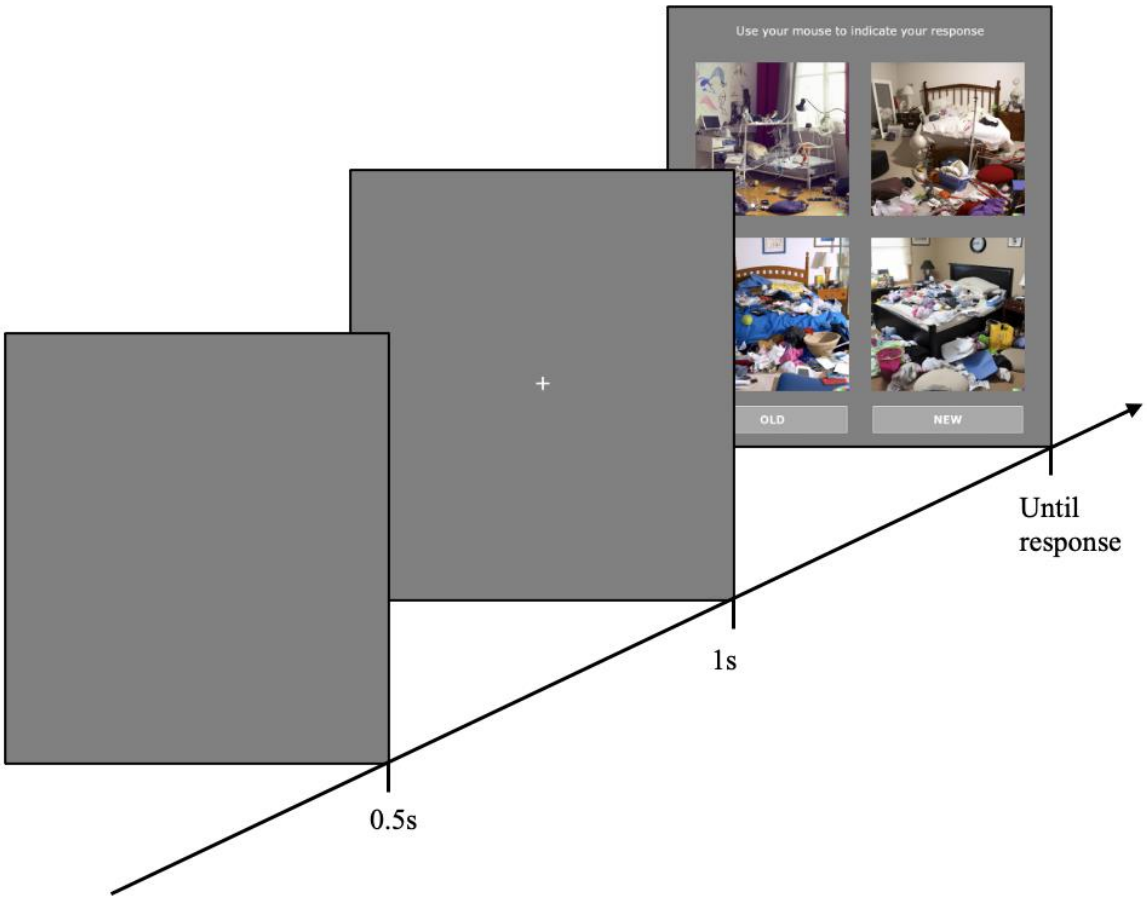
The training phase comprised 40 blocks in total and each block consisted of four different trials (one per trial type shown in Table 2), presented in a random order within a block.

Test Phase. Following the training phase, participants were told that they will be presented with further collections of images but that they will be asked to judge whether a trial was old or new (an example trial was shown). On every trial, participants were shown four images; as during the training phase, the location of these images was randomly

determined on each trial. The two buttons that represented the outcomes during the training phase were now used as response options for ‘old’ and ‘new’ (Figure 2 shows an example trial). Each trial consisted of three displays: a grey screen (0.5s), a fixation cross (1s), and the response display (until response). However, no feedback was provided during this phase, and each participant’s response initiated the next trial. There were 16 trials during this phase: each of the four old training trials was presented twice, and eight new trials. Four new test trials involved a change in a predictive compound, and four involved a change in a non-predictive compound (see Table 1); therefore, each new trial involved a unique combination of stimuli.

Figure 2

An Example of a Test Trial in Experiment 1



Note. The timings represent the duration of each display.

Results

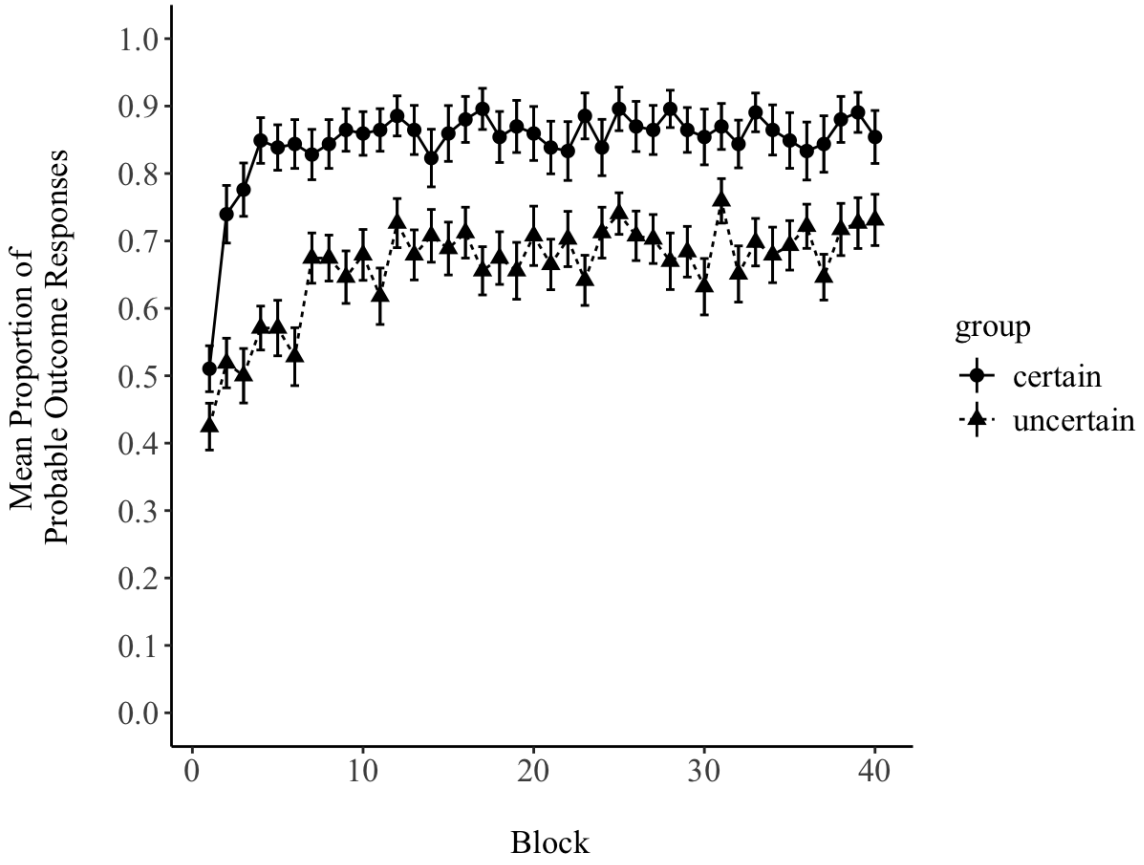
In all statistical tests, we adopt a significance level of .05. Greenhouse–Geisser corrected degrees of freedom were used where Mauchly's test indicated that the assumption of sphericity was violated.

Training accuracy. To investigate the extent to which participants learned the associations between the compounds and the outcomes, we calculated each participant's accuracy for each block during the training phase. Due to the outcome uncertainty in the

design, participants in Group Uncertain could not have made a correct response on every trial even if they had learnt the associations between the predictive compounds and the outcomes. Therefore, if they responded with the more likely outcome (e.g., O1 on an ABWX trial), this was defined as a "correct response", irrespective of the outcome presented on that trial. Figure 3 shows the mean proportion of probable outcome responses across the training blocks, separately for Group Certain and Group Uncertain. A two-way mixed analysis of variance (ANOVA) of individual proportions of probable outcome responses with the between-subjects factors of group (Certain vs Uncertain) and the within-subjects variable of block (1 to 40) revealed a main effect of group: accuracy was higher in Group Certain relative to Group Uncertain, $F(1, 99) = 26.16$, $p < .001$, $\eta_p^2 = 0.21$. A main effect of block indicated that accuracy increased across the training blocks, $F(19.25\ 1905.42) = 9.86$, $p < .001$, $\eta_p^2 = 0.09$; there was also an interaction between block and group, $F(19.25, 1905.42) = 1.67$, $p = .033$, $\eta_p^2 = .02$. Simple main effects revealed that the effect of block was significant in both Group Certain, $F(39, 1833) = 7.09$, $p < .001$, $\eta_p^2 = 0.13$, and in Group Uncertain, $F(18.68, 971.38) = 5.30$, $p < .001$, $\eta_p^2 = 0.09$.

Figure 3

Training Accuracy in Experiment 1: the Mean Proportion of Probable Outcome Responses in Group Certain and Uncertain across the Training Blocks



Note. The error bars represent ±1 standard error of the mean (SE).

Recognition test accuracy. To investigate whether participants were more accurate in detecting a change in a predictive relative to a non-predictive compound, we conducted separate analyses for the new and the old test trials. For the new trials, for each participant, we calculated the hit rate - the proportion of correct responses on the test trials separately for

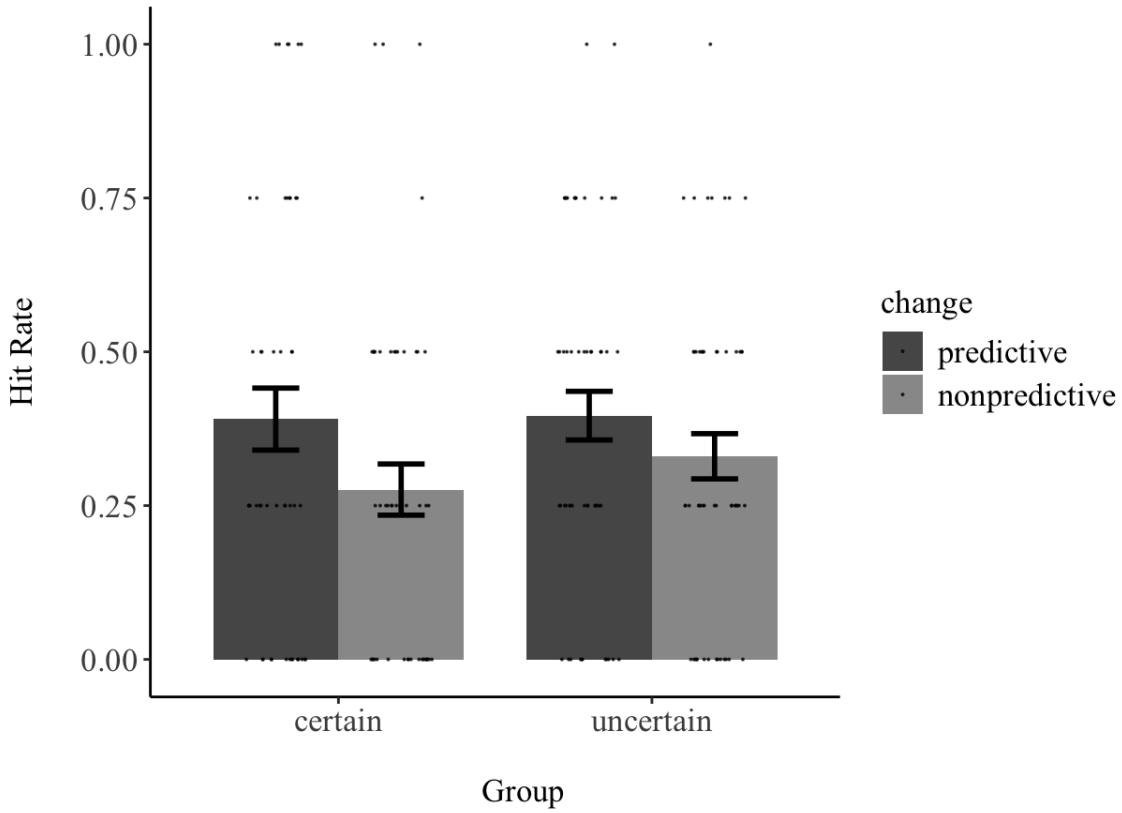
the predictive and the non-predictive change trials[2]. Figure 4 shows that, for both Group Certain and Group Uncertain, hit rates were higher when a change was made to the predictive pairs than to the non-predictive pairs of stimuli, although this effect was slightly more pronounced in Group Certain. There was no indication of any overall difference in hit rates between the groups. These impressions were confirmed with a two-way mixed ANOVA with the factors of group (Certain/Uncertain) and change (predictive/non-predictive) which showed that the overall hit rate was higher when the change involved a predictive relative to a non-predictive compound, $F(1, 99) = 7.07$, $p = .009$, $\eta_p^2 = 0.07$, but that the overall hit rate did not differ significantly between Group Certain and Group Uncertain, $F(1, 99) = 0.37$, $p = .542$, $\eta_p^2 = 0.004$. The interaction between change and group was not significant, $F(1, 99) = 0.51$, $p = .477$, $\eta_p^2 = 0.005$. A Bayesian mixed ANOVA showed consistent results: there was moderate support for the effect of change, $BF_{Inclusion} = 3.51$, moderate support for the null for the effect of group, $BF_{Inclusion} = 0.27$, and moderate support for the null for the interaction between change and group, $BF_{Inclusion} = 0.26$.[3]

[2] The same-different task in our experiment did not involve any novel stimuli. The old test trials involved cue configurations that participants had seen during the training phase, and the new test trials involved novel configurations of familiar cues. Therefore, the new test trials in our experiment were defined as the signal trials, and the proportion of correct responses on these trials was defined as the hit rate. Correspondingly, then, we defined correct rejections as the proportion of correct responses on the old test trials.

[3] We investigated the reliability of this analysis by running non-parametric alternatives to a mixed ANOVA. To investigate whether the effect of predictiveness varied with uncertainty, we calculated the difference score between the predictive and the non-predictive hit rate for each participant and investigated whether these difference scores differed between the groups. A Mann-Whitney U test showed that these scores did not vary with Group, $W = 1269.5$, $p = .989$. Due to the high number of ties in our ranks, we also ran permutation tests: the difference scores were randomly shuffled between the groups 10 000 times to investigate the proportion of tests on which this statistic would occur by chance. The proportion of cases that were more extreme than the observed value (W = 1269.5) was p = .514, which is in line with our previous analysis that indicated that the interaction between group and uncertainty was not significant.

Figure 4

The Hit Rate in Experiment 1: the Mean Proportion of Correct Responses on the New Test Trials by Change (Predictive/Non-predictive) and Group (Certain/Uncertain)



Note. Bars represent mean proportion of correct responses; points represent proportion correct responses for individual participants; error bars represent ±1 SE.

To investigate whether the differences in the hit rate between the predictive and the non-predictive changes were accompanied by a speed-accuracy trade-off, we also measured reaction times (RT) for the new test trials (RTs were measured in milliseconds and log transformed due to a positive skew in the data), and these RTs are shown in Table 3.

Table 3

The Mean Reaction Times and Standard Errors on the New Test Trials by Uncertainty and Change in Experiment 1

Uncertainty	Change	Mean	SE
Certain	Predictive	8.07	0.07
Certain	Non-predictive	8.08	0.08
Uncertain	Predictive	8.03	0.07
Uncertain	Non- predictive	7.92	0.08

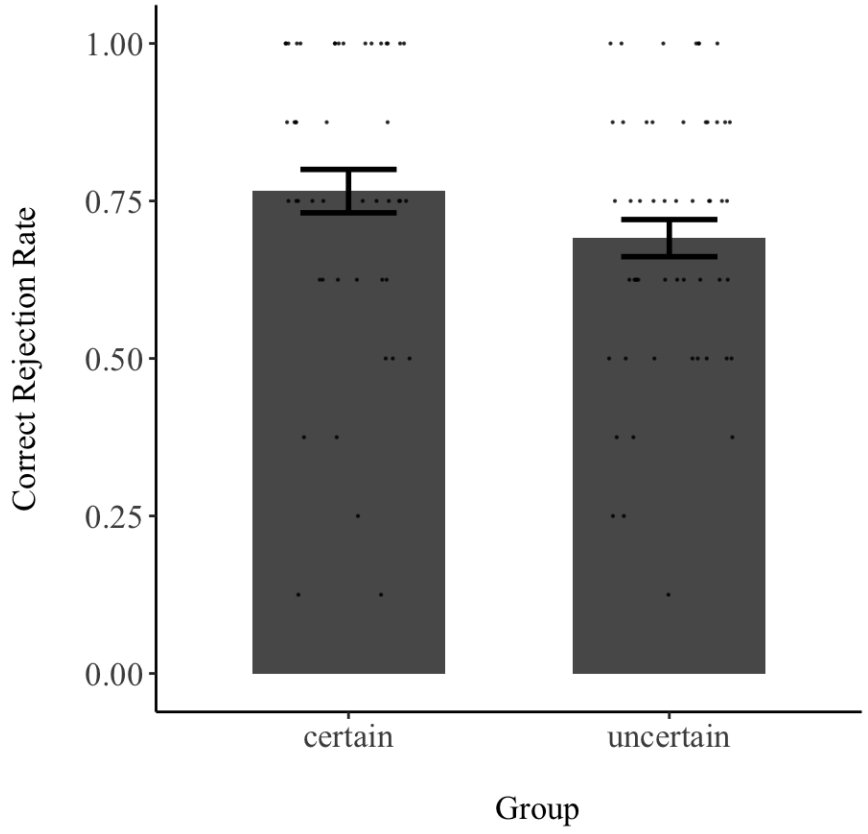
Note. The RTs were expressed in milliseconds and log transformed.

The RTs did not vary significantly with change, $F(1, 99) = 0.87$, $p = .353$, $\eta_p^2 = 0.009$, nor with group, $F(1, 99) = 1.12$, $p = .294$, $\eta_p^2 = 0.01$. The interaction between change and predictiveness was not significant, $F(1, 99) = 1.09$, $p = .299$, $\eta_p^2 = 0.01$.

We conducted a separate one-way ANOVA to investigate accuracy on the old test trials; see Figure 5 for mean accuracy across participants in Group Certain and Group Uncertain. The proportion of correct responses on the old test trials (the correct rejection rate) did not differ significantly between Group Certain and Group Uncertain, $F(1, 99) = 2.74$, $p = .101$, $\eta_p^2 = 0.03$. Similarly, the RTs did not differ significantly between Group Certain and Group Uncertain, $F(1, 99) = 0.37$, $p = .543$, $\eta_p^2 = 0.004$. The mean RTs (and SEs) were $M = 8.05$ ($SE = 0.07$) and $M = 7.99$ ($SE = 0.06$) log ms for Group Certain and Group Uncertain respectively.

Figure 5

The Correct Rejection Rate in Experiment 1: the Proportion of Correct Responses on the Old Test Trials by Uncertainty



Note. Bars represent mean proportion of correct responses; points represent proportion correct responses for individual participants; error bars represent ±1 SE.

Discussion

The purpose of Experiment 1 was to reproduce, using an indirect method of testing, the results reported by Lagator et al. (2025). In their Experiment 1, participants received identical training to the participants in the current experiment, but following this training, on each test trial, Lagator et al. presented participants with a target-stimulus beside two options and required them to select the option that accompanied the target. The results showed that

participants made more correct selections on trials with the predictive stimuli than on trials with the non-predictive stimuli; and furthermore, that there was no effect of uncertainty. Using an indirect testing procedure, Experiment 1 reproduced these effects. At test, participants were presented with compounds that were either the same or different to those presented during training and participants were required to respond "old" or "new". When the different trials comprised a change in the predictive pairs of stimuli then the hit rate was higher than when the change was made to the non-predictive pairs of stimuli. Again, there was no effect of uncertainty. It is important to note that the predictive and the non-predictive pairs of stimuli were presented together equally often, and thus any differences at test could not, therefore, be a consequence of differences in mere exposure alone.

Across both groups and both the test trials with predictive and non-predictive changes, it is notable that the hit rate on the "new" trials at test was relatively low; indeed, lower than 0.5 (chance). On the other hand, the correct-rejection rate on the old trials was relatively high (> 0.7). Together these results imply that participants had a bias towards responding "old" during the test trials. Traditionally, bias has been conceptualised in terms of a characteristic of the response (e.g., Green & Swets, 1966; Luce, 1963). However, other discussions of bias have considered it sensible to also speak of stimulus bias (e.g., Witt et al., 2015), and indeed Nosofsky (1991) discusses in substantial detail the relationship between stimulus similarity and bias. In the current experiment, the similarity of the test stimuli to the training stimuli was high; all the elements presented during the test trials were also presented during training. The feature of the "new" test trials that distinguished them from the "old" test trials was the conjunction of stimuli rather than the stimuli themselves. On this basis, then, the overall level of the hit rate observed during the test trials of the current experiment is, perhaps, not surprising.

Our expectation was that the training phase would establish variations in the representational accuracy of the elements of complex compound stimuli. The theoretical motivation for this expectation stems from combining the predictions of attentional theories of associative learning (e.g., Esber & Haselgrove, 2011; Le Pelley, 2004; Mackintosh, 1975; Pearce & Hall 1980) with connectionist theories of stimulus representation (e.g. McClelland & Rumelhart, 1985; McLaren & Mackintosh, 2001). The latter class of theory aligns an accurate stimulus representation with one whose elements are more successfully interconnected. It can therefore be hypothesised that successful interconnection of the elements of a compound would be more likely under circumstances when those elements possess high associability – such as when they have been established, during training, with good relative predictive validity (a la Mackintosh, 1975) or have been followed by an inconsistent outcome (a la Pearce & Hall, 1980). The data from the current experiment and those from Experiment 1 of Lagator et al. (2025) converge, both providing support for the first hypothesis, but not the second. The goal of Experiment 2 was to determine whether this pattern of results persists when predictive validity and outcome uncertainty were both manipulated within subjects.

Experiment 2

On each trial of the training phase of Experiment 1, compounds of four stimuli were presented to participants, and within each quartet of stimuli, two stimuli were established as predictive of the trial outcome while the other two stimuli were non-predictive. During the subsequent test phase, the stimuli were re-presented to participants with changes made, on different trials, to either the predictive or non-predictive pairs of stimuli; consequently, this comparison is within subjects. In contrast, the effect of outcome uncertainty was manipulated between-subjects - participants received trials in the training phase which were either consistently followed by the same outcome (Group Certain) or inconsistently followed by an

outcome (Group Uncertain). It is possible that the reason why an effect of predictive validity, but not outcome uncertainty, was detected in Experiment 1 (see also Lagator et al.) was because of this feature of the experimental design. Within-subjects designs certainly have statistical benefits over between-subjects design for detecting differences between conditions (e.g., increased power; Erlbacher, 1977; Greenwald, 1976). Beyond this, however, employing a within-subjects design may also impact upon the detection of an effect for more behavioural reasons – such as behavioural contrast. Behavioural contrast refers to a case where the effect of a given stimulus is impacted not solely by its own value, but also by its value relative to other stimuli presented either concurrently, or in the recent past (Lattal & Miles, 2024; Williams, 2002), for example, the rate of responding to one stimulus being impacted by varying the schedule of reinforcement in another stimulus (e.g., Thomas & Cameron, 1974). The manipulation of relative predictive validity may have benefitted from a behavioural contrast as the predictive stimulus was presented concurrently with the non-predictive stimulus, affording the opportunity for a contrast between the values of these two stimuli during training. However, the same was not true for the manipulation of outcome uncertainty, as here there was no opportunity for the stimuli that were treated probabilistically to be contrasted with another stimulus that was, for example, treated deterministically. It is possible, then, that the effect of uncertainty would be more prominent if the training trials with stimuli followed by an inconsistent outcome were experienced with trials in which the stimuli were followed by a certain outcome.

Experiment 2 replicated the general design of Experiment 1, but now with a fully within-subjects design. Thus, all participants received trials that differed in terms of outcome uncertainty and relative predictive validity (see Table 4). Consequently, participants experience of uncertainty on the trials where the stimulus compound was only intermittently

followed by one particular outcome may now benefit from a contrast with trials in which the stimulus compound was consistently followed by the same outcome.

Method

Design.

During the training phase, all the participants were presented with a set of certain trials that were used in Experiment 1 (ABWX – CDYZ; see Table 2), as well as an additional set of uncertain trials (EFPQ – GHRS). As was the case in Experiment 1, following the training phase, participants were presented with further collections of images, and they were asked to indicate whether these were old or new. See Table 4 for the design of Experiment 2, including the full list of old-new trials used during the test phase.

Table 4

The Design of Experiment 2.

	Training phase			Test phase		
Trial Type	Compound	Outcome Probability		Trials		
		O1	O2	Old (x2)	New Predictive	New Non-predictive
Certain	ABWX	1	0	ABWX	**AD**WX	AB**WY**
	ABYZ	1	0	CDWX	**BD**WX	CD**ZX**
	CDWX	0	1	ABYZ	**CB**YZ	AB**WZ**
	CDYZ	0	1	CDYZ	**CA**YZ	CD**YX**
Uncertain	EFPQ	0.8	0.2	EFPQ	**EH**PQ	EF**PS**
	EFRS	0.8	0.2	EFRS	**FH**PQ	GH**RQ**
	GHPQ	0.2	0.8	GHPQ	**GF**RS	EF**QS**
	GHRS	0.2	0.8	GHRS	**GE**RS	GH**RP**

Note. Bold typeface indicates the change in the compound.

Participants.

103 participants were recruited in the same manner as described in Experiment 1. Eighteen files were excluded due to double participation attempts made by participants, and 10 files were excluded because participants did not finish the task. Following the exclusions, there were 75 participants remaining: 34 men, 35 women, 3 non-binary individuals, and 3 people who did not report their gender and age. The age range was 18-34 (M = 24.5, SD = 4).

The experiment took 60 minutes to complete, and participants were given £9 as an inconvenience allowance.

Stimuli and Apparatus.

Experiment 2 used the same set of images as Experiment 1; these images were randomly assigned to each of the cues at the start of the experiment for each participant.

Procedure.

Each trial during the training phase proceeded as described in Experiment 1. However, each block consisted of four certain (ABWX – CDYZ) and four uncertain trials (EFPQ – GHRS). The trials within each block were randomly ordered, and there were 40 blocks (320 trials in total). Each trial during the old-new task consisted of the same displays as in Experiment 1. To account for the additional training trials, there were 16 old test trials (two per type), and 16 new test trials (each new trial combination was unique; see Table 4 for the full list of test trials).

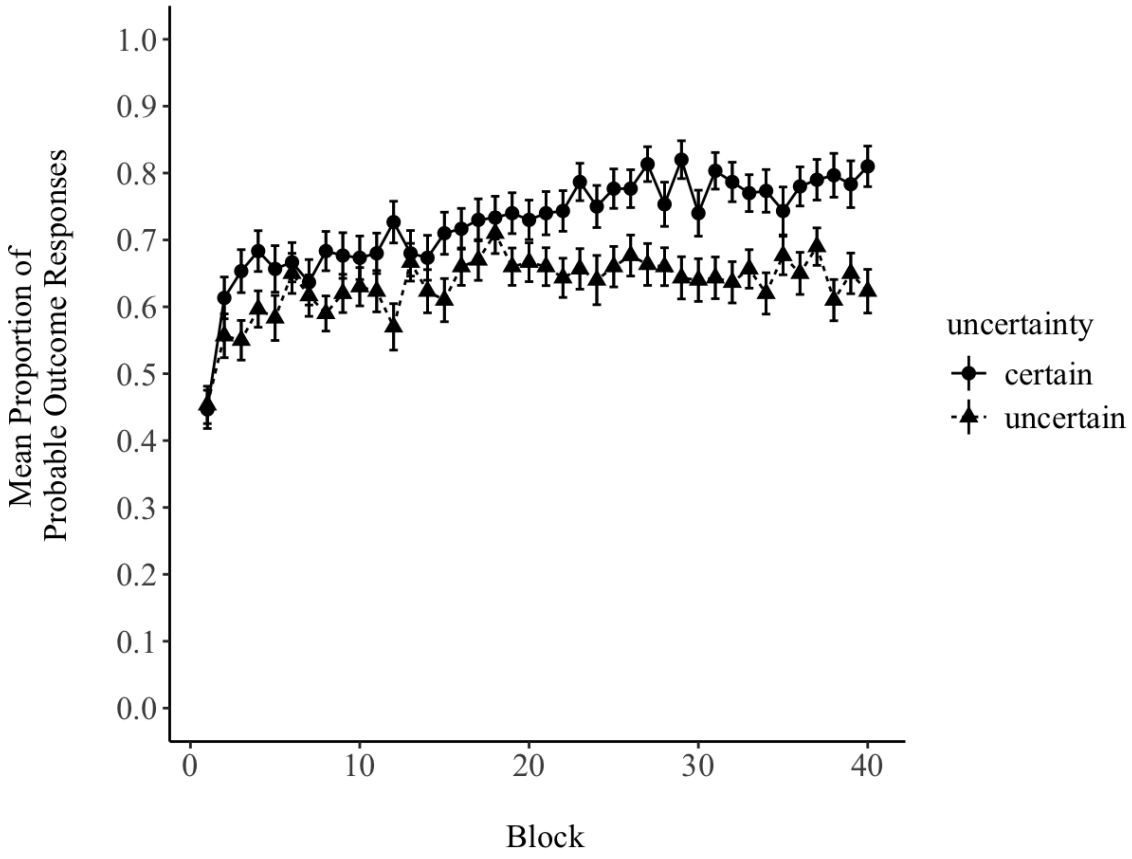
Results

Training phase accuracy. To investigate performance during the training phase, we once again calculated the proportion of probable outcome responses for each participant across the training blocks. Figure 6 shows that performance increased steadily across training, and that learning on the certain trials was more successful than on the uncertain trials. These impressions were confirmed with a two-way ANOVA with within-subjects factors of uncertainty and block, which revealed that accuracy was higher on the certain relative to the uncertain trials, $F(1, 74) = 19.01$, $p < .001$, $\eta_p^2 = 0.20$, and that accuracy increased across the training blocks, $F(18.45, 1365.37) = 9.12$, $p < .001$, $\eta_p^2 = 0.11$. The interaction between uncertainty and block was also significant, $F(22.57, 1670.41) = 2.02$, $p = .003$, $\eta_p^2 = 0.03$;

accuracy improved at a faster rate on the certain than on the uncertain trials. The effect of block was significant for both the certain, $F(19.06, 1410.56) = 8.67$, $p < .001$, $\eta_p^2 = 0.11$, and for the uncertain trials, $F(20.76, 1536.32) = 3.09$, $p < .001$, $\eta_p^2 = 0.04$.

Figure 6

Training Accuracy in Experiment 2: the Mean Proportion of Probable Outcome Responses for the Certain and the Uncertain Trials across the Training Blocks



Note. The error bars represent ±1 standard error of the mean (SE).

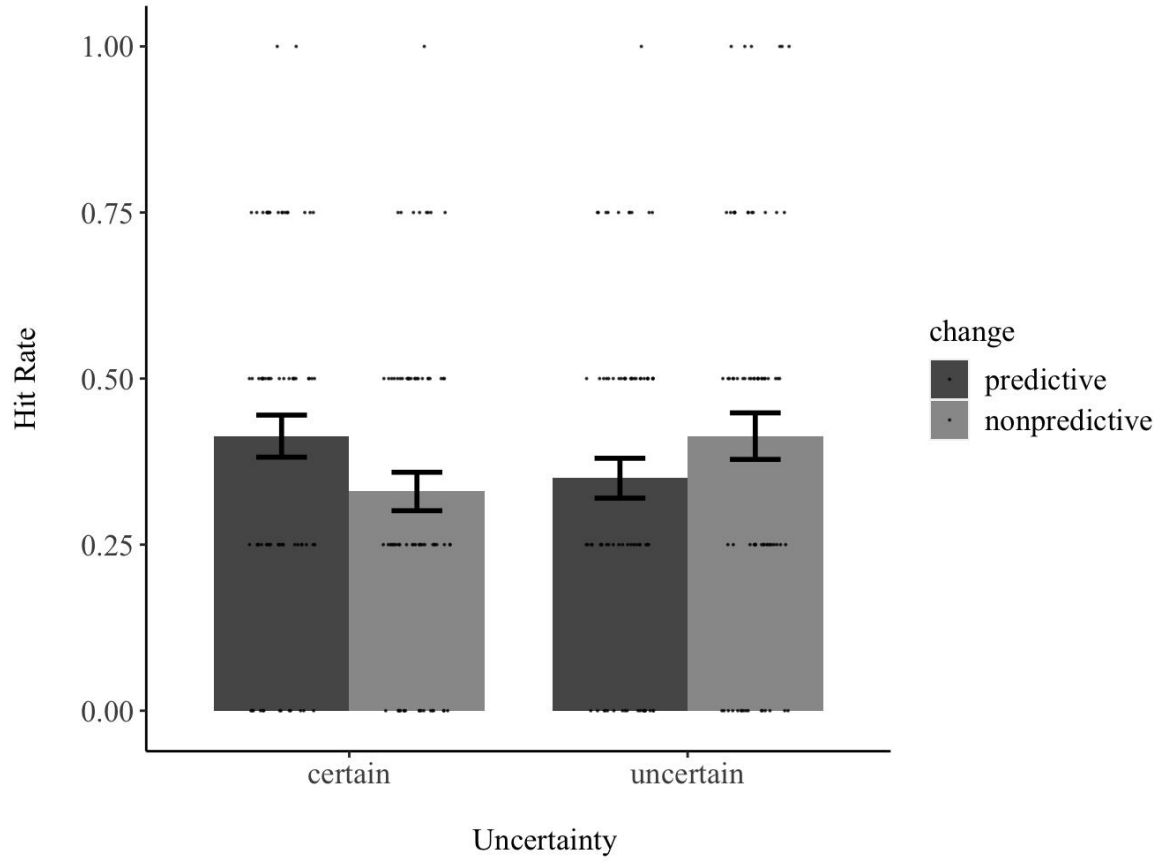
Test accuracy. As was the case in Experiment 1, test accuracy analyses were conducted separately for the new and the old test trials. For each participant, we calculated the hit rate separately for the certain and the uncertain trials, and for the predictive and the non-predictive changes. Figure 7 shows the mean hit rate by uncertainty and change. In

keeping with the results of Experiment 1, performance was higher when a change was made to the predictive than the non-predictive pairs of stimuli, but this effect was exclusive to the certain condition - if anything, this effect was reversed in the uncertain condition. There was no evidence of an overall improvement in performance on the uncertain trials relative to the certain test trials, however, there is an indication that hit rate for the non-predictive change trials was higher in the uncertain condition compared to the certain condition. These impressions were confirmed by a two-way ANOVA of hit rates with the within-subjects factors of change (predictive vs non-predictive) and uncertainty (certain vs uncertain); the hit rate did not vary significantly with uncertainty, $F(1, 74) = 0.11$, $p = .741$, $\eta_p^2 = 0.001$, $BF_{Inclusion} = 0.13$, nor with change, $F(1, 74) = 0.14$, $p = .710$, $\eta_p^2 = 0.002$, $BF_{Inclusion} = 0.13$. However, the interaction between uncertainty and change was significant, $F(1, 74) = 11.78$, $p < .001$, $\eta_p^2 = 0.14$, $BF_{Inclusion} = 6.15$. Like the pattern of recognition accuracy in Experiment 1, the hit rate was higher for the predictive relative to the non-predictive changes on the certain trials, $F(1, 74) = 5.89$, $p = .018$, $\eta_p^2 = 0.07$, $BF_{Inclusion} = 2.49$, but there was inconclusive evidence for differences on the uncertain trials, $F(1, 74) = 3.42$, $p = .068$, $\eta_p^2 = 0.04$, $BF_{Inclusion} = 0.83$. When the change happened in one of the predictive compounds, the difference between the certain and the uncertain trials was not significant, $F(1, 74) = 2.74$, $p = .102$, $\eta_p^2 = 0.04$, $BF_{Inclusion} = 0.64$. However, for the non-predictive changes, the hit rate was higher on the uncertain than on the certain trials, $F(1, 74) = 5.47$, $p = .022$, $\eta_p^2 = 0.07$, $BF_{Inclusion} = 2.02$.[4]

[4] For each participant, we calculated the difference in the hit rate between the predictive and the non-predictive trials, and separately for the certain and the uncertain trials. A Wilcoxon signed rank tested indicated that these scores varied significantly with uncertainty, $W = 1065.5$, $p = .002$. Permutation tests showed that the proportion of tests with a more extreme statistic was $p < .001$. In Group Certain, a Wilcoxon signed rank test showed a significant difference between the predictive and the non-predictive trials, $V = 1002.5$, $p = .018$; the proportion of more extreme cases in permutation tests was $p = .008$. In Group Uncertain, this difference was not significant, $V = 313.5$, $p = .074$, while the proportion of more extreme cases in permutation tests was $p = .040$. The hit rate on the predictive trials did not vary with uncertainty, $V = 918$, $p = .120$; the proportion of more extreme values in permutation tests was $p = .058$. The hit rate on the non-predictive trials varied significantly with uncertainty, $V = 394$, $p = .025$; the proportion of permutation tests with a more extreme value was $p = .013$.

Figure 7

The Hit Rate in Experiment 2: the Mean Proportion of Correct Responses on the New Test Trials by Change and Uncertainty



Note. Bars represent mean proportion of correct responses; points represent proportion correct responses for individual participants; error bars represent ±1 SE.

We also calculated the mean RT on the new test trials for each participant by uncertainty and change. The mean RTs did not vary significantly with change, $F(1, 74) = 0.10$, $p = .753$, $\eta_p^2 = 0.001$; however, the effect of uncertainty was (just) significant, $F(1, 74) = 3.99$, $p = .050$, $\eta_p^2 = 0.05$, the RTs were somewhat longer on the certain relative to the uncertain trials (see Table 5 for the mean RTs and SEs by uncertainty and change). The

interaction between uncertainty and change was not significant, $F(1, 74) = 0.63$, $p = .432$, $\eta_p^2 = 0.008$.

Table 5

The Mean Reaction Times and Standard Errors on the New Test Trials by Uncertainty and Change in Experiment 2

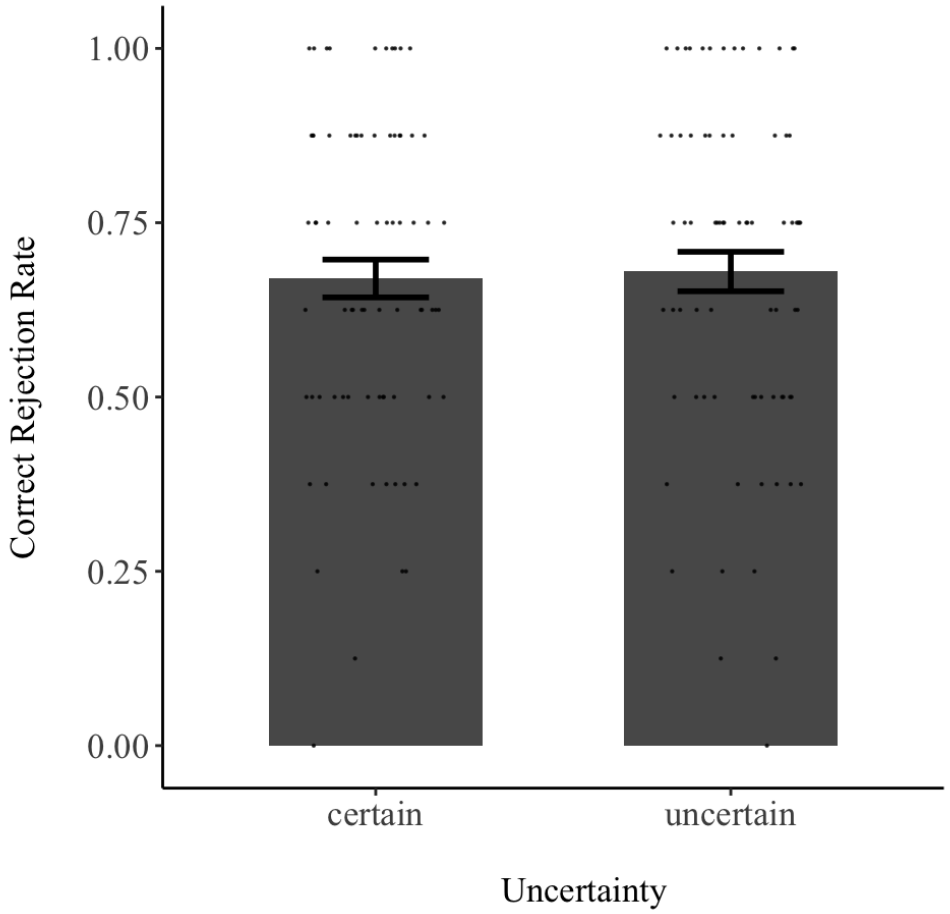
Uncertainty	Change	Mean	SE
Certain	Predictive	8.08	0.06
Certain	Non-predictive	8.03	0.05
Uncertain	Predictive	7.98	0.06
Uncertain	Non- predictive	8.00	0.08

Note. The RTs were expressed in milliseconds and log transformed.

We also calculated the correct rejection rate separately for the certain and the uncertain trials for each participant; see Figure 8 for the mean correct rejection rate across participants by uncertainty. The correct rejection rate did not vary with uncertainty, $F(1, 74) = 0.09$, $p = .767$, $\eta_p^2 = 0.001$. The mean RTs on these trials also did not vary with uncertainty, $F(1, 74) = 1.31$, $p = .257$, $\eta_p^2 = 0.02$. The mean RTs (and SEs) were $M = 8.07$ ($SE = 0.06$) and $M = 8.02$ ($SE = 0.06$) for the certain and the uncertain trials, respectively.

Figure 8

The Correct Rejection Rate in Experiment 2: the Proportion of Correct Responses on the Old Test Trials by Uncertainty



Note. Bars represent mean proportion of correct responses; points represent proportion correct responses for individual participants; error bars represent ±1 SE.

Discussion

Experiment 2 provided results that were consistent with Experiment 1. During the "new" test trials, a change made to the predictive pairs of a "certain" compound of four stimuli resulted in a significantly higher hit rate compared to the same change made to the non-predictive pairs of stimuli. Also, in-keeping with Experiment 1, there was no effect of

uncertainty on the overall hit rate for "new" test trials, or the correct reject rate for the "old" test trials. Experiment 2 did, however, reveal an unexpected interaction effect on the "new" test trials: simple effects analysis revealed that, for the test trials in which the change in stimuli was made to non-predictive stimuli, the hit rate was higher for those cues trained in uncertain relationships compared to those trained in certain relationships (Figure 7). This effect was accompanied by an overall effect of uncertainty on reaction time during the "New trials" - RTs on the uncertain test trials were, overall, (just) significantly faster than on the certain test trials.

The effect of uncertainty that was specific to the hit rate on the test trials with changes to non-predictive stimuli is unexpected and potentially theoretically interesting. One possibility is that elevated attention towards non-predictive stimuli in the uncertain condition reflects exploratory attention (e.g. Beesley et al., 2015; Dickinson, 1980; Luque et al., 2017). The reason why uncertainty might elevate attention towards non-predictive stimuli is because prediction error cannot be fully resolved by the predictive stimuli (which can be achieved in the certain condition). This effect is not predicted by theories that aim to capture the role of uncertainty in learning and attention (e.g. Pearce & Hall, 1980; Pearce, Kaye & Hall, 1982). These theories suggest that an update in associability following an inconsistent outcome would happen both for the predictive and non-predictive stimuli. These intuitions were tested and confirmed with a series of simulations of the Pearce, Kaye, and Hall (1982) theory (See: Supplementary Materials A). While potentially interesting, our results regarding this interaction have been inconsistent – while there was evidence for it in the current experiment and in Experiment 1 by Lagator et al., (2025), we did not find evidence for it in Experiment 1 of the current manuscript. In addition, the specific facilitation of hit rates to the non-predictive stimuli in the uncertain condition relative to the certain condition, while significant

with frequentist statistics, was accompanied with an inconclusive Bayes factor (2.12). Therefore, our aim in Experiment 3 was to further investigate the reliability of Experiment 2.

Experiment 3

The results of Experiment 2 were both theoretically interesting and surprising, we were therefore keen to investigate their reliability. The purpose of Experiment 3, therefore, was to conduct a replication of Experiment 2 and comprised an identical experimental design and procedure, and employed identical stimuli, participant selection procedures, exclusion criteria and statistical analyses. Based on the consistent findings of Experiment 1 and 2, we were reasonably confident that at test, the hit rate for the certain predictive "new" trials would be higher than for the certain non-predictive "new" trials. Of particular interest, however, was whether the hit rate during the uncertain non-predictive test trials would, again, be selectively higher than during the certain non-predictive test trials.

Method

Participants.

To make the sample size comparable to Experiment 2, 102 participants were recruited via Prolific in the same manner as described in Experiment 2. 16 files were excluded due to double attempts made to begin the experiment, and nine because participants did not finish the task. Two participants were excluded because they participated in a previous experiment that used a similar design. Following the exclusions, there were 75 participants remaining: 36 men, 32 women, one non-binary individual, and six people who did not report their gender and age. The age range of the remaining participants was 18-30 (M = 25.7, SD = 4). The experiment took 60 minutes to complete, and participants were given £9 as their inconvenience allowance.

Stimuli and Apparatus.

The stimuli involved the same 16 images of cluttered bedrooms that were used in Experiment 2.

Design and Procedure.

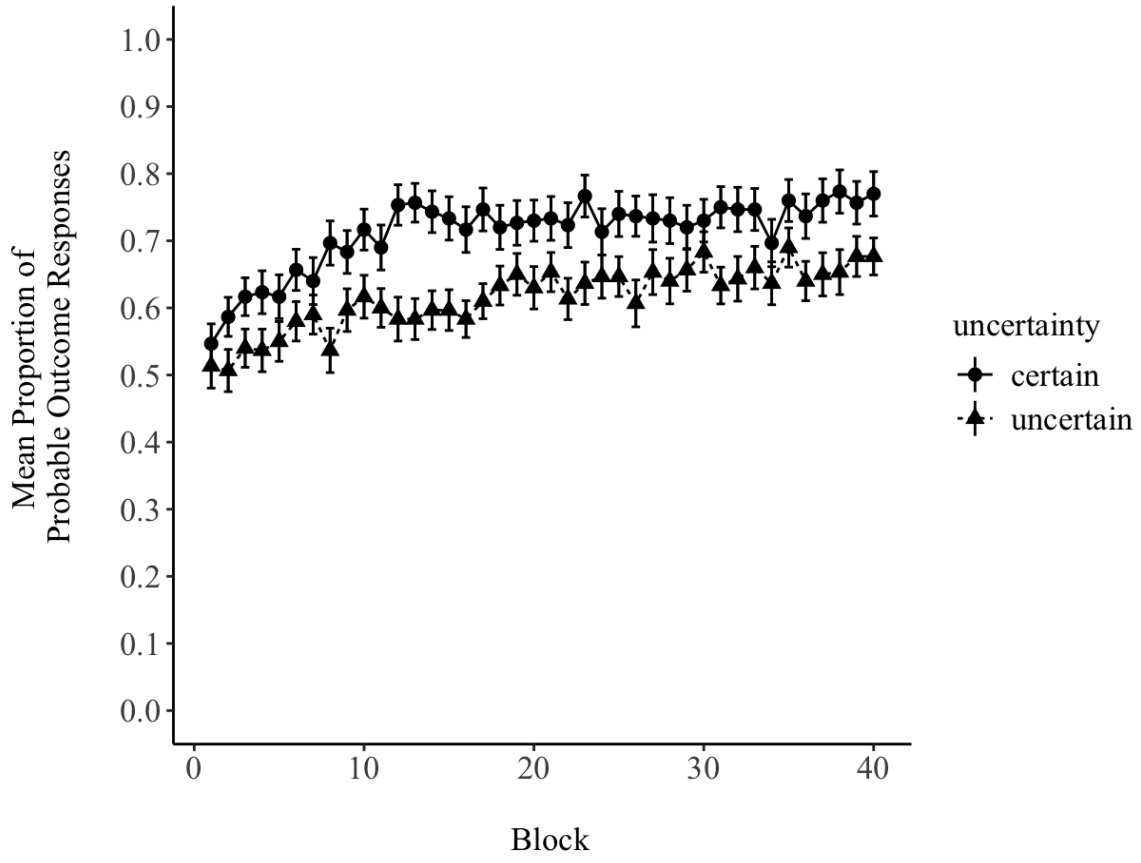
Training and Test phases. The instructions, timing of events, stimulus locations, and stimulus sizes in Experiment 3 were identical to Experiment 2.

Results

Training phase accuracy. Accuracy during the training phase was defined and calculated as described in Experiment 2. In keeping with the training data from Experiment 2, performance increased steadily across training, and learning on the certain trials was more successful than on the uncertain trials (Figure 9). A two-way ANOVA showed that accuracy increased across the training blocks, $F(20.08, 1485.65) = 6.77$, $p < .001$, $\eta_p^2 = 0.08$, and that participants performed more accurately on the certain relative to the uncertain trials, $F(1, 74) = 30.79$, $p < .001$, $\eta_p^2 = 0.29$. Unlike Experiment 2, the interaction between uncertainty and block was not significant, $F(22.64, 1675.56) = 0.95$, $p = .566$, $\eta_p^2 = 0.01$.

Figure 9

Training Accuracy in Experiment 3: the Mean Proportion of Probable Outcome Responses by Uncertainty across the Training Blocks



Note. The error bars represent ±1 standard error of the mean (SE).

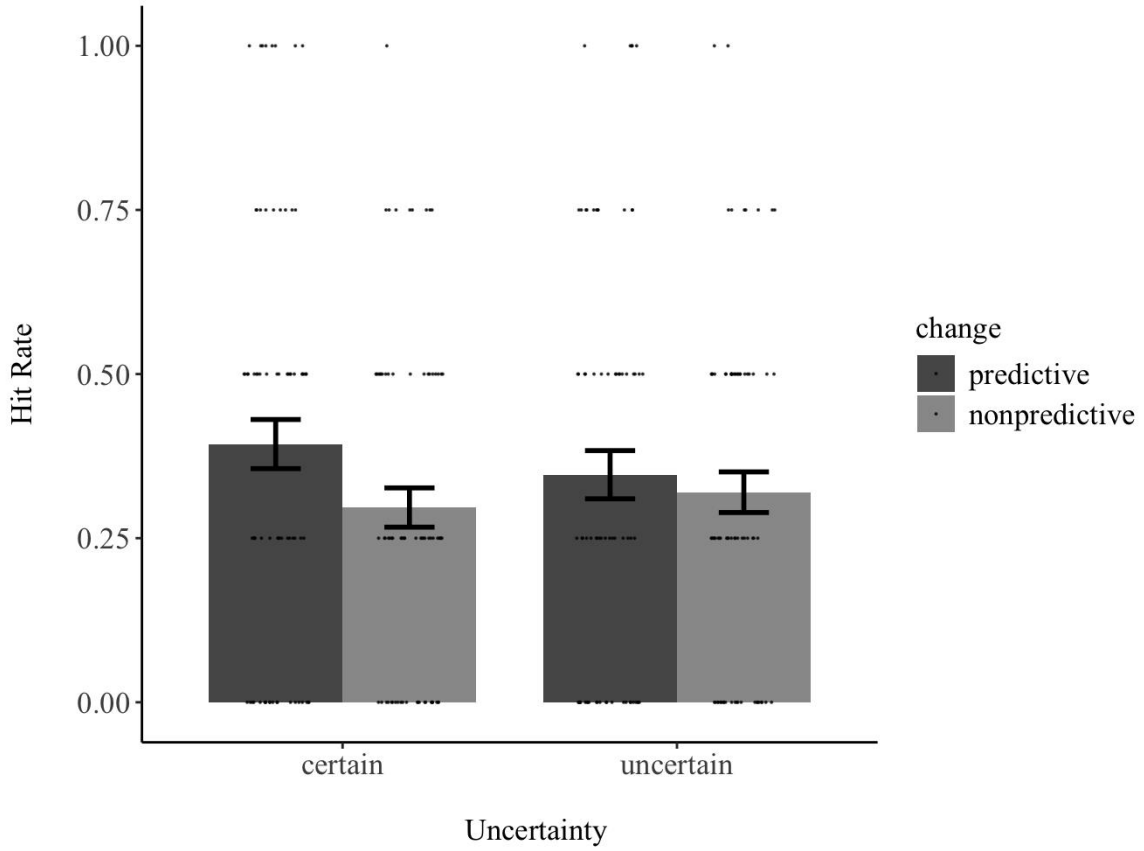
Test accuracy. As in the previous two experiments, test accuracy was analysed separately for the new and the old test trials. Figure 10 shows the mean hit rate for the "new" test trials, separated by the different levels of uncertainty and change. In keeping with the results of Experiments 1 and 2, the hit rate was higher on the predictive test trials than on non-predictive trials, in particular when training trials had a deterministic outcome (certain trials). A smaller difference between the predictive and non-predictive stimuli was evident on the uncertain test trials. Importantly, and in contrast to the results of Experiment 2, there was

no indication of the hit rate being selectively higher on the uncertain non-predictive test trials relative to the certain non-predictive test trials. These impressions were confirmed with a two-way ANOVA of hit rate with the within subjects factors of certainty (certain vs uncertain) and change (predictive vs non predictive), which revealed a main effect of change, $F(1, 74) = 5.14$, $p = .026$, $\eta_p^2 = 0.07$, $BF_{Inclusion} = 1.61$, but no effect of uncertainty, $F(1, 74) = 0.16$, $p = .691$, $\eta_p^2 = 0.002$, $BF_{Inclusion} = 0.14$. Unlike Experiment 2, the interaction between uncertainty and change was not significant, $F(1, 74) = 2.38$, $p = .127$, $\eta_p^2 = 0.03$, $BF_{Inclusion} = 0.39$.[5]

[5] As in the previous experiments, we calculated the difference in the hit rate between the predictive and the non-predictive trials, separately for the certain and the uncertain trials for each participant. A Wilcoxon signed rank test showed that these difference scores did not vary significantly with uncertainty, $V = 980$, $p = .210$, and the proportion of permutation tests that showed a more extreme test statistic was p = .104.

Figure 10

The Hit Rate in Experiment 3: the Mean Proportion of Correct Responses on the New Test Trials by Change and Uncertainty



Note. Bars represent mean proportion of correct responses; points represent proportion correct responses for individual participants; error bars represent ±1 SE.

An identical two-way ANOVA of log transformed reaction times showed that the mean RTs did not vary significantly with uncertainty, $F(1, 74) = 0.14$, $p = .713$, $\eta_p^2 = 0.002$, nor with change, $F(1, 74) = 0.04$, $p = .840$, $\eta_p^2 = 0.0005$. The interaction between uncertainty and change was not significant, $F(1, 74) = 1.39$, $p = .241$, $\eta_p^2 = 0.02$. See Table 6 for the mean RTs and SEs across participants by uncertainty and change.

Table 6

The Mean Reaction Times and Standard Errors on the New Test Trials by Uncertainty and Change in Experiment 2

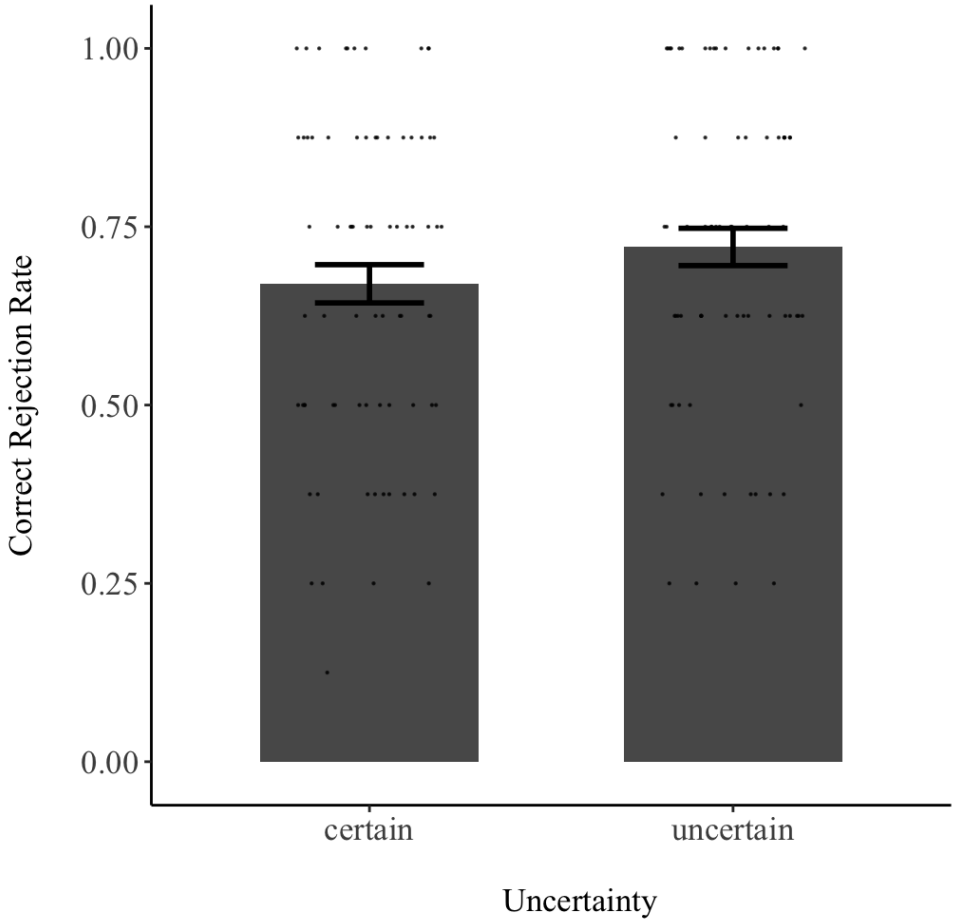
Uncertainty	Change	Mean	SE
Certain	Predictive	7.95	0.06
Certain	Nonpredictive	7.90	0.07
Uncertain	Predictive	7.93	0.07
Uncertain	Nonpredictive	7.96	0.06

Note. The RTs were expressed in milliseconds and log transformed.

Figure 11 shows the mean correct rejection rate on "old" trials for the certain and uncertain group. The correct rejection rate did not differ significantly between the certain and the uncertain trials, $F(1, 74) = 3.48$, $p = .066$, $\eta_p^2 = 0.05$. The difference in the mean RTs on the certain and the uncertain trials was not significant, $F(1, 74) = 2.78$, $p = .099$, $\eta_p^2 = 0.04$. The mean log RTs (and SEs) for the certain and the uncertain trials were $M = 7.86$ ($SE = 0.06$) and $M = 7.93$ ($SE = 0.06$) respectively.

Figure 11

The Correct Rejection Rate in Experiment 3: the Proportion of Correct Responses on the Old Test Trials by Uncertainty



Note. Bars represent mean proportion of correct responses; points represent proportion correct responses for individual participants; error bars represent ±1 SE.

Discussion

The purpose of Experiment 3 was to replicate Experiment 2 to determine the reliability of its results, and, reassuringly, a number of the results from Experiment 2 were successfully reproduced. First, perhaps not surprisingly, accuracy during the training phase

was more successful on the certain than on the uncertain trials. Second, and more importantly, during the "new" trials of the test phase, the hit rate during the predictive trials was higher than during the non-predictive test trials when training was conducted with a deterministic outcome (certain trials). This is a result which we have consistently observed across all three experiments, and which is also consistent with the results of Lagator et al. (2025) who employed a more direct test of the strength of the associations between pairs of predictive or non-predictive stimuli. Although in each specific experiment we report here, the effect of predictiveness (which is significant with frequentist statistics) is not always accompanied by conclusive Bayesian evidence, collapsing across Experiments 1 to 3 reveals this effect to be robust: the difference between the predictive and non-predictive test trials in the certain conditions revealed a small-to-medium effect size with a conclusive Bayes factor, $t(197) = 4.19$, $p < .001$, $d = 0.30$, $BF_{10} = 321.57$. The same comparison is not true, across these experiments, for the uncertain test trials, $t(202) = 0.17$, $p = .863$, $d = 0.01$, BF10 = 0.08, where the data strongly support the null. One possible reason for the absence of a difference between the hit rates on the test trials with the predictive and non-predictive stimuli is that the validity of the predictive stimuli for the outcomes was degraded during training in the uncertain conditions (p = .08) relative to the certain conditions (p =1). Consistent with this possibility, Lagator et al. (2025) showed that when participants with overall training accuracy below .6 were excluded from analyses (a criterion used by McLaren & LePelley, 2003; Le Pelley et al. 2011) then the proportion of correct responses during the predictive test trials was higher than on the nonpredictive test trials in both the certain and the uncertain groups. The explanation offered by Lagator et al. for this finding was that the apparent within-compound associations were, to some degree, mediated by the strength of the forward (and backward) associations from the cues to the outcomes. We will return to this matter in the General Discussion.

Despite Experiments 2 and 3 having identical experimental designs and similar sample sizes, the effect of uncertainty on test performance was not consistent. In Experiment 2 we observed a higher hit rate on the non-predictive test trials for the uncertain condition relative to the certain condition. However, this was not reproduced in Experiment 3. Together, Experiments 2 and 3 analysed data from 150 participants. This sample size, coupled with the fact that these experiments were fully within-subjects designs, provides us with an opportunity to investigate some of the variables that may predict variations in test performance. Therefore, we performed a set of exploratory analyses using multiple regression and principal component analysis from the data aggregated across Experiments 2 and 3 (See: Supplementary Materials B). These revealed three features of interest. First, that the superior hit rate on the certain-predictive test trials over the certain-nonpredictive test trials is positively, and specifically, related to performance during training. Second, hit rates on the uncertain test trials were higher than on the certain trials when learning on the uncertain training trials was more successful. And third, despite there being scant evidence for an effect of uncertainty upon performance at test, principal component analysis of all the types of test trial nonetheless extracted two factors that loaded onto certain and uncertain trial types.

General Discussion

Across three experiments participants were presented with trials during a training phase in which compounds of four stimuli were followed by one of two outcomes. Within each compound, pairs of stimuli were always presented together and themselves established as either predictive (e.g. AB and CD) or nonpredictive (e.g. WX and YZ) of one of two outcomes (O1 or O2). In experiment 1, outcome uncertainty was manipulated between subjects: thus, for group Certain, O1 and O2 always followed compounds containing AB and CD respectively; whereas for group Uncertain, O1 and O2 probabilistically (80%) followed

compounds that contained AB and CD, respectively. The same procedure was employed in Experiments 2 and 3, but outcome uncertainty was manipulated within-subjects for these studies. In a subsequent test phase, participants were again presented with compounds of four stimuli but now were asked to respond "old" or "new" depending upon whether they had seen them in the preceding training phase. Old trials comprised presentations of the compounds previously presented during the training phase, whereas for the crucial "new" trials, stimuli within either the predictive or non-predictive pairs were changed with another stimulus of the same type to make novel compounds. Across all three experiments, hit rates on these new test trials were consistently higher when the change was made to the predictive pairs of stimuli relative to a change made to the nonpredictive pairs. The effect of outcome uncertainty on test performance was, on the one hand, more variable. The results of Experiment 2 suggested that hit rates were higher when a change was made to the nonpredictive uncertain stimuli relative to when a change was made to the nonpredictive certain stimuli. However, this result was not reproduced in Experiment 3 and was not observed in Experiment 1. On the other hand, however, there was some consistency associated with the effect of uncertainty - across all three experiments, there was no indication that hit rates (or indeed correct rejections) were higher on test trials with uncertain relative to certain stimuli.

These experiments were motivated by theories of distributed memory and representation (e.g., McLaren & Mackintosh 2000; McClelland & Rumelhart, 1985). According to these theories, accurate representations of complex stimuli should develop as associations are acquired among the elements of those stimuli. According to attentional theories of learning, themselves governed by error correction (e.g., Esber & Haselgrove, 2011; Le Pelley, 2004; Mackintosh, 1975, Pearce & Hall, 1980), the associability of the elements of stimuli are influenced by (a) the relative predictive validity of the elements for an outcome, and/or (b) the overall outcome uncertainty of the stimulus. It follows from (a) that

the associability of the predictive elements of a compound stimulus should be higher than the nonpredictive elements of a compound stimulus – thus facilitating the interconnectivity of the predictive elements and hence their representational accuracy. All three experiments reported here supported this prediction. If follows from (b) that the associability of the elements of a stimulus that is followed by an inconsistent outcome will be higher than the elements of a stimulus that is followed by a consistent outcome – thus facilitating the interconnectivity of the stimuli associated with uncertainty. However, none of the experiments reported here were able to find evidence for this prediction. It remains to be determined precisely why relative predictive validity, but not outcome uncertainty, influences the representational accuracy of a stimulus. One possibility is that the current task is simply not suited to the deployment of exploratory attention. Experiments 1 to 3 require the classification of complex compounds of stimuli as cues for one of two outcomes, and in the uncertain conditions participants were achieving this on between 60 - 70% of trials by the end of the training phase. Obviously, then, there is still significant error on these trials, but it is conceivable that this level of performance may not be sufficiently low to motivate additional stimulus exploration, and from there an improvement in representational accuracy. It is at this point that the exploratory analyses conducted upon experiments 2 and 3 (Supplementary Materials B) may be instructive. We could detect no evidence of there being a negative relationship between hit rates to the uncertain stimuli (specifically) and performance during training. Indeed, if anything, the relationship here was positive: a non-significant increase in hit rate to the uncertain stimuli (minus hit rate to the certain stimuli) as performance during training increased. An alternative possibility is that the impact of uncertainty on stimulus associability, at least in humans, is relatively slight, a possibility to which we turn now.

The current experiments reproduce the findings of Lagator et al., (2025) who, using a more direct measure, also showed that (a) the elements of a compound stimulus were better

associated when they were predictive rather than nonpredictive of a trial outcome and (b) that the elements of a compound stimulus were no better associated when they were trained under conditions of uncertainty relative to certainty. Together, the results of Lagator et al. and the current experiments join others which have shown that while relative predictive validity can change the associability of stimuli, outcome uncertainty may not (e.g., Beesley et al, 2015; Kattner, 2015; Le Pelley et al., 2010; Livesey et al., 2011, but see Chao et al, 2021). To explain the differential effect of predictive validity and outcome uncertainty on stimulus associability, Beesley et al. suggested that exploratory attention, but not exploitative attention, is particularly sensitive to changes in context and hence the former (but not the latter) may be reset if the conditions of training (which induced uncertainty) are subsequently altered. For example, in Beesley et al., following initial AX-O1, AY-O1, BX-O2, BY-O2 training (either with or without uncertainty) a new stage of training was employed where learning about entirely new outcomes (e.g., O3 and O4) was used to measure the associability of the various stimuli. If exploratory attention is "reset" upon the introduction of the new training context, but exploitative attention is not, then we would expect to see a maintenance of the effects of predictive validity (i.e., the associability of A and B would be higher than X and Y) but there would be no effect of outcome uncertainty – which was precisely the results that Beesley et al. observed. Unfortunately, this analysis is less easy to apply to the results of the current experiments as the associations that were used to assess the impact of outcome uncertainty and predictive validity (i.e. the within-compound associations) were being established *at the same time* as the stimulus-outcome relationships that created those variations in outcome uncertainty and predictive validity were being manipulated. Consequently, there was no change in context between training stages that would selectively reset exploratory attention in the current experiments as there was only one training stage.

Perhaps, instead, outcome uncertainty did facilitate the acquisition of within-compound associations, but the *expression* of these associations in performance was disrupted by the contextual transition from the training phase to the old/new test. For example, Le Pelley et al. (2009), like Mackintosh (1975) before them, considered the possibility that the associability of a stimulus could directly impact the response that it could elicit. Thus, responding to some stimulus p was suggested to be the product of its associative strength and associability: $R_p = k\ \alpha_p\ V_p$ (Le Pelley et al., 2009; p. 313). If this was the case then the effects of exploitative attention (induced through differential predictive validity) established during training should transfer to the test and improve the hit rate when changes were made to the predictive stimuli relative to the nonpredictive stimuli, however the effects of exploratory attention (induced through outcome uncertainty) should not, as its values would be reset. It remains to be determined if this possibility constitutes a reasonable explanation for the current results. It is worth highlighting that the conditions of training and testing were specifically designed to be very similar – thus, as in training, each test trial comprised four stimuli (indeed on the "old" trials the stimuli were identical to the training trials). Furthermore, at test, the response requirements were spatially identical to training. It is always possible that the instructional difference between training and test disrupted the expression of exploratory attention in the current experiments; this being the case though, it starts to become difficult to imagine any circumstances under which exploratory attention could be of general use to behaviour.

Thus far, our explanation for why performance on the "New predictive" test trials was higher than on the "New nonpredictive" test trials has been based upon the assumption that the predictive cues acquired higher associability than the non-predictive cues during the training phase, allowing them to form stronger within-compound associations. Consequently, the change in the configuration of the compounds on the "New Predictive" test trials should

be easier to detect than on the "New Non-predictive" test trials and the proportion of correct responses should be higher during the former trials than on the latter trials. However, there is an alternative "outcome mediation" account for the current results, which Lagator et al., (2025) considered (and found evidence for) in their Experiment 2. When participants were presented with an old test trial (e.g., ABWX) the predictive cues would both activate O1, while the non-predictive cues would activate O1 and O2 to a lesser but equivalent degree. In contrast, on a new test trial (e.g., ADWX), one of the predictive cues would activate O1 and another predictive cue would activate O2. An elevated level of concurrent activation of O1 and O2 could be a signal that participants used to discriminate between the new and the old test trials. When a change occurred in a non-predictive compound (e.g., from ABWX to ABWY) there would be no difference in the level of concurrent activation of O1 and O2, since all the non-predictive cues are expected to be equally associated with both the outcomes. This analysis can also explain the reduced difference in performance between the predictive and the non-predictive cues on the uncertain test trials. For example, consider the following types of the uncertain test trials: EFPQ (old) and E**G**PQ (new). E and F were paired with O1 on 80% of the training trials but were also paired with O2 on 20% of the training trials, while the reverse was true for G. Therefore, each of the predictive cues on the old test trials will activate both the outcomes (although to a different extent), and so the difference in concurrent outcome activation between the old and the new test trials would be reduced when the cues are associated with outcome uncertainty.

Lagator et al. (2025) noted that the foregoing account of their results could have broader generality as it aligns with previous demonstrations which show that predictive cues enter into novel associations more quickly than non-predictive cues (e.g., Le Pelley et al., 2011). If we assume, based on connectionist theories of representation (e.g., e.g. McClelland & Rumelhart, 1985; McLaren & Mackintosh, 2001), that only a sample of elements of each

stimulus can be detected on any one trial, then each sample of a predictive stimulus will be associated with the same outcome, unlike the samples of a non-predictive stimulus. As a result, samples of a predictive stimulus would be more likely to activate each other through outcome mediation, supporting the establishment of a stronger network of within-compound associations and hence a more consistent stimulus representation across training trials – facilitating the formation of associations with novel outcomes.

The current experiments revealed that participants' accuracy in detecting a change to the elements of a complex compound stimulus was influenced by the predictive history of those elements. When a change was made to the previously predictive elements, then hit rates were higher than when a change was made to the previously nonpredictive elements. Given that, objectively, the correlation between the pairs of predictive elements was identical to the correlation between the nonpredictive elements, the current results imply that experience established a bias in the representation of the compound away from the veridical. In this sense the representation was not accurate. The representation was, however, an *appropriate* extraction from the conditions established during training. Learned predictiveness has, as Rumelhart et al. (1986) put it, permitted "an arbitrarily connected neural network to develop an internal structure that is appropriate for a particular task domain."

References

Beesley, T., Nguyen, K. P., Pearson, D., & Le Pelley, M. E. (2015). Uncertainty and predictiveness determine attention to cues during human associative learning. *Quarterly Journal of Experimental Psychology*, *68*(11), 2175-2199.

Biederman, I. (1972). Perceiving real-world scenes. *Science, 177*, 77–80.

Chao, C. M., McGregor, A., & Sanderson, D. J. (2021). Uncertainty and predictiveness modulate attention in human predictive learning. *Journal of Experimental Psychology: General*, *150*(6), 1177-12002.

Cho, S. A., & Cho, Y. S. (2021). Uncertainty modulates value-driven attentional capture. *Attention, Perception, & Psychophysics*, *83*, 142-155.

Collins, L., & Pearce, J. M. (1985). Predictive accuracy and the effects of partial reinforcement on serial autoshaping. *Journal of Experimental Psychology: Animal Behavior Processes*, *11*(4), 548-564.

Dickinson, A. (1980). *Contemporary animal learning theory* (Vol. 1). Cambridge University Press.

Eatherington, C. J., & Haselgrove, M. (2022). The role of prediction in learned predictiveness. *Journal of Experimental Psychology: Animal Learning and Cognition*, *48*(3), 203-221.

Erlebacher, A. (1977). Design and analysis of experiments contrasting the within-and between-subjects manipulation of the independent variable. *Psychological Bulletin*, *84*(2), 212-219.

Esber, G. R., & Haselgrove, M. (2011). Reconciling the influence of predictiveness and uncertainty on stimulus salience: a model of attention in associative learning. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1718), 2553-2561.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41(4), 1149–1160. https://doi.org/10.3758/ BRM.41.4.1149

Fei-Fei, L., Iyer, A., Koch, C., & Perona, P. (2007). What do we perceive in a glance of a real-world scene?. *Journal of vision*, *7*(1), 1-29.

Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. NewYork: Wiley.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological science*, *20*(4), 464-472.

Greenwald, A. G. (1976). Within-subjects designs: To use or not to use?. *Psychological Bulletin*, *83*(2), 314-320

Hall, G. (1991). *Perceptual and Associative Learning*. Oxford: Clarendon.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Hogarth, L., Dickinson, A., Austin, A., Brown, C., & Duka, T. (2008). Attention and expectation in human predictive learning: The role of uncertainty. *Quarterly Journal of Experimental Psychology*, *61*(11), 1658-1668.

JASP Team (2023). JASP (Version 0.17.1) [Computer software].

Ju, J., & Cho, Y. S. (2023). The modulation of value-driven attentional capture by exploration for reward information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *49*(2), 181-197.

Kattner, F. (2015). Transfer of absolute and relative predictiveness in human contingency learning. *Learning & Behavior*, *43*, 32-43.

Lagator, S., Muñiz-Diez, C., Beesley, T., & Haselgrove, M. (2025). Mechanisms underlying the accuracy of stimulus representations: Within-event learning and outcome mediation. *Journal of Experimental Psychology: Animal Learning and Cognition*.

Larkin, M. J., Aitken, M. R., & Dickinson, A. (1998). Retrospective revaluation of causal judgments under positive and negative contingencies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*(6), 1331.

Lattal, K. A., & Miles, A. K. (2024). Resurgence and behavioral contrast, compared and contrasted. *Perspectives on behavior science*, *47*(2), 315-334.

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 Section B*, *57*(3), 193-243.

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*(2), 220-229.

Le Pelley, M. E., Haselgrove, M., & Esber, G. R. (2012). Modelling attention in associative learning: Two processes or one?. *Learning & behavior*, *40*, 292-304.

Le Pelley, M. E., & McLaren, I. P. L. (2003). Learned associability and associative change in human causal learning. *The Quarterly Journal of Experimental Psychology: Section B*, *56*(1), 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*(10), 1111-1140.

Le Pelley, M. E., Suret, M. B., & Beesley, T. (2009). Learned predictiveness effects in humans: A function of learning, performance, or both?. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*(3), 312-327

Le Pelley, M. E., Turnbull, M. N., Reimers, S. J., & Knipe, R. L. (2010). Learned predictiveness effects following single-cue training in humans. *Learning & Behavior*, *38*(2), 126-144.

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*(4), 446-465.

Lochmann, T., & Wills, A.J. (2003). Predictive history in an allergy prediction task. In F. Schmalhofer, R.M. Young & G. Katz (Eds.), *Proceedings of EuroCogSci 03* (pp. 217222). Mahwah, NJ: Lawrence Erlbaum Associates.

Luque, D., Vadillo, M. A., Le Pelley, M. E., & Beesley, T. (2017). Prediction and uncertainty in associative learning: Examining controlled and automatic components of learned attentional biases. *Quarterly Journal of Experimental Psychology*, *70*(8), 1485-1503.

Lute, R. D. (1%3). Detection and recognition. In R. D. Lute, R. R. Bush, & E. Galanter. (Eds.), *Handbook of mathematical psychology* (pp. 103-189). New York: Wiley.

Mack, A. (2003). Inattentional blindness: Looking without seeing. *Current directions in psychological science*, *12*(5), 180-184.

Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological review*, *82*(4), 276-298.

McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*, 159-188

McLaren, I. P. L., & Mackintosh, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, *28*(3), 211-246.

Melchers, K., Lachnit, H., & Shanks, D. (2004). Within-compound associations in retrospective revaluation and in direct learning: A challenge for comparator theory. *The Quarterly Journal of Experimental Psychology Section B*, *57*(1), 25-53.

Melchers, K. G. , Lachnit, H. & Shanks, D. R. (2006). The Comparator Theory Fails to Account for the Selective Role of Within-Compound Associations in Cue-Selection Effects. *Experimental Psychology, 53* (4), 316-320.

Neisser U. (1979). The control of information pickup in selective looking. In: Pick A.D.(Ed.), *Perception and its development: A tribute to Eleanor Gibso n*(pp. 201–219). Hillsdale, NJ: Erlbaum.

Nosofsky, R. M. (1991). Stimulus bias, asymmetric similarity , and classification, *Cognitive Psychology, 23,* p. 94-140

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological review*, *87*(6), 532-552.

Pearce, J. M., Kaye, H., & Hall, G. (1982). Predictive accuracy and stimulus associability: Development of a model for Pavlovian learning. In M. Commons, R. Herrnstein, & A. R. Wagner (Eds.), *Quantitative analysis of behavior* (Vol. 3, pp. 241-256).

Pearson, D., Chong, A., Chow, J. Y. L., Garner, K. G., Theeuwes, J., & Le Pelley, M. E. (2024). Uncertainty-modulated attentional capture: Outcome variance increases attentional priority. *Journal of Experimental Psychology: General, 153*(6), 1628–1643.

Potter, M. C. (1976). Short-term conceptual memory for pictures. *Journal of Experimental Psychology: Human Learning and Memory, 2,* 509–522.

Reger, M., Vrabie, O., Volberg, G., & Lingnau, A. (2025). Actions at a glance: The time course of action, object, and scene recognition in a free recall paradigm. *Cognitive, Affective, & Behavioral Neuroscience*, 1-15.

Rensink, R. A., O'regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological science*, *8*(5), 368-373.

Rescorla, R. A. (1980). Simultaneous and successive associations in sensory preconditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, *6*(3), 207-216.

Rescorla, R. A. (1981). Within-signal learning in autoshaping. *Animal Learning & Behavior*, *9*(2), 245-252.

Rescorla, R. A., & Colwill, R. M. (1983). Within-compound associations in unblocking. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*(4), 390-400.

Rescorla, R. A., & Cunningham, C. L. (1978). Within-compound flavor associations. *Journal of Experimental Psychology: Animal Behavior Processes*, *4*(3), 267-275.

Rescorla, R. A., & Durlach, P. J. (1981). Within-Event Learning In Pavlovian Conditioning. In N. E. Spear & R. R. Miller (Eds) Information Processing in Animals: Memory Mechanisms (pp. 81-111). Psychology Press.

Rescorla, R. A., & Freberg, L. (1978). The extinction of within-compound flavor associations. *Learning and Motivation*, *9*(4), 411-427.

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

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, *323*(6088), 533-536.

Schmajuk, N. A., Lam, Y. W., & Gray, J. A. (1996). Latent inhibition: a neural network approach. *Journal of Experimental Psychology: Animal Behavior Processes*, *22*(3), 321-349.

Simons D.J., Chabris C.F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, 28, 1059–1074.

Simons, D. J., & Rensink, R. A. (2005). Change blindness: Past, present, and future. *Trends in cognitive sciences*, *9*(1), 16-20.

Torrents-Rodas, D., Koenig, S., Uengoer, M., & Lachnit, H. (2021). Evidence for two attentional mechanisms during learning. *Quarterly Journal of Experimental Psychology*, *74*(12), 2112–2123. https://doi.org/10.1177/17470218211019308

Torrents-Rodas, D., Koenig, S., Uengoer, M., & Lachnit, H. (2023). The effect of prediction error on overt attention and learning in humans. *Behavioural Processes*, *206*, 104843. https://doi.org/10.1016/j.beproc.2023.104843

Thomas, G. V., & Cameron, G. N. (1974). Response rate, reinforcement frequency, and behavioral contrast. *Journal of the Experimental Analysis of Behavior*, *22*(2), 427-432.

Walker, A. R., Luque, D., Le Pelley, M. E., & Beesley, T. (2019). The role of uncertainty in attentional and choice exploration. *Psychonomic Bulletin & Review*, *26*(6), 1911-1916.

Walker, A. R., Navarro, D. J., Newell, B. R., & Beesley, T. (2022). Protection from uncertainty in the exploration/exploitation trade-off. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(4), 547–568.

Wasserman, E., & Berglan, L. R. (1998). Backward blocking and recovery from overshadowing in human causal judgement: The role of within-compound associations. *The Quarterly Journal of Experimental Psychology: Section B*, *51*(2), 121-138.

Welham, A. K., & Wills, A. J. (2011). Unitization, similarity, and overt attention in categorization and exposure. *Memory & Cognition*, *39*(8), 1518-1533.

Williams, B. A. (2002). Behavioral contrast redux. *Animal learning & behavior*, *30*(1), 1-20.

Witt, J. K., Taylor, J. E. T., Sugovic, M., & Wixted, J. T. (2015). Signal Detection Measures Cannot Distinguish Perceptual Biases from Response Biases. *Perception*, *44*(3), 289-300.

Statements and declarations

Ethical considerations

The experiments reported in this paper received ethical approval (approval no. F1565) by the Ethics Committee at the School of Psychology, University of Nottingham, UK on 04/09/2024.

Consent to participate

All participants provided informed consent prior to participating.

Consent for publication

Not applicable.

Declaration of conflicting interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding statement

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Economic and Social Research Council of the UK [grant number ES/W013215/1].

Data availability

Experimental materials, data, and data analysis scripts are available at https://osf.io/s56wj/overview?view_only=5f2528f099f747619f1b35d6023e8e80.