

Learned Changes in Outcome Associability

Martyn C. Quigley

4251631

Thesis submitted to the University of Nottingham for the
degree of Doctor of Philosophy [2019]

Abstract

The ability to learn associations between events is of crucial importance for human and non-human animals. Without this ability we would be unable to identify when danger is imminent (e.g., the signs of a looming predator) or how to process the causal structure of our environment (e.g., understanding the role of causes and their corresponding effects). As such, it is essential to understand the factors which influence the formation of learned associations. Indeed, a plethora of studies have been conducted with this goal in mind and many models have been devised to elucidate the factors which influence how associative learning between events arises. A critical factor that a number of these models have identified concerns how the predictiveness of a cue, or a conditioned stimulus, can influence associative learning. For instance, studies have identified that a cue that has reliably predicted an outcome in the past, enters into subsequent associations better than a cue that has been an unreliable predictor of an outcome (e.g., the learned predictiveness effect; see Le Pelley & McLaren, 2003).

The predictive ability of a cue can also interact with and influence learning about other cues. For example, prior learning about a cue and an outcome can block learning between an additional cue and the same outcome, when these cues are presented together (i.e., blocking; see Kamin, 1968; 1969). Furthermore, simply presenting a cue in isolation or in compound – with no prior exposure to either – can also influence learning, with the elements of a cue compound entering into an association with an outcome to a lesser extent than a cue presented in isolation (i.e., overshadowing, see Pavlov, 1927). As

such, a considerable amount of evidence has clearly identified the importance of the prior predictive ability of a cue and the extent to which cues can interact with one another to influence learning. However, much less research has explored the extent to which the *predictability* of an outcome can influence learning and the extent to which outcomes can interact with one another. Interestingly, key associative learning models assume that the prior predictability of an outcome should have no bearing on subsequent learning about this outcome, once separated from the cue which it was first paired with (e.g. Esber & Haselgrove, 2011; Le Pelley, 2004; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972).

These models also assume that outcomes are unable to interact with other outcomes, when presented in compound, in the same manner as cues. This approach appears logical: cues can provide information which allows an organism to change its behaviour in anticipation of an event, therefore providing the organism with an evolutionary advantage. Consequently, it appears reasonable to suggest that the associability of cues can be modified based on their associative history or due to the presence of other cues. In contrast, however, when an outcome is presented the event of significance has already happened. As such, there appears to be no evolutionary advantage to the associability of an outcome being modified based on its prior associability, or the presence of other outcomes: once the event has happened it is too late for any advantage to be gained.

Nevertheless, there is evidence to suggest that the prior predictability of an outcome can influence subsequent learning about this outcome. In light of this, Experiments 1 – 4 of Chapter 2 sought to further investigate whether the

prior predictability of an outcome can influence subsequent learning about this outcome when separated from the cue which the outcome was first paired with. Experiments 1 and 2 revealed that participants learned better about a previously well-predicted outcome than a previously less well-predicted outcome, a finding which is inconsistent with key associative learning models. Experiments 3 and 4 respectively revealed that the concurrent predictability of the outcomes and Stage 1 training is important in obtaining this effect.

Chapter 3 sought to assess whether the prior predictability of an outcome could influence subsequent learning about an additional outcome, when these outcomes were presented in compound. The experiments in this chapter employed an outcome blocking procedure using a food allergist task. In Experiment 5 an outcome blocking procedure was employed where an outcome compound was presented consisting of a previously well-predicted outcome and a novel outcome. In addition, an outcome compound was presented featuring a novel outcome and a surprising outcome. In this experiment an interaction effect was observed providing evidence to suggest that the prior predictability of an outcome can influence learning about other outcomes, as has been suggested by some other studies. In contrast to some previous studies in animals, however, an outcome facilitation effect was observed as opposed to the blocking effect which is typically – although not exclusively – seen in cues (cf. Batson & Batsell, 2000). Experiment 6 replicated this effect using a diagnostic task where the outcomes were causes and the cues were effects, thus demonstrating that the effect was present regardless of the type of judgement task participants completed. These findings

are inconsistent with the predictions of causal model theory (e.g. Waldmann & Holyoak, 1992).

Experiment 7 employed an additional type of control trial within the outcome blocking procedure employed in Experiments 5 – 6. When employing this additional type of control, which received exposure only in Stage 2, an outcome interaction effect was again observed. However, when this additional control trial was employed an outcome competition effect was observed in contrast to the outcome facilitation effect. Experiment 8 sought to tease apart two theoretical accounts derived from the Mackintosh model (1975) and the Standard Operating Procedures (SOP) model proposed by Wagner (1981) to account for the findings of Experiment 7. According to the Mackintosh account the effects reported in Experiments 5-7 should only be present following the second trial when the outcomes are presented in compound (i.e., when the associability of the cues have had an opportunity to be modified). The SOP account, however, stipulates that the effect should be present following the first trial. As such, Experiment 8 employed a one-trial outcome blocking procedure to test these accounts. This experiment provided partial support for the SOP account of the experiments reported in this chapter (whilst facilitation was observed amongst the outcomes competition was not).

To further explore whether outcome interaction effects can be observed when additional types of procedures are employed, Chapter 4 sought to assess whether interaction effects could also be observed when outcome overshadowing and outcome relative validity procedures were employed. These experiments also used the food allergist task. In the outcome overshadowing experiment, both a single (i.e., elemental) outcome and a

compound outcome (i.e., two outcomes) were reliably predicted. In the outcome relative validity task outcome compounds were presented which contained a common outcome (i.e., an outcome which was predicted by more than one cue) and unique outcomes. Some of these outcome compounds were reliably predicted and some were partially predicted. When an outcome relative validity procedure was employed no interaction effects were observed. One possible account of this finding is to suggest that both competition and facilitation effects were operating simultaneously, thus obscuring any differences. Yet, when an outcome overshadowing procedure was employed a competition effect was observed which is typically seen in cues. That is, an outcome which was presented alone was learned about better than when it was presented in compound with an additional outcome.

Taken together these results provide evidence to suggest that the associative history of an outcome can influence subsequent learning about this outcome when separated from the cue which the outcome was first paired with. The results of Chapter 3 suggest that both facilitation and competition effects can be observed amongst outcomes, thus being similar to cues, which have also displayed both competition and facilitation effects. Although there was an absence of an interaction effect when a relative validity procedure was employed in Experiments 10 and 11 in Chapter 4, Experiment 9 of Chapter 4 also provides evidence to suggest that the associability of an outcome can interact with and influence subsequent learning about additional outcomes when an outcome overshadowing procedure is employed. As such, these experiments suggest that outcomes might be more similar to cues than originally conceived of in associative learning models. There are, however,

several important factors to consider and methodological issues which arise when employing outcome associability and interaction procedures. The challenge of future work will be to identify the key factors which influence outcome associability, to understand the impact of this on key associative learning models, and to consider how outcomes are conceived of and distinguished from cues.

Publications

The data contained in this thesis have been published, or submitted for publication, as follows:

Chapter 2

Quigley, M. C., Eatherington, C. J., & Haselgrove, M. (2019). Learned changes in outcome associability. *The Quarterly Journal of Experimental Psychology*, 72(2), 209 – 221. doi: 10.1080/17470218.2017.1344258

Acknowledgements

I am deeply grateful for the support and guidance of my supervisor Dr Mark Haselgrove. Throughout the course of my PhD I could not have asked for a better supervisor. Your accommodating approach, patience and sense of humour has kept my neuroticism at bay. Moreover, your infectious passion for associative learning and intuitive ability to design experiments has inspired and (hopefully) improved me as a researcher and made writing and preparing presentations a pleasurable experience. I continue to admire your ability to formulate experimental designs to test almost any question. Please do realise how thankful I am for all your help.

I would also like to thank Dr Jasper Robinson for his advice and guidance throughout this process; my internal examiner Dr Charlotte Bonardi for her insightful comments and considered feedback; and my external examiner, Dr Anna Thorwart, for taking the time to read this work and raising some excellent points. I am also grateful to the staff, students and officemates who have made my time at Nottingham so enjoyable, and of course all of those who have participated in the experiments. It has been a pleasure meeting you all.

The research in this thesis would also not be possible without the funding provided by the Economic and Social Research Council (ESRC) and the School of Psychology, I will forever be grateful for this opportunity. I am also indebted to my family and friends, and the support they have provided outside of work. None of these opportunities would be possible without you. Thank you to all of you for putting up with me!

Table of Contents

Abstract	2
Publications	8
Acknowledgements	9
Table of Contents	10
List of Figures	14
List of Tables	16
 Chapter 1: General Introduction.....	 17
1.1 Learning by Association	17
1.2 Associative Learning Models.....	19
1.3 CS History Influences Learning	27
1.4 CSs Interact to Influence Learning	32
1.5 The Associability of USs	43
1.6 Interactions Amongst Outcome Elements	50
1.7 Aims of the thesis	54
 Chapter 2: Outcome Predictability Biases Learning	 57
2.1 Experiment 1	63
2.1.1 Method	64
2.1.1.1 Participants	64
2.1.1.2 Apparatus and Stimuli.....	65
2.1.1.3 Procedure.....	65
2.1.1.3 Stage 1.....	66
2.1.1.3 Stage 2.....	66
2.1.1.4 Scoring and Error Classification.....	67
2.1.2 Results and Discussion	69
2.1.2.1 Response Time Data	69
2.1.2.2 Error Data.....	72
2.2 Experiment 2	74
2.2.1 Method	76
2.2.1.1 Participants	76
2.2.1.2 Apparatus and Stimuli.....	76
2.2.1.3 Procedure.....	77
2.2.2 Results and Discussion	78
2.2.2.1 Response Time Data	78

2.2.2.2 Error Data	81
2.2.2.3 Ratings data	83
2.3 Experiment 3	85
2.3.1 Method	87
2.3.1.1 Participants	87
2.3.1.2 Apparatus and Stimuli.....	87
2.3.1.3 Procedure.....	88
2.3.2 Results and Discussion	89
2.3.2.1 Stage 1 – Response Time	89
2.3.2.2 Stage 2 – Response Time	91
2.3.2.3 Error Data.....	93
2.4 Experiment 4	96
2.4.1 Method	97
2.4.1.1 Participants	97
2.4.1.2 Apparatus and Stimuli.....	97
2.4.1.3 Procedure.....	98
2.4.2 Results and Discussion	99
2.4.2.1 Stage 2 – Response Times.....	99
2.4.2.2 Error Data.....	101
2.5 General Discussion	102
Chapter 3: Outcome Facilitation and Outcome Blocking.....	112
3.1. Experiment 5	118
3.1.1 Method	119
3.1.1.1 Participants	119
3.1.1.2 Apparatus and Stimuli.....	120
3.1.1.3 Procedure.....	121
3.1.1.3 Stage 1.....	122
3.1.1.3 Stage 2.....	123
3.1.1.3 Test.....	124
3.1.2 Results and Discussion	126
3.1.2.1 Stage 1.....	126
3.1.2.2 Stage 2.....	128
3.1.2.3 Test Data - Ratings	130
3.1.2.4 Test Data (Difference Ratings).....	132
3.2 Experiment 6	133
3.2.1 Method	138

3.2.1.1 Participants	138
3.2.1.2 Apparatus and Stimuli.....	138
3.2.1.3 Procedure.....	140
3.2.1.3 Stage 1.....	141
3.2.1.3 Stage 2.....	143
3.2.1.3 Test.....	143
3.2.2. Results and Discussion	144
3.2.2.1 Stage 1.....	145
3.2.2.2 Stage 2.....	147
3.2.2.3 Test Data	149
3.3 Experiment 7	151
3.3.1 Method	153
3.3.1.1 Participants	153
3.3.1.2 Apparatus and Stimuli.....	153
3.3.1.3 Procedure.....	154
3.3.2. Results and Discussion	155
3.3.2.1 Stage 1.....	156
3.3.2.2 Stage 2.....	157
3.3.2.3 Test Data	160
3.4 Experiment 8	163
3.4.1 Method	168
3.4.1.1 Participants	168
3.4.1.2 Apparatus, Stimuli and Procedure	169
3.4.2 Results and Discussion	170
3.4.2.1 Stage 1.....	171
3.4.2.2 Stage 2.....	173
3.4.2.3 Test Data	174
3.5 General Discussion	177
Chapter 4: Outcome Overshadowing and Outcome Relative Validity ...	186
4.1 Experiment 9	193
4.1.1 Method	196
4.1.1.1 Participants	196
4.1.1.2 Apparatus and Stimuli.....	196
4.1.1.3 Procedure.....	198
4.1.1.3 Training Stage	199
4.1.2 Results and Discussion	200

4.1.2.1 Training Stage	201
4.1.2.2 Test Data	202
4.1.2.3 Type of Overshadowing	204
4.2 Experiment 10	209
4.2.1. Method	212
4.2.1.1 Participants	212
4.2.1.2 Apparatus and Stimuli.....	213
4.2.1.3 Procedure.....	213
4.2.2 Results and Discussion	215
4.2.2.1 Training Stage	216
4.2.2.2 Test Data	217
4.3. Experiment 11	219
4.3.1 Method	221
4.3.1.1 Participants	221
4.3.1.2 Apparatus and Stimuli.....	221
4.3.1.3 Procedure.....	221
4.3.2 Results and Discussion	222
4.3.2.1 Training Stage	223
4.3.2.2 Test Data	224
4.3 General Discussion	227
 Chapter 5: General Discussion	 236
5.1. Summary of findings	236
5.2. Outcome Predictability	240
5.3. Outcomes Interactions	244
5.4. Future work	253
5.5. Conclusion.....	260
 References	 262

List of Figures

Chapter 1: General Introduction

<i>Figure 1</i>	Illustration of Wanger's SOP model (1981)	26
-----------------	---	----

Chapter 2: Outcome Predictability Biases Learning

<i>Figure 2</i>	Example (target) trials for both Stage 1 and Stage 2	67
<i>Figure 3</i>	Response times in Experiment 1, Stage 1	69
<i>Figure 4</i>	Response times in Experiment 1, Stage 2	71
<i>Figure 5</i>	Response times in Experiment 2, Stage 1	78
<i>Figure 6</i>	Response times in Experiment 2, Stage 2	80
<i>Figure 7</i>	Mean Ratings for the outcomes in Experiment 2	83
<i>Figure 8</i>	Response times in Experiment 3, Stage 1	89
<i>Figure 9</i>	Response times in Experiment 3, Stage 2	91
<i>Figure 10</i>	Response times in Experiment 4, Stage 2	99

Chapter 3: Outcome Facilitation and Outcome Blocking

<i>Figure 11</i>	Food stimuli used in Experiment 5	120
<i>Figure 12</i>	Example screens of each stage of Experiment 5	125
<i>Figure 13</i>	Mean correct responses in Experiment 5, Stage 1	126
<i>Figure 14</i>	Mean correct responses in Experiment 5, Stage 2	128
<i>Figure 15</i>	Test ratings for Experiment 5	130
<i>Figure 16</i>	Test ratings (outcome specific) for Experiment 5	131
<i>Figure 17</i>	Additional Food stimuli used in Experiment 6	139
<i>Figure 18</i>	Example screens of the Diagnostic task in Experiment 6	142
<i>Figure 19</i>	Group Causal responses in Experiment 6, Stage 1	144
<i>Figure 20</i>	Group Diagnostic responses in Experiment 6, Stage 1	144
<i>Figure 21</i>	Group Causal responses in Experiment 6, Stage 2	146
<i>Figure 22</i>	Group Diagnostic responses in Experiment 6, Stage 2	146
<i>Figure 23</i>	Test ratings (outcome specific) for Experiment 6	148
<i>Figure 24</i>	Food stimuli used in Experiment 7	153
<i>Figure 25</i>	Group Causal responses in Experiment 7, Stage 1	155

<i>Figure 26</i>	Group Diagnostic responses in Experiment 7, Stage 1	155
<i>Figure 27</i>	Group Causal responses in Experiment 7, Stage 2	157
<i>Figure 28</i>	Group Diagnostic responses in Experiment 7, Stage 2	157
<i>Figure 29</i>	Test ratings (outcome specific) for Experiment 7	160
<i>Figure 30</i>	Group Causal responses in Experiment 8, Stage 1	170
<i>Figure 31</i>	Group Diagnostic responses in Experiment 8, Stage 1	170
<i>Figure 32</i>	Group Causal responses in Experiment 8, Stage 2	172
<i>Figure 33</i>	Group Diagnostic responses in Experiment 8, Stage 2	172
<i>Figure 34</i>	Test ratings (outcome specific) for Experiment 8	174

Chapter 4: Outcome Overshadowing and Outcome Relative Validity

<i>Figure 35</i>	Example screens of each stage of Experiment 5	197
<i>Figure 36</i>	Group Causal responses, Experiment 9, Training	200
<i>Figure 37</i>	Group Diagnostic responses, Experiment 9, Training	200
<i>Figure 38</i>	Test ratings (outcome specific) for Experiment 9	202
<i>Figure 39</i>	Test ratings (outcome specific), Experiment 9, OS Type	204
<i>Figure 40</i>	Examples screens of each stage of Experiment 10	214
<i>Figure 41</i>	Group Causal responses Experiment 10, Training	215
<i>Figure 42</i>	Group Diagnostic responses in Experiment 10, Training	215
<i>Figure 43</i>	Test ratings (outcome specific) for Experiment 10	217
<i>Figure 44</i>	Group Causal responses, Experiment 11, Training	222
<i>Figure 45</i>	Group Diagnostic responses, Experiment 11, Training	222
<i>Figure 46</i>	Test ratings (outcome specific) Experiment 11	224

List of Tables

Chapter 1: General Introduction

<i>Table 1</i>	Design of Le Pelley and McLaren (2003)	31
<i>Table 2</i>	Simplified Design of Durlach and Rescorla (1980)	40
<i>Table 3</i>	Design of Miller and Matute (1998)	53

Chapter 2: Outcome Predictability Biases Learning

<i>Table 4</i>	Design of Griffiths et al. (2015)	59
<i>Table 5</i>	Design of Experiment 1	63
<i>Table 6</i>	Design of Experiment 2	75
<i>Table 7</i>	Design of Experiment 3	87
<i>Table 8</i>	Design of Experiment 4	97

Chapter 3: Outcome Facilitation and Outcome Blocking

<i>Table 9</i>	Design of Experiment 5	119
<i>Table 10</i>	Design of Experiment 6	138
<i>Table 11</i>	Design of Experiment 7	152
<i>Table 12</i>	Design of Experiment 8	168

Chapter 4: Outcome Overshadowing and Outcome Relative Validity

<i>Table 13</i>	Design of Matute, Arcediano and Miller (1996)	191
<i>Table 14</i>	Design of Experiment 9	194
<i>Table 15</i>	Design of Experiment 10	211
<i>Table 16</i>	Design of Experiment 11	221

Chapter 1: General Introduction

1.1 Learning by Association

Learning associations between events is a crucial ability which human and non-human animals require to complete a range of essential and trivial tasks. Whether it be an animal detecting the events that signal the presence of a predator, or a human determining which signs they need follow to navigate safely to their chosen destination, in both cases an association must be learned. Moreover, this association must be updated based on subsequent information. For instance, let us imagine the case of an animal which has identified a visual cue that signals a predator is near, perhaps a sudden movement in its visual array. It is likely that this signal in and of itself does not always indicate the presence of a predator, it could be that this movement was the consequence of the wind blowing. In this instance the association between the visual cue and the presence of a predator would need to be modified (especially in the presence of reliable or more salient cues). The same applies to a human navigating a route to their destination; perhaps the visual cues which previously signalled the necessary route no longer apply (due to recent alterations in the terrain), in which case the association between these cues and the goal destination will need to be revised.

Given the importance of this associative process, a plethora of studies have been conducted with animals (e.g., Hall & Pearce, 1979; Haselgrove, Esber, Pearce & Jones, 2010; Kamin, 1968; 1969; Mackintosh & Little, 1969; Pavlov, 1927; Rescorla, 1968) and humans (e.g., Bonardi, Graham, Hall, & Mitchell, 2005; Le Pelley & McLaren, 2003; Livesey & McLaren, 2007; see

Le Pelley, Mitchell, Beesley, George, Wills, 2016, for review), which have aimed to identify the mechanisms underpinning the formation and subsequent modification of these associations. Perhaps most notably, Pavlov (1927) highlighted the importance of simply pairing two types of stimuli together in order to form associations between the two stimuli. For instance, in his experiments Pavlov established a predictive relationship between an auditory stimulus (e.g., a metronome) and the presentation of a biologically significant stimulus (i.e., food) to dogs in a laboratory (i.e., metronome → food). After several instances of this arrangement Pavlov observed that upon hearing the auditory stimulus, the dogs preformed a salivary response in anticipation of receiving the food. With this observation, Pavlov therefore demonstrated that by pairing two stimuli together an association can be formed between them. In the context of Pavlov's study, the animals became conditioned to respond to the neutral stimulus (i.e., the metronome) in the same way they would to the biologically significant stimulus (i.e., the food). As such, the neutral stimulus became a 'conditioned stimulus' (i.e., a CS, a stimulus which produces a response after conditioning or training), whilst the biologically significant stimulus, which produces a response in and of itself, served as an unconditioned stimulus (i.e., a US).

Rescorla (1968) provided further evidence that an association can be formed by establishing a predictive relationship between two events. In a series of fear conditioning experiments with rats, Rescorla varied the extent to which a shock (i.e., the US) was presented in the presence or absence of a tone (i.e., the CS). When the delivery of the shock was equally probably in both the absence and presence of a tone (i.e., there was no predictive relationship

between the two stimuli) no learning took place concerning the association between the tone and the shock. However, when the shock had the same probability of occurring, but only following the presence of the tone, substantial learning took place regarding the association between the tone and the shock. In a second experiment Rescorla further manipulated the probabilities of the shock occurring in the presence or absence of the tone. When the probability of the shock occurring in the presence of the tone was higher, better learning took place, yet when the probability of the shock occurring in the absence of the tone was higher, poorer learning was observed. These results demonstrate the importance of establishing a predictive relationship between stimuli to ensure that an association is formed.

1.2 Associative Learning Models

In order to account for the variables that influence how an association is formed between stimuli which share a predictive relationship, a number of formal learning models have been developed. Generally, these models can be categorised into two classes by identifying whether they place their focus on the processing of a CS or the US, when determining the strength of the association which is formed between these stimuli. Models which place focus on the US are referred to as US-processing models, whilst models which focus on the role of the CS are referred to as CS-processing models. According to the first class of these models, the US can support a finite amount of associative strength which is divided amongst the CSs, or cues, presented prior to the presentation of a US, or an outcome.¹ One of the most influential of these

¹ The terms cue and conditioned stimulus (CS), and outcome and unconditioned stimulus (US) are used interchangeably in this thesis.

models is that proposed by Rescorla and Wagner (1972). This model states that the change in the associative strength of a cue A (ΔV_A) is determined by the following equation:

$$\Delta V_A = \alpha \beta (\lambda - \Sigma V) \quad (1)$$

where α and β respectively denote the salience of the CS and the US, λ signifies that maximum amount of conditioning supported by the US, and ΣV represents the summed associative strength of all stimuli present on a trial. The model states that the change in the associative strength between stimuli is a function of the salience of the CS (α) and the US (β), and crucially the difference between the extent to which the US is anticipated as a result of the CSs which precede it (ΣV), and the maximum amount of conditioning supported by the US (λ). Simply stated the model proposes that the US must be unexpected for learning to occur. Applied to a simple fear conditioning procedure, for example that employed by Rescorla (1968) the model suggests that upon the initial pairing of the tone (i.e., the CS) and the shock (i.e., the US), considerable learning would take place concerning the association between these stimuli, as the US is unexpected. However, each subsequent trial will provide diminishing returns in terms of the amount of associative strength gained between these stimuli (resulting in a negatively accelerated learning curve), as the US is expected due to the previous trials.

This model has achieved considerable attention since its inception, and it remains a major success due to its ability to accommodate a range of associative learning phenomena (which will be discussed subsequently).

Nevertheless, the model does have several shortcomings (see Miller, Barnet & Grahame, 1995; Pearce & Bouton, 2001) and many key models developed following the Rescorla and Wagner model have shifted their focus to the role of the CS in determining the extent to which an association is formed between stimuli. According to these CS processing models or “attentional models”, the *predictiveness* of a cue, that is, how predictive a cue has been in the past, influences its *associability*.² Associability refers to the extent to which a stimulus enters into associations with other stimuli. If a stimulus has high associability, it will more readily enter into associations than a stimulus with low associability. One of the most influential CS processing models was proposed by Mackintosh (1975), henceforth the ‘Mackintosh model’. According to this model the change in the associative strength of a stimulus e.g. ‘A’ (i.e. ΔV_A) is determined by the following equation:

$$\Delta V_A = \alpha_A \beta (\lambda - V_A) \quad (2)$$

here, α_A refers to the associability of stimulus A, β is, again, a fixed learning-rate parameter determined by the salience of the outcome, λ the asymptotic value of learning possible on that trial, and V_A the current associative strength of Stimulus A. Crucially, in this model the parameter α is deemed to alter as a function of experience: Stimulus A will gain in associability if it is a better predictor of an outcome than all other stimuli

² The term predictiveness in this thesis refers to how predictive a cue or CS has been of an outcome or US in the past. This is differentiated from the term predictability which refers to how well-predicted an outcome or US has been predicted in the past.

present on a given trial, otherwise the associability of A will decrease. These rules can be expressed in the following formulae:

$$\Delta\alpha_A > 0 \text{ if } |\lambda - V_A| < |\lambda - V_r| \quad (3)$$

$$\Delta\alpha_A < 0 \text{ if } |\lambda - V_A| \geq |\lambda - V_r| \quad (4)$$

in which $\Delta\alpha_A$ is the change in the associability of stimulus A; $\lambda - V_A$ is the difference between the asymptotic value and the associative strength of stimulus A, and $\lambda - V_r$ is the discrepancy between the asymptotic value and the associative strength of all cues present on that trial, apart from A (that is to say, the remainder). Informally, this associability rule permits organisms to pay attention to features of an environment that are predictive of important outcomes and ignore those features which are less predictive. This rule therefore provides organisms with an evolutionary advantage. For example, an animal may learn to focus its attention on the call of a predator which has been a reliable predictor of danger in the past, but to pay less attention to the noise made by a companion which has been less reliable.

The shift of focus toward the role of the CS in determining the strength of association formation is not unique to the Mackintosh Model. Alternative attentional-learning models such as that proposed by Pearce and Hall (1980) come to the same conclusion regarding the importance of CS processing, despite offering a contradictory approach to Mackintosh regarding the mechanisms which determine whether a cue will experience an alteration in its associability. In the Pearce and Hall model, cues which are relatively unreliable or ‘poor’ predictors of an outcome are deemed to receive

an enhancement in associability. Formally, the Pearce and Hall model states that a change in the associative strength (ΔV) of a stimulus (e.g. A), is determined by the following equation:

$$\Delta V_A = S_A \alpha_A \lambda \quad (5)$$

in this equation S refers to the salience or intensity of the CS which can range from 0 to 1, α_A refers to the associability the stimulus, whilst λ refers to the intensity of the US. Importantly, in this model the intensity of a CS is differentiated from its associability, unlike the Mackintosh model. Furthermore, the Pearce and Hall model stipulates that the rule governing the changes in the associability of A is determined by the following equation:

$$\alpha_n = |\lambda - \sum V|_{n-1} \quad (6)$$

again, α represents the associability of a CS with n representing the trial. The change in the associability of the CS is determined by the absolute difference between the intensity of the US (λ) and the summed associative strength of the stimuli ($\sum V$) on the previous trial ($n - 1$). This associability modification rule enables a stimulus which is rendered a reliable predictor of an outcome to undergo decrements in its associability as learning progresses. Informally, this enables an organism to pay less attention to a good predictor of an event (as they know the outcome), and thus devote their attention to learning about those features of an environment which are less certain

predictors of an event. Given the conflicting, yet logical, approaches adopted by both the Mackintosh and Pearce-Hall models, more recently developed hybrid models have sought to resolve the discrepancies between these models (Esber & Haselgrove, 2011; Le Pelley, 2004), which also place their focus on the role of changes in processing of the CS.

Alternatively, a different type of associative learning model to those outlined above has been provided by Wagner (1981; also see Brandon, Vogel & Wagner 2003). According to this model, referred to as *SOP* (Standard Operating Procedures), each stimulus consists of multiple elements which represent the features of a stimulus. For example, a food stimulus may have elements which reflect the odour and appearance of the food. Each of these elements of a stimulus can occupy one of three different states which influence how an association may be formed and the type of response which is elicited: a primary state (A1), a secondary state (A2) or an inactive state (I). Prior to a stimulus's initial presentation, the stimulus is in the inactive state, where it is unable to attract attention or produce a response. However, following the first presentation of a stimulus its elements enter into their A1 activity state. In this state the stimulus can produce its strongest response, being at the focal point of the organism's attention, before decaying into the A2 activity state. When residing in the A2 state the stimulus's elements are still able to produce a response, albeit weaker in magnitude; during this state the stimulus's elements are considered at the periphery of an organism's focus. Once the stimulus's elements pass through the A2 state they will eventually return to their inactive state.

Importantly, the model prescribes conditions under which stimuli elements can enter into their A1 and A2 states. According to the model, stimuli elements can only enter their A1 state, the focal point of attention, by being presented. Alternatively, stimulus elements can enter into their A2 state by the process of decay from A1 or through associative activation from another stimulus directly from their inactive state. Both these states (i.e., A1, A2) are also deemed to have a limited capacity, although A1 is granted much less capacity than A2. For instance, Wagner proposed that perhaps only 2-3 nodes (i.e., stimuli) can be fully in their A1 state (i.e., being truly at the focal point of an organism's attention), however, somewhere in the region of 10-15 nodes could be fully in their A2 state (i.e., at the periphery of an organism's attention). The rate at which the elements of a stimulus decays from each of these states differs accordingly as well, with decay from A1 – A2 assumed to occur rapidly and decay from A2 to inactive to occur at a slower rate. The limited capacity principle inherent in this model would appear to be an essential mechanism to any model, whether it is an explicitly stated component of a model or implicitly implied, as is the case in the Rescorla and Wagner model (1972).

The state which the elements of a stimulus reside in is crucial to the type of learning which will take place. For instance, when elements of the CS and US are both in their A1 state, thus being at the focal point of the organisms' attention, excitatory learning is assumed to take place (if, for example, a tone is reliably paired with food). Alternatively, if the elements of a CS are in their A1 state, when, the elements of the US are in their A2 state, the model suggests inhibitory learning will take place. Furthermore, if the

CS's elements are in their A2 state and the US's elements are in their A1 state, no learning will take place. This also applies if CS and US elements are both in their A2 state. This model can accommodate a diverse range of associative learning phenomena which have been difficult for some models to account for (e.g., Rescorla & Wagner, 1972), such as latent inhibition and timing effects (i.e., inter-trial interval effects), and has resulted in novel predictions (e.g., context specific habituation and latent inhibition). To avoid undue repetition a fuller description of the model will be provided in Chapter 3.

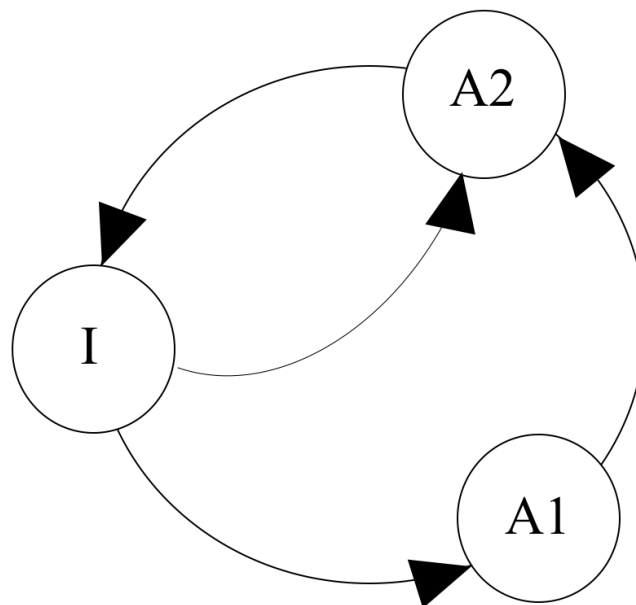


Figure 1. The Standard Operating Procedures (SOP) model by Wagner (1981) proposes that stimuli consist of elements which can reside in one of three states: an inactive state (I), a primary state (A1), and a secondary state (A2). When a stimulus is first presented its elements are activated from their inactive state to A1. These elements will then rapidly decay from A1 to their A2 state, before gradually decaying back to their inactive state. Arrows represent the methods by which elements may enter each of the respective states. To enter A1, elements must be activated from their inactive state. However, elements can enter A2 as a result of decay from A1, or by being associatively activated from their inactive state to A2.

A brief consideration of some of the key associative learning theories, noted above, reveals a variety of different ideas concerning how associations may form between stimuli. Although some models agree with respect to their point of focus when it comes to variations in stimulus processing (i.e., CS or US), they differ regarding the mechanisms responsible for the modification of associations. Alternatively, other models differ with regards to how stimuli are perceived and processed. Nevertheless, despite these differences, some core assumptions are shared by all these models in order to account for a range of associative learning phenomena (e.g., Kamin, 1969; Lubow, 1959; Pavlov, 1927) reported prior to the formation of these models, to which we shall now turn our attention.

1.3 CS History Influences Learning

A key assumption in associative learning models, whether formally stated or informally implied, is the idea that the associability of a CS is determined not only by its salience (i.e., the intensity of the CS), but also through prior exposure with the CS. For instance, Rescorla and Wagner (1972) acknowledge the influence of this factor, although they do not provide a rule explicitly stating how such exposure may influence the associability of a CS. In contrast, Mackintosh (1975) and other CS processing models explicitly prescribe a set of formulae detailing how the associability of a stimulus can be modified based on the history of this stimulus.

This notion, inherent in these models, is empirically substantiated by a range of associative learning phenomena, such as the latent inhibition effect first reported by Lubow and Moore (1959). In their study Lubow and Moore employed a classical conditioning procedure with goats and sheep. A flashing

light or turning rotor served as the CSs and a shock delivered to the right foreleg of the animal served as the US. One of the CSs (e.g., the flashing light) was first exposed ten times without being paired with the US in a pre-exposure stage. In a subsequent stage both CSs (i.e., the flashing light and the turning rotor) were reliably paired with the US. During this second stage animals demonstrated better conditioning with the novel CS, which had not previously been presented in Stage 1, than the CS which had been pre-exposed in Stage 1. These results demonstrate that pre-exposure with a non-reinforced CS impacted upon learning about this stimulus. CS processing accounts of this effect suggest that attention to the pre-exposed stimulus declines (as it does not predict a US), therefore impairing subsequent learning about this stimulus (Mackintosh, 1975; Pearce & Hall, 1980). According to SOP, impaired learning about the pre-exposed stimulus occurs because the pre-exposed stimulus enters into an association with the context (as a result of presenting this stimulus repeatedly in the same context in the absence of an outcome). This then results in the context priming the CS into the A2 state therefore preventing learning about the stimulus. A prediction of this account is that the latent inhibition effect should in fact be context-specific (i.e., by changing contexts the effect should be absent) which studies have revealed to be the case (i.e., Channell & Hall, 1983; Hall & Channell, 1985; also see Honey & Hall, 1989).

Channell and Hall (1983) demonstrated the context-specific nature of latent inhibition in three experiments using an appetitive conditioning procedure with rats. In their first experiment pre-exposure to a light impaired subsequent learning about the light when it was reliably paired with food,

relative to a control group which received no pre-exposure to the light (i.e., latent inhibition). In their second experiment, one group of rats received both pre-exposure and conditioning between the light and the food in the same context (Group Same). However, a second group received pre-exposure to the light in one context and conditioning between the light and food in a different context (Group Different). In order to produce a different context, Channell and Hall altered the background noise in the chamber in addition to the odour of the chamber. In Group Same impaired conditioning was observed to the light, relative to Group Different, therefore demonstrating that context can modulate latent inhibition.

The latent inhibition effect has been observed in a diverse range of species, including: bees, rats, mice, rabbits and fish (e.g., Abramson & Bitterman, 1986; Chamizo & Mackintosh, 1989; Lubow, Markman & Allen, 1968; Shishimi, 1985; for review see Lubow, 1973). The effect has also been demonstrated in humans (see Byrom, Mstefi & Murphy, 2018), both with children and adults, although in adults the use of a masking task has been deemed necessary when certain procedures are employed (although not all, see Forrest, Mather & Harris, 2016), in order to ensure the participant's attention is engaged and the pre-exposed stimulus does not appear too conspicuous (Ginton, Urca & Lubow, 1975). For instance, Schnur and Ksir (1969) demonstrated latent inhibition with humans using an eyelid conditioning procedure when a masking task was employed. In the masking task participants were required to predict when one of two lights would be presented (they were informed it was a problem-solving task). When a group was first exposed to a light during this task, which would subsequently be

paired with the delivery of a puff of air to the eye, they demonstrated impaired learning relative to controls who received no pre-exposure to the light. These results therefore further demonstrate the ubiquitous nature of latent inhibition and highlight how the history of a CS can influence learning in animals and humans.

Perhaps the most compelling evidence that the history of a CS can influence its subsequent associability, however, is demonstrated by the learned predictiveness effect. In a learned predictiveness experiment participants are exposed in a first stage of training to a set of cues, some of which are reliable predictors of an outcome and some of which are unreliable predictors of an outcome. In Stage 2, these cues are arranged in a manner which makes them equally reliable predictors of novel outcomes. What is typically observed in such experiments is a bias in learning during Stage 2 to the cues that had a greater predictive history in Stage 1. Le Pelley and McLaren (2003, see Table 1 for design) demonstrated this in a computerised food allergist task with humans, where participants were presented with a number of different food compounds. Some foods reliably predicted reactions in a first stage of training, however, others were poor predictors of reactions. In Stage 2, previously reliably predictors were presented in compound with poor predictors and reliably predicted a novel outcome. Following this a test stage was presented where compounds were presented consisting of stimuli which served as good predictors or poor predictors in Stage 1 and participants were asked to rate the extent to which the compounds predicted the novel outcomes in Stage 2. At test participants learned better about the compounds consisting of good predictors in Stage 1, relative to compounds consisting of poor predictors.

This effect appears robust having been replicated a number of times with different procedures (e.g., Bonardi, Graham, Hall & Mitchell, 2005; Le Pelley, Turnbull, Reimers & Knipe, 2010; Le Pelley, Vadillo & Luque, 2013; O’Brien & Raymond, 2012). The Mackintosh model accounts for – and indeed predicts – this effect based on the formulae the model prescribes to account for changes in the associability of a stimulus. Recall, that according to these formulae the cue which serves as a better predictor of an outcome, relative to another cue, will receive enhancements in its associability. The Mackintosh model also assumes that these changes in cue associability occur independently of the outcome which the cue was first paired with (which is demonstrated in the learned predictiveness effect). The model therefore suggests that changes in a cue’s associability can be transferred to situations where the cue is paired with novel outcomes.

Table 1. Design of Le Pelley and McLaren (2003).

<i>Stage 1</i>	<i>Stage 2</i>	<i>Test</i>
AV – O1	AX – O3	AC
BV – O2	BY – O4	BD
AW – O1	CV – O3	VX
BW – O2	DW – O4	WY
CX – O2	EF – O3	EH
DX – O1	GH – O4	FG
CY – O2	IJ – O3	IJ
DY – O1	KL – O4	KL

Note. A-Z refers to cues, represented by foods. O1-O4 refer to outcomes, represented by allergic reactions. A-D served as good predictors in Stage 1, whilst V-Z served as poor predictors. E-L served as filler trials.

1.4 CSs Interact to Influence Learning

A second notion inherent in these models pertains to the impact of presenting multiple CSs in compound (i.e., together) prior to a US. A large body of literature has demonstrated (i.e., Honey & Hall, 1988; Jones & Haselgrove, 2013; Kamin, 1968, 1969; Le Pelley, Oakeshott & McLaren, 2005) that when two or more cues are presented in compound this can impact upon the way in which these cues are learned about relative to when they are presented alone, either resulting in decrements in learning about these stimuli (termed cue competition effects) or enhancements (termed cue facilitation effects). The earliest observation of these cue interaction effects was first reported by Pavlov (1927, p. 141). According to Pavlov, when two CSs are presented together (i.e., AB+), learning about each stimulus will be attenuated, relative to a stimulus conditioned alone. This effect is termed the overshadowing effect.

Accounts of this effect typically appeal to the idea that there is some form of competition between the elements of the compounds as they seek to enter into an association with the US. According to the SOP model both elements of the compound enter into their A1 state, which has limited capacity, and thus reciprocally detract from each other's ability to fully enter into an association with the US to the same extent as they would if presented alone. The Rescorla and Wagner model accounts for the effect in a similar manner, as the model assumes that the US can support a finite amount of associative strength which is divided amongst the CSs presented prior to a US, thus each stimulus prevents the other from being learned about as well as they would be if they were presented alone. According to the Mackintosh model, however,

overshadowing arises not because of some form of competition for a limited resource, but because of a failure to attend to the stimuli which are not deemed the best predictor of an outcome. According to the equations outlined by Mackintosh (1975), the associability of the stimuli conditioned together will decline as they are equally predictive of the US.

Another manifestation of this type of cue-interaction effect was reported by Kamin (1968, 1969), referred to as “blocking”, and is arguably one of the most pivotal findings in associative learning over the last 50 years. In these experiments Kamin employed a shock conditioning procedure with two groups of rats: an experimental group (i.e. Group Blocking) and a control group. In Group Blocking rats were first presented with a white noise, the CS, which reliably predicted the onset of a US, in this case a 1-milliampere foot shock. In a second stage rats were exposed to a cue compound (AB+) featuring the white noise (A) and a light (B) which reliably predicted the same US. In the control group rats were only exposed to Stage 2 training. Following Stage 2 training both groups were presented with the light (the ‘B’ element of the compound) and the amount of fear produced was measured with conditioned suppression. Crucially, Group Blocking displayed little fear toward the light at test in contrast to the control group, despite both groups being presented with an equal number of pairings of the light (B). This indicates that Stage 1 training trials with the white noise (A) attenuated learning about the light (B) shock association, in Group Blocking.

This effect has been reported in a diverse range of species when a variety of procedures have been employed, including rats (e.g. Bonardi, 1991; Dopson, Pearce, Haselgrove, 2009; Jones & Haselgrove, 2013, Mackintosh,

1975), mice (Bonardi, 2010, Sanderson, Jones & Austen, 2016), snails (Prados et al., 2013), honeybees (Cheng & Spetch, 2001; Guerrieri, Lachit, Gerber & Giurfa, 2005) and humans (Crookes & Moran, 2003; Dickinson, Shanks & Eveden, 1984; Le Pelley, Oakeshott, McLaren, 2005). Kamin (1968, 1969) accounted for this effect by suggesting that learning would only take place between stimuli when the outcome was surprising. Evidence of unblocking would certainly seem to support this proposition. For instance, studies have demonstrated that when a feature of the outcome changes (such as its intensity, see Bakal, Johnson & Rescorla, 1974; Holland, 1984) in the presence of the additional element of the compound cue, the element can be ‘unblocked’, resulting in learning taking place between this additional element of the cue compound and the outcome.

To account for blocking the Rescorla and Wagner model suggests that learning will not take place between the additional element (e.g. B) of the compound CS and the US, following training with A+ trials, as the additional component of the compound does not signify anything surprising about the US (the US is already predicted by A). This therefore renders $\lambda - \sum V = 0$ on all trials featuring B, ensuring $\Delta V = 0$ on these trials (i.e., no learning will take place). Recall, that according to the Rescorla and Wagner model the amount of learning which is supported by the US is finite and therefore the US can support only a limited amount of learning. This therefore means that once the association between A and the US is fully learned, the US is depleted of its ability to support conditioning and is incapable of entering into an association with the novel component of the compound (B).

The Mackintosh model, however, accounts for blocking by suggesting that A's associability is enhanced, relative to B, when these stimuli are first presented in compound because of Stage 1 training with A. As such, given the formulae Mackintosh prescribed regarding changes in the associability of stimuli, A serves as a better predictor of the US and continues to accrue greater associability, whilst B undergoes a decrease in its associability, as it is presented in compound with a more reliable predictor of the outcome. Alternatively, the SOP account of blocking assumes that the Stage 1 training with A results in the US being primed into its A2 state, therefore when the additional component of the compound is presented, B, little excitatory learning can take place between B and the US.

Another similar type of cue-interaction effect to that of blocking was also observed by Wagner, Logan, Haberlandt and Price (1968), referred to as relative stimulus validity, or the relative validity effect. In the study of interest, Wagner et al. conducted a shock conditioning procedure with two groups of rats. In both groups compound CSs were presented prior to the delivery of a shock in a single stage of training. In a true discrimination (TD) group, rats were presented with the compound cues AX+ and BX – (where + signals shock and – denotes the absence of a shock). In this group the rats could learn that the presence of stimulus A signalled that a shock would be delivered, B signalled the absence of a shock and X was paired with the outcome 50% of the time as a result of being presented alongside stimuli A and B. In a second group rats were presented with the compound cues AX +/- and BX +/- (where +/- denotes that 50% of the time the compound was followed by the US and 50% of the time it was not). This group was referred to as the pseudo-

discrimination (PD) group as it was impossible to discriminate between the stimuli regarding their association with the US (i.e., all three stimuli were paired with the US an equal number of times). The central finding of this study was that despite X being with paired with the US an equal number of times in both groups, rats displayed greater learning about the relationship between X and the US in the PD group, relative to group TD. Wagner et al. accounted for this effect, informally, by suggesting that X's validity in predicting the US *relative* to other stimuli determines the strength of the association formed between X and the US, as opposed to its absolute validity as a predictor. In the TD group A and B serve as good predictors relative to X and thus little is learned about X. In the PD group all the stimuli serve as equal predictors of X, thus they are learned about equally.

The Mackintosh analysis of this effect can be easily inferred from the account of this effect provided by Wagner et al. According to the Mackintosh model, in the TD group Stimuli A and B serve as better predictors of the US, thus their associability increases relative to X, whilst X's decreases. In Group PD, however, all stimuli are learned about equally as there is no difference in their predictive ability. The Rescorla and Wagner model can also account for the effect assuming the associability of the US (β) is deemed to be greater when predicted (β_1) than when the US is not predicted (β_2) (see, Miller, Barnet & Grahame, 1995). If these parameters are set then learning about the stimuli in the PD group is assumed to be comparable for A, B and X, as the limited amount of learning supported by the US is divided amongst the cues which are equally predictive of the US. In the TD group, however, the model suggests that initially X accrues some associative strength on AX+ trials, however, on

the BX- trials, X loses some of the associative strength it has gained which is subsequently accrued by A on the following AX+ trials. SOP accounts for the effect in a comparable manner. According to SOP, on AX+ trials, learning takes place about both A and X, however learning about X and the US is extinguished on the BX- trials, allowing for A to occupy more of A1 and to enter into a stronger relationship with X. In the PD group, however, this is not the case as elements of the compound are both reinforced and non-reinforced at a comparable rate.

Thus far the types of interaction effects noted above demonstrate how presenting multiple CSs in compound can result in a decrement in the associability of these stimuli. However, presenting multiple stimuli prior to a US does not exclusively cause competition effects. Although historically key associative learning models have been developed to account for competition effects such as those reported above, the range of effects which can be observed when presenting multiple CSs in compound prior to a US can vary from competition, to no effect at all, to facilitation (i.e., enhanced learning about the elements of a compound, relative to when the elements are presented alone).

Several studies employing flavour-aversion procedures have demonstrated cue-facilitation effects under conditions where the opposite effect would be expected according to associative learning models (see Clarke, Westbrook & Irwin, 1979; Galef & Osborne, 1978; Rusiniak, Hankins, Garcia & Brett, 1979). For instance, Durlach and Rescorla (1980), conducted a flavour-aversion experiment with rats where a compound (see Experiment 1, Durlach & Rescorla, 1980) consisting of an odor and a taste was followed by

the delivery of a US (i.e., lithium chloride). When presented with this compound rats demonstrated better learning about the relationship between the odor and the US, relative to when an odor was presented in elemental form (i.e., without the taste). Given that the experimental paradigm employed involves presenting a compound stimulus consisting of a more salient element, an overshadowing effect would be expected according to most associative learning models. However, the data from these experiments demonstrated “potentiation” (i.e., enhanced learning about the elements of a compound as a result of being presented in compound with another stimulus).

A comparable effect has also been demonstrated when procedures have been employed that would traditionally be expected to produce blocking. In a flavour conditioning procedure Batson and Batsell (2000) presented a group of rats with an almond odor-solution which was followed by an injection of lithium chloride (i.e., the US) to induce toxicosis (i.e., A+). In a second conditioning stage the rats were again presented with the almond odor solution, which was followed by lithium chloride, however, this time it was presented in compound (i.e., AB+) with a taste stimulus, denatonium saccharide solution. Surprisingly rats that received conditioning trials with the almond odor before being presented with the compound, demonstrated better learning about the relationship between the taste stimulus and the US than a control group of rats who had not received this conditioning stage. These results run contrary to the expectations of many associative learning models. As noted above, according to these models (and a plethora of studies), pre-training with an element of a compound should result in attenuated learning about the relationship between the additional element in Stage 2 and the US, however, this study observed the

opposite effect. The authors termed this effect augmentation (see also Batsell & Batson, 1999; Good, Allswede, Curley, Batsell, 2015; Vadillo & Matute, 2010; Vadillo & Matute, 2011).

Given the surprising nature of these cue-facilitation effects and their theoretical implications, several explanations have been provided to account for these findings (see LoLordo & Droungas, 1989). For instance, Durlach and Rescorla (1980) suggested that potentiation may arise as a result of within-compound associations (i.e., associations forming between the elements of a compound). Recall, that potentiation arises when a more salient stimulus is presented in compound with a less salient stimulus, thus it is possible that better learning is produced as a result of the enhanced associability of the more salient CS, which is associated with the other element of the compound through within-compound associations. Interestingly, potentiation is indeed sensitive to differences in the saliences of the compound elements, with studies demonstrating that the potentiated element of the compound must have lower salience than its companion CS to produce potentiation (see Slotnick, Westbrook & Darling, 1997; also see Bouton, Jones, McPhillips & Swartzentruber, 1986).

However, perhaps the most compelling evidence of the role of within-compound associations in producing potentiation is provided by a series of experiments by Durlach and Rescorla (1980). To assess if within-compound associations can indeed occur in arrangements used to study potentiation, these authors employed a procedure where two odor-taste compounds were conditioned to rats in an initial Training Stage. Following this, the taste of one of these compounds was then extinguished by being presented repeatedly on

its own. At a subsequent Test Stage, the rats displayed a weaker aversion to the odor that had been presented alongside its extinguished taste partner, than the odor which had been presented alongside a non-extinguished taste-partner, thus demonstrating that within-compound associations can indeed form in procedures which produce potentiation.

Table 2. Simplified Design of Durlach and Rescorla (1980), Experiment 4

<i>Conditioning</i>	<i>Extinction</i>	<i>Test 1</i>	<i>Test 2</i>
O1+, O2S+, O3N+	N –	O1 vs O2	O3

Note. + = toxin; – = nonreinforcement; N = sodium chloride; S = saccharin; O1 = almond; O2 = banana; O3 = Strawberry. Counterbalanced conditions are omitted for clarity.

Furthermore, in a subsequent experiment (see Table 2) these authors directly assessed the extent to which within-compound associations can contribute to the potentiation effect. To assess this, Durlach and Rescorla (1980) presented rats with three different solutions in the form of an odor element (O1+), and two odor-taste compounds (O2S+ and O3N+). In an initial Training Stage, the rats were conditioned to each of these solutions. Following this, the taste element of only one of the compounds was extinguished (N –). At a first test Stage, rats displayed more aversion to the odor element of the compound whose taste partner had not been extinguished (i.e., O2), than the odor presented in elemental form (O1+) when they were presented with these two odors, thus demonstrating the standard potentiation effect. Importantly, however, in a second test stage, when presented with the odor whose taste-partner had been extinguished (O3), there was no difference between the amount of odor solution consumed for this odor and the odor presented in elemental form (O1). These results suggest that potentiation is dependent on the

continued aversion to the taste element of the compound, as expected by a within-compound association account of the effect.

This within-compound account of potentiation can also be used to accommodate the augmentation effect reported by Batson and Batsell (2000). Although in an augmentation procedure, a single odor (or taste) element is presented prior to being part of an odor-taste compound, it is possible that the enhanced associative strength (as a result of pre-training with the odor) enhances the associability of the taste component, when presented in compound with the odor, as a result of a within-compound association. Indeed, evidence of the role of within-compound associations mediating augmentation in a comparable manner to potentiation has been provided by Allswede, Curley, Cullen and Batsell (2014). To assess the extent to which within-compound learning may mediate augmentation, these authors performed post-conditioning extinction of the pre-trained element of the compound. When doing so the augmentation effect was abolished, thus being consistent with the results reported by Durlach and Rescorla (1980).

Given the above studies, it is clear that cue-facilitation effects can, at the very least, be mediated by within-compound associations. However, the fact that extinguishing one of the cue-elements in the above studies abolished a cue-facilitation effect but did not reveal a competition effect, suggests that perhaps there is an additional mechanism influencing results. For example, Urcelay (2017), suggests there are three primary variables that influence the type of interaction effect which will be displayed when presenting CSs in compound: (1) relative stimulus duration, (2) variations in contingency, and (3) variations in contiguity. Regarding stimulus duration, Urcelay (2017) notes

that competition type effects are typically observed when trials are spaced by long inter-trial intervals (i.e., ITIs) and the stimulus duration is short (e.g., Kamin, 1968). However, in studies where an ITI is absent or shorter than the duration of the stimulus, facilitation is typically observed (i.e., Sissons, Urcelay & Miller, 2009). With regards to variations in contingency, Urcelay notes that decreasing the contingency between a CS and US, by exposing a CS alone, before presenting this stimulus in compound with a novel element predictive of a US, can also result in a facilitation effect as opposed to a competition effect (e.g., Blaisdell, Bristol, Gunther and Miller, 1998). Finally, concerning variations in temporal contiguity between a CS and a US, Urcelay notes that when temporal contiguity is minimal competition effects are typically produced. However, when temporal contiguity is increased these differences tend to disappear or be replaced by facilitation effects (e.g., Batsell, Wakefield, Ulrey, Reimink, Rowe, & Dexheimer, 2012). In addition to these variables, Urcelay also notes that there is evidence to suggest factors such as stimulus salience, number of trials and timing can also influence the type of cue-interactions which are observed. In summary, many factors can influence whether competition or facilitation effects are observed beyond within-compound associations. Moreover, it is clear that cue-interaction effects are not exclusively unilateral (i.e., only displaying one type of interaction effect, typically deemed to be competitive) and as such it appears best to conceive of these effects as appearing on a dimension ranging from competition, to no effect, to facilitation.

1.5 The Associability of USs

Thus far the literature surveyed above has focused on various factors known to influence CS associability. Primarily, two key factors have been considered: (1) CS predictiveness; and (2) the impact of presenting CSs in compound with novel or previously presented CSs. However, there is an additional critical component which must be learned about when an association is formed, which has not yet been considered to the same extent: the US, or the outcome. The US in animal studies typically refers to a biologically significant stimulus that produces a response from an organism without prior training. The US is often presented in conjunction with the CS or following the presentation of the CS (e.g., tone – shock; although see, Spetch, Wilkie & Pinel, 1981). In human studies using computerised tasks, where the US or outcome is not always explicitly biologically significant (e.g., a reaction a fictitious patient might experience), the temporal order can be key to defining the US or outcome. For instance, USs or outcomes are deemed to be those stimuli which follow the presentation of a cue (or CS). In this thesis, an outcome is defined as a stimulus which follows the presentation of an antecedent event. Importantly, the US or outcome further imbues the CS with value (in addition to its predetermined salience), by ensuring that the predictive information provided by the CS warrants the attention of an organism (e.g., Le Pelley, Mitchell & Johnson, 2013). As such, the US is essential in determining the extent to which an association is formed between it and the CS. Despite providing an essential role in ensuring that an association is learned between a CS and US, however, these two components are treated differently in associative learning models.

As noted above, an integral assumption in associative models is that a CS's associability can change as a result of its associative history, and the associative history of other CSs which it is presented in compound with. Moreover, any change in the associability of a CS is assumed to persist when the same CS is no longer paired with the US it was initially presented with. As such, the CS is granted an independence of processing, ensuring that any changes in CS associability may persist when the same CS is paired with a novel outcome. This is perhaps best demonstrated by the learned predictiveness effect (e.g., Le Pelley & McLaren, 2003). In a learned predictiveness experiment, the high predictive ability of a CS in a previous stage of training aids subsequent learning about the CS when the original outcome it was paired with is removed, and the CS is paired with a novel outcome. Yet, in such circumstances the predictability of the outcome is also high. As such, given that the high predictive ability of a CS aids subsequent learning about this CS, a key question which will be addressed in this thesis is whether the high predictability of an outcome also aids learning about this outcome when separated from the CS which it was first paired with.

According to key learning models, the associability of the US is fixed, being determined only by its salience. Furthermore, although the extent to which a US can support learning is dependent upon how surprising the US is (e.g., Rescorla & Wagner, 1972), this is dependent upon the US being paired with the CS which first established the predictability of the US (e.g., λ - ΣV in Equation 1). As such, unlike the CS, variations in US processing are deemed to be *cue dependent*. Notably, this approach is not exclusive to traditional associative learning models. Other, well established and more recent hybrid

models, (e.g. Esber & Haselgrove, 2011; Le Pelley, 2004; Pearce & Hall, 1980), also assume that learned variations in processing of an outcome are either non-existent or cue-dependent.

Although this differential treatment may at first appear curious, the reason for this asymmetry in these models may appear justified. In the framework of these models, it makes sense for the associability of a CS to change as a function of its predictive ability, yet there is little reason for the associability of a US to change. When a CS is presented it can signal the absence or presence of an event, thus learning about the predictive nature of a CS is beneficial to the organism and is an ability which has a clear evolutionary advantage (Hullinger, Kruschke & Todd, 2015). Yet, when a US is presented the event has already happened, as such it would appear counter-intuitive to suggest a mechanism exists which mediates associability changes of a US as it would appear to provide little benefit. Perhaps because of this reasoning, considerably less research in the associative learning literature has explicitly explored US associability and the role of the US in learning relative to the CS.

This can be demonstrated by a basic search of simple key terms relating to CSs and USs. For instance, searching the terms “conditioned stimulus learning” in the search engines Google Scholar and Web of Science produces the following number of results: 414,000 and 21,436³ respectively. However, when searching for the terms “unconditioned stimulus learning”, the following number of results are produced: 65,600; 3,719, an average of only 17% of the number of results produced when the term conditioned stimulus is used

³ Search conducted September 11, 2018.

instead. The same is also true when searching the terms: “conditioned stimulus associations” and “unconditioned stimulus associations” on these databases.

Searching the terms “conditioned stimulus associations” respectively produces results of 233,000 and 8,831. Yet, when searching the terms “unconditioned stimulus associations”, 42,100 and 1,567, results are respectively produced, an average of 18% of the number of results produced when using the term conditioned stimulus instead.

Nevertheless, despite receiving considerably less attention than CSs, there is evidence to suggest that outcomes can indeed play a critical role in determining how organisms learn associations beyond that anticipated by most learning theories. For instance, one such example of outcomes influencing learning in a manner inconsistent with, in this case, traditional *instrumental* theories of conditioning, was reported by Trapold (1970). In Trapold’s study rats completed a discrimination task where they were trained to press one of two levers (left or right) based on the presence of two different auditory stimuli. In the experimental group the outcome when one auditory stimulus was presented was food, whilst the outcome when the other auditory stimulus was presented was liquid sucrose. In the control group only one type of outcome was presented to animals regardless of the two auditory stimuli. Trapold found that the experimental group who were exposed to two different outcomes learned the correct responses faster than those in the control group. These results have also been replicated in a range of species when a variety of different procedures have been employed (e.g., Gonzalez & Alvarez, 2001; Miyashita, Nakajima, & Imada, 2000; Mok & Overmier, 2007; Overmier, Bull & Trapold, 1971; Urcuioli, 1991).

The results of these studies are of interest as prior to Trapold's study the role of the outcome in learning associations was considered to be that of "facilitating" learning between a conditioned stimulus and the response produced by this stimulus (see Urcuioli, 2005 for further context).⁴ However, the results from this study challenge the notion inherent in some theories of learning that the outcome serves exclusively as the mechanism which facilitates learning between a stimulus and a response. Indeed, to account for this effect Trapold and Overmier (1972) suggested that subjects form an expectation about the subsequent outcome which facilitates their learning about which specific lever press to make (i.e., outcome-specific expectations cued which responses to make; for a thorough review of how this may happen, see: Urcuioli, 2005).

Additional phenomena also suggest that the prior predictability of an outcome can influence how this outcome is subsequently responded to and (potentially) learned about. Predictability, in this sense, refers to the extent to which an outcome is expected to occur based on previous experience with the outcome. For example, a number of studies in both the human and animal literature have reported variations in the response produced by an outcome as a consequence of how well-predicted it has been in the past (see Marcos & Redondo, 2001). Interestingly, establishing an outcome as predictable has produced both diminution (i.e. attenuated responses; see Kimble & Ost, 1961; Marcos & Redondo, 1999) and facilitation (i.e. enhanced responding; see Donegan, & Wagner, 1987). A range of explanations have also been provided

⁴ Interestingly, some researchers (e.g., Hull, 1930) initially disagreed with the notion that outcomes merely facilitate learning between a CS and a CR, and instead suggested that the outcome is a unique component which is learned about during conditioning.

for these effects. Typically, however, these explanations invoke variations in the effectiveness of the outcome as a function of the presence (or absence) of the cue (λ - ΣV). This therefore means that explanations of these studies do not need to appeal to any change in cue-independent variations in outcome processing.

Other effects of interest have also been reported which could suggest that the associative past of outcomes can influence their subsequent associability, however, these effects have also typically been accounted for by appealing to the presence or absence of cues which have signalled the outcome. One such example is the US pre-exposure effect, where exposure to the US, prior to CS-US pairings, impairs subsequent learning supported by that US (e.g. Kremer, 1971; see Randich & LoLordo, 1979, for review). This finding has typically been understood in terms of context blocking (e.g. Baker & Mercier, 1982), according to which, pre-exposure with the US results in the formation of an association between the context and the US, therefore permitting the context to compete with the CS for associative strength on subsequent CS-US trials. This account has received much support (e.g. Baker & Mackintosh, 1979), however, notably, Baker and Mercier (1982), came to the conclusion that whilst the context may mediate the US preexposure effect, it is also conceivable that during pre-exposure to the US, learning is acquired regarding the unpredictable nature of the US, which influences subsequent learning in Stage 2 (See also: Randich & Lolordo, 1979). That is, the outcome is coded as unpredictable, thus impairing the formation of predictive relationships (associations) between the CS and this US at a subsequent stage. This latter explanation would seem to indicate that previous exposures to the

US could potentially change the associability of the outcome itself regardless of the cue, or context, which it was trained with, thus having important implications for many associative learning models.

Similarly, Baker and Mackintosh (1977) demonstrated that when a CS and US are presented in an uncorrelated manner, these presentations impair subsequent conditioning between the two stimuli. This ‘learned irrelevance’ effect has been accounted for by assuming that these uncorrelated pre-exposures result in the CS and the US being deemed ‘irrelevant’ (e.g. Baker & Mackintosh, 1979). However, this interpretation has been challenged. For instance, Matzel, Schachtman and Miller (1988) suggested that the deficit in learning is underpinned by pre-exposure to the CS, pre-exposure to the US, and a learned irrelevance effect. Furthermore, Bonardi and Ong (2003; also see Bonardi & Hall, 1996) revealed, in a review of the literature that there is no evidence of a genuine learned irrelevance effect above and beyond the effects of CS and US pre-exposure. It is also conceivable, however, as suggested by Baker and Mercier (1982), that subjects learn that the US itself is unpredictable preventing subsequent learning about this outcome to proceed as it would without any experience with the outcome.

Nevertheless, although the phenomena discussed above suggest that prior experience with an outcome can influence subsequent learning about the same outcome, there is little evidence in these studies to suggest that these changes persist when the outcome is separated from the cue which the outcome was first paired with (whether this be an explicit cue, or the experimental context operating as a cue). As such, these effects pose little challenge to associative learning models. More recently, however, a number of studies have

sought to assess if outcome predictability can indeed bias subsequent learning about these outcomes when separated from the cues with which they were first paired with (Griffiths, Erlinger, Beesley & Le Pelley, 2018; Griffiths, Livesey, & Thorwart, 2019; Griffiths, Mitchell, Bethmont & Lovibond, 2015; Thorwart, Livesey, Wilhelm, Liu, Lachnit, 2017). The results of these studies suggest that the prior associability of outcomes can indeed bias subsequent learning, thus posing a challenge to associative learning models and questioning the extent to which cues and outcomes do differ from one another. These studies will be discussed further in Chapter 2.

1.6 Interactions Amongst Outcome Elements

As noted previously, the associability of a cue (e.g., Cue A) can be influenced by the associative past of the cue, and the associative past of other cues presented in compound with the cue (e.g., cue interaction effects such as blocking). In light of the above literature, which leaves open the possibility that the associability of an outcome can also be influenced by previous experience with the outcome, it appears pertinent to also consider whether learning about an outcome can also be influenced by the associative past of other outcomes, when these outcomes are presented together. According to associative learning models there is little reason for the associability of an outcome to change based on its past associability, or the past associability of accompanying outcomes. To clarify, learning about a novel outcome should not be influenced by being presented in compound with an outcome which has previously been presented. The same logic applies as previously outlined when discussing why these models do not permit outcome associability to change as a result of its associative history. That is, once the outcome has been presented

- whether elemental or compound - the event has happened, therefore it would appear that a mechanism which modulates the associability of an outcome would be redundant. Nevertheless, there is evidence to suggest that outcomes may interact to produce interaction effects comparable to those observed in cues.

For instance, in his 1980 book “Second Order Conditioning”, Rescorla described a series of experiments exploring whether interactions can occur between the elements of a compound reinforcer. In these experiments a second-order conditioning procedure was employed with pigeons completing an autoshaping task where the presentation of visual cues was reinforced by the delivery of food. In a ‘blocking’ group, pigeons were presented with presentations of A – O1 and A – O1 O2, following first order conditioning where A was rendered as a non-reinforced stimulus. In a control group subjects were also presented with A – O1 O2, following first-order conditioning trials where A was rendered as a non-reinforced stimulus. In this group subjects were not exposed to the A – O1 pairings. The question of interest concerned the extent to which each group learned the A – O2 association. The results revealed that subjects in the control group, who had no exposure to A-O1, demonstrated greater learning about the association between A – O2 than those in the experimental group. These results therefore appear to demonstrate an ‘outcome blocking’ effect – learning about the A-O1 association attenuated the formation of the A-O2 association. Rescorla’s interpretation of these results however were primarily descriptive, focusing on the possibility that O1 in the experimental group distracted participants from attending to O2, when these were presented in compound. This interpretation suggests that rather than the

associative properties of the previously exposed outcome blocking learning about the novel outcome, the animals simply failed to notice it, thus providing an alternative non-associative account of the findings. Nevertheless, it is possible that the previous associability of the outcome prevented learning about the novel outcomes, therefore suggesting that outcomes may interact in a comparable way to cues.

Miller and Matute (1998) also further investigated this phenomenon in a series of sensory preconditioning experiments with rats (see Table 3). In their experiment rats in the experimental group (Group Outcome Blocking) were first presented with a click train which served as a signal (A) for a white noise (O1). In Stage 2, rats were then presented with the cue (A) and a compound outcome (O1 O2) featuring the white noise (O1) and a novel outcome (e.g. a tone, O2). In a subsequent stage the novel element of the compound outcome (O2) was reliably paired with a footshock. At test subjects were presented with the cue (A). The question of interest was whether the A – O1 pairings in Stage 1 blocked learning about the A – O2 association, thus resulting in the presentation of the cue at test failing to elicit the anticipation of O2, and hence the expectation of a footshock. This was precisely the effect observed: subjects displayed less fear suppression behaviour towards A when they received pairings of the cue with O1 in Stage 1, as opposed to a control group which had not received exposure to the cue in Stage 1, therefore producing an outcome blocking effect comparable to that reported by Rescorla (1980). The same effect was also observed with a control group who experienced unpaired presentations of the cue and O1 in Stage 1. These results suggest that blocking is not an effect unique to cues. As such, they are inconsistent with an

assumption of key associative learning models. Indeed, Miller and Matute (1998) note that demonstrations of outcome blocking appear to challenge the differential treatments of cues and outcomes in these models.

Table 3. Design of Miller and Matute (1998)

<i>Group</i>	<i>Stage 1</i>	<i>Stage 2</i>	<i>Stage 3</i>	<i>Test</i>
Outcome blocking	A – O1	A – O1 O2	O2 - Shock	A
Control group 1	Context only	A – O1 O2	O2 - Shock	A
Control group 2	A / O1	A – O1 O2	O2 - Shock	A

Note. A = Cue; O1 = Outcome 1; O2 = Outcome 2. ‘-’ = followed; / = unpaired.

Further evidence of outcomes interacting to produce comparable effects to those in cues was observed by Burger, Mallemat and Miller (2000). In their study these authors employed a conditioned lick suppression task with rats to assess whether an overshadowing effect could be observed in outcomes. A backward conditioning procedure was employed, thus the CSs were presented as the subsequent events as opposed to the US. In Stage 1 rats were first presented with a 1.0-mA, 0.5-sec foot-shock which served as the US. For the overshadowing group rats were presented with the US, followed by presentations of the compound CS consisting of a white noise (which served as the ‘overshadowed’ stimulus) and a flashing light (which served as the ‘overshadowing stimulus’). In a control group the rats were simply presented with an element of the compound (e.g., the white noise) following presentations with the US. At a Test Stage the ‘overshadowed’ stimulus alone was presented to the overshadowing group and a control group and the suppression of rats licking was measured. The results of the study revealed that

rats in the overshadowing group displayed little suppression of licking relative to a control group which had not received any trials with the CS compound. These results appear to demonstrate that overshadowing can be observed with outcomes in a comparable way to cues, although it is important to note that the stimuli which served as the outcomes or subsequent events in this study are those typically used to assess cue competition effects. For instance, in their study a shock was presented prior to presentation of a white noise and a flashing light, thus the outcomes for this study (the flashing light and white noise) are not the stimuli which typically serve as outcomes in a conditioning study. Nevertheless, the results of this study in addition to those of Rescorla (1980) and Miller and Matute (1998), suggest that comparable interaction effects to those observed in cues can also be observed in outcomes, with the above studies appearing to provide evidence of blocking and overshadowing in outcomes. In summary, the results of the studies above pose a further question to associative learning models which, at best, remain silent on the matter of whether the elements of an outcome compound interact.

1.7 Aims of the thesis

The literature reviewed above has identified factors which influence the formation of associations between stimuli, including the associative past of the specific stimulus being tested, and the associative past of accompanying stimuli when presenting a compound stimulus. Evidence has been reported which demonstrates that the prior associability of a cue influences subsequent learning involving this cue, when separated from the outcome which the cue was first paired with (e.g., Le Pelley and McLaren, 2003). In addition, a range

of interaction effects have also been reported which demonstrate that when presented in compound, the elements of a cue can interact to influence learning. These effects are largely anticipated by associative learning models (which some of these models are designed precisely to do), however, these models do not anticipate comparable effects to persist amongst outcomes. Despite this, however, the literature reviewed above has provided reason to believe such effects may also be observed amongst outcomes. These studies have provided evidence – or at the very least failed to deny an interpretation – which suggests that the associability of an outcome can influence subsequent learning about that outcome, or an accompanying outcome. In light of the above evidence, the studies reported in this thesis sought to assess whether the previous associability of an outcome can influence subsequent learning involving that outcome with human participants.

Experiments 1 – 4 sought to directly assess whether the prior associability of an outcome can influence subsequent learning about the outcome when separated from the cue which it was first paired with. The studies reported above, exploring outcome associability, have provided evidence to suggest that the prior predictability of an outcome can influence subsequent learning involving that outcome. Yet in these studies the outcome has never been separated from the cues with which it was first paired with. More recently such studies have been performed (to be discussed further in Chapter 2), however, given the relative scarcity of such studies these experiments sought to assess the generality of the findings from these studies using a serial letter-prediction task.

Experiments 5 – 8 sought to assess whether the cue-interaction effect of blocking could also be observed amongst outcomes with humans completing a food allergist task. The results from outcome interaction studies reported above have appeared to demonstrate blocking with outcomes with non-human animals using indirect conditioning methods. For example, Rescorla observed an outcome blocking effect with pigeons when a second order conditioning blocking procedure was used, whilst Miller and Matute observed the same effect when employing a sensory preconditioning experiment with rats. The aim of Experiments 5-8 was to assess the generality of this effect in humans using causal and diagnostic learning procedures, and to examine the potential mechanisms mediating any potential blocking effect.

Experiments 9 – 11 sought to assess whether the cue-interaction effects of relative validity and overshadowing could also be observed amongst outcomes. Given that cue interaction effects such as blocking, overshadowing and relative validity are deemed to be different manifestations of an overarching cue-competition effect, it is possible that the presence of outcome blocking might suggest comparable effects will also be observed when employing different outcome interaction designs. That is, if a comparable mechanism mediates interaction effects amongst cues and outcomes, it might be expected that previous observations of outcome interaction effects in animals would suggest that overshadowing and relative validity effects observed in cues may also be observed in outcomes. Notably, the absence or presence of these effects could reveal potential mechanisms mediating any interaction effects observed in outcomes.

Chapter 2: Outcome Predictability Biases Learning

A large body of evidence suggests that changes in the associability of a cue persist in the absence of the outcome which the cue was first paired with. For example, a cue which has previously served as a reliable predictor of an outcome will continue to attract an organisms' attention, relative to other poorer predictors, when this cue is paired with a different outcome (e.g., Le Pelley & McLaren, 2003). There remains, however, some ambiguity concerning whether the associability of outcomes can also change based on their associative history, and the extent to which any changes will persist in the absence of the cue which the outcome was first paired with.

Key associative learning models assume that changes in the associability of a cue will persist when separated from the outcome which it was first paired with, but do not consider the same be true of outcomes. When an outcome is separated from the cue which it was first paired with, the outcome's associability is reset according to these models. For example, if A – O1 training is provided to an organism, then according to associative theories of learning (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972), the processing of O1 will diminish as its prediction error reduces. However, if B – O1 trials are then given, the processing of the O1 will be restored, despite having a history of being well-predicted in the past. This therefore means that cues and outcomes receive differential treatment in these models, for a CS (e.g., A) can be established as a good predictor of an outcome, and then come to differentially control behaviour with respect to other outcomes (or tasks), whereas outcomes cannot. This differential

treatment – although at first curious – is justifiable. When presented with cues an organism can gain valuable predictive information about events which may follow, allowing them to prepare and respond. Therefore, modifying behaviour based on previously presented cues would appear to be a sensible strategy which would provide an organism with an evolutionary advantage (see Hullinger, Kruschke & Todd, 2014). Yet, modifying behaviour based on previously presented outcomes appears to bear little evolutionary value - by the time an outcome is presented the event has already happened, therefore it is too late to gain any advantage by preparing for the event.

As seen in Chapter 1, however, despite the assumption of these models a number of studies have been performed which question the validity of the differential treatment afforded to cues and outcomes in these models (e.g., Baker & Mackintosh, 1977; Baker & Mercier, 1982; Kimble & Ost, 1961; Kremer, 1971; Marcos & Redondo, 2001; Randich & LoLordo, 1979). These studies suggest that the associative history of an outcome can indeed influence subsequent learning involving that outcome. However, they do so without explicitly separating the outcome from the cue which it was first paired with (or it is unclear what served as the cue, e.g., context blocking). As such, alternative explanations can be derived based on the presence of the cue with which the outcome was first paired with, and it is not possible to assess whether the associability of the outcome independently influenced subsequent learning about this outcome.

More recently, however, given the theoretical relevance of this question, studies have sought to directly assess whether prior learning about an outcome can influence new learning about this outcome, when it is separated

from the cue which it was first paired with. The first of these studies was conducted by Griffiths, Mitchell, Bethmont, and Lovibond (2015). In their study, these authors employed a procedure in which participants were required to take on the role of a food allergist and identify which symptoms followed the consumption of certain foods in (fictitious) patients. The design of this study can be seen in Table 4.

Table 4. Design of Griffiths et al. (2015), Experiment 1

<i>Phase 1</i>	<i>Phase 2</i>	<i>Test</i>
A – p1, uØ	EY – p1, u2	E: pØ, p1, p2, uØ, u1, u2
B – p2, uØ	FY – p2, u1	F: pØ, p1, p2, uØ, u1, u2
X – pØ, u1	GY – p1, u1	G: pØ, p1, p2, uØ, u1, u2
X – pØ, u2	HY – p2, u2	H: pØ, p1, p2, uØ, u1, u2
AX – p1, u1		Y: pØ, p1, p2, uØ, u1, u2
AX – p1, u2		
BX – p2, u1		
BX – p2, u2		

Note. Letters A – Y denote foods [cues]; p1, p2 = predicted outcome dimensions; u1, u2 = unpredicted outcome dimensions; Ø = the absence of the outcome.

In Stage 1 of their experiment, compounds of foods (AX and BX) were followed by an outcome that comprised a compound of symptoms. One component of the outcome compound (e.g. itchiness, p1) was reliably predicted by cue A whilst another component (e.g. swelling, p2) was reliably predicted by cue B. Both of these foods, plus a third food, X, was paired with one outcome component (u1) 50% of the time and another outcome component (u2) the remaining 50% of the time, thus rendering these outcome components as less-well predicted. In Stage 2 novel cues (E, F, G, H), were reliably paired

with the previously less well-predicted component of an outcome and a previously well-predicted component of an outcome. During this stage participants were required to predict the components produced by the cues. The results of a final test demonstrated that participants better learned associations involving the previously well-predicted outcome component than the less well-predicted outcome component. These results are important as they are the first to demonstrate that the associative past of an outcome can influence learning independently of the cue which it was first paired with.

These findings have been replicated by Thorwart, Livesey, Wilhelm, Liu and Lachnit (2017). In light of Griffiths et al., these authors sought to assess whether the predictability of an outcome influenced the extent to which an organism subsequently learns about it regardless of its functional role (i.e., whether the outcome goes on to serve as a cue or remain an outcome). These authors noted that if a comparable effect to learned predictiveness can be observed in outcomes (i.e., an “outcome predictability effect”), perhaps both learned predictiveness and outcome predictability are manifestations of the same phenomenon. According to this view, the predictive history of a stimulus is the crucial factor which determines how its associative past will influence learning, as opposed to the functional role of the stimulus (i.e., whether it serves as a cue or an outcome). If this is indeed the case then a previously well-predicted stimulus (be it a cue or an outcome in Stage 1), should be learned about better in Stage 2 regardless of whether its functional role switches between the two stages.

In order to address this question these authors conducted a series of experiments employing the food allergist task. In Experiment 1, the authors

replicated the standard learned predictiveness effect. In Experiment 2, the authors replicated the outcome predictability effect observed by Griffiths et al. using the same paradigm. Experiments 3 and 4, however, sought to assess whether the predictiveness of cues or the predictability of outcomes from Stage 1 would transfer to Stage 2 when these cues or outcomes served a different functional role in Stage 2. For instance, in Stage 1 of these experiments the cues were represented by foods and the outcomes by reactions. In Stage 2 of these experiments, however, the cues were now represented by reactions and the outcomes were represented by foods, thus Stage 2 took on the form of a diagnostic task. In both these experiments the effects demonstrated in Experiments 1 and 2 did not persist when the functional nature of the stimuli changed in Stage 2. These results suggest that the predictive history of stimuli alone is not sufficient in determining learning and that the functional nature of the stimulus must remain the same. However, the replication of the outcome **predictability** effect does constitute a challenge for standard associative learning models. Both Griffiths et al. and Thorwart et al. have demonstrated that when separating the outcome from the cue which it was first paired with, the outcome's associative past can influence subsequent learning about this outcome.

In light of the potential theoretical significance of the data reported by these authors, the experiments reported in this chapter explored the generality of this effect with a serial letter-prediction procedure in which participants were tasked with learning the predictive relationships between certain target letters. This experimental procedure has been successfully used in a range of studies of learned variations of cue processing (e.g. Evans, Gray & Snowden,

2007; Granger, Moran, Buckley & Haselgrove, 2016; Young, Kumari, Mehrotra, Hemsley, Andrew, Sharma, Williams & Gray, 2005) and permits measurements of learning derived from response times. This is important as it allows an online measure of performance, and therefore provides an insight into how learning is influenced by prior outcome predictability during the learning stages. As such, the aim of the current experiments was to assess whether the prior predictability of outcomes influenced the rate at which participants learned about these outcomes, when subsequently predicted by novel cues. If the associative history of an outcome produces an effect comparable to a learned predictiveness effect, it would be expected that a previously well-predicted outcome would be learned about more readily than a previously less well-predicted outcome. Such a finding would constitute a conceptual replication of the results reported by Griffiths et al. (2015) and Thorwart et al. (2017) and pose a major challenge to a central assumption of many associative models of learning.

2.1 Experiment 1

Experiment 1 comprised a serial letter-prediction task in which participants were required to make a response to certain target letters or ‘outcomes’. The design of Experiment 1 can be seen in Table 5.

Table 5. Design of Experiment 1

<i>Stage 1</i>			<i>Stage 2</i>		
P – X,	P – X,	P – X	P – X,	P – X,	P – X
F – Z,	G – Z,	W – Z	F – Z,	G – Z,	W – Z
F – H,	G – H,	W – H	H – X,	S – Z	
F – S,	G – S,	W – S	F – G,	G – F	
			F – W,	W – F	
			G – W,	W – G	

Note. X refers to the predicted outcome. Z refers to the unpredicted outcome. P refers to the cue which reliably predicts the outcome X. F/G/W refer to the cues which partially predict the less well-predicted outcome Z. H and S refer to the stimuli paired with F/G/W to form non-target trials in Stage 1, and the cues of the outcomes X and Z in Stage 2.

In Stage 1, letters X and Z were established as a well-predicted and a less well-predicted target stimulus, respectively (referred to henceforth as the *predicted* outcome, and the *unpredicted* outcome) and participants were asked to respond either upon their presentation or beforehand if they thought they could anticipate their presentation. To establish one letter (e.g. X) as a predicted outcome, this letter was only ever preceded by one other letter (e.g. P). To establish an outcome as being unpredictable (e.g. Z), three letters (F, G and W) preceded this outcome with equal frequency. In Stage 1 the letters F, G and W also preceded the letters S and H an equal number of times. These trials, in which F, G and W were paired with S and H, served three purposes. First, they ensured that the unpredicted outcome was in fact unpredictable. Without

these trials Z would merely have a greater number of predictors. Second, these trials ensured that S and H (which would serve as cues for X and Z in Stage 2) were familiarised to participants prior to the crucial test stage. And third, these trials served as non-target trials during which participants were not required to make a response. Stage 2 followed seamlessly from Stage 1. Here, trials in which P was paired with X, and trials in which F, G and W were paired with Z continued to be presented. In addition, however, H and S were now paired equally frequently with X and Z respectively. The question of central interest was whether participants would come to acquire a more rapid response time to H (which was paired with the previously predicted outcome, X) than to S (which was paired with the previously unpredicted outcome, Z). Additional non-target trials were also included in Stage 2 in which F, G and W preceded each other equally frequently.

2.1.1 Method

2.1.1.1 Participants

Thirty-two participants (18 females; 14 males) were recruited from the University of Nottingham's School of Psychology. Participants ranged from 19 – 51 years of age ($M = 23.8$; $SEM = 1.40$). All participants had normal or corrected-to-normal vision. Participants received course credit for their participation. The study was conducted in accordance with the British Psychological Society (2009) guidelines, receiving institutional ethical approval from the University of Nottingham's Psychology ethics committee.

2.1.1.2 Apparatus and Stimuli

All stimuli were presented (and responses recorded) in the graphical experimental software package PsychoPy2, v1.82.02 (see Peirce, 2007; Peirce 2008), running on Windows 7 on a standard desktop computer. The stimuli were white capital letters, Arial [7mm × 5mm; h × w], which appeared in the centre of the screen [27cm × 46cm; h × w] against a grey background. The stimulus letters were, ‘F’, ‘G’, ‘H’, ‘P’, ‘S’ and ‘W’. The target stimuli [i.e. the outcomes] were ‘X’ and ‘Z’. The actual stimulus letters which served as the cues, outcomes, and non-target trials were counterbalanced across experimental versions using a Latin-Squared counterbalancing technique. Stimulus letters never appeared simultaneously on screen.

2.1.1.3 Procedure

All participants were tested individually. Once participants had provided their consent, they read the following instructions on screen before being exposed to Stage 1:

“Thank you for participating in this experiment.

In this experiment you will see individual letters appear in the centre of the screen.

It is your job to press X when you see X appear and press Z when you see Z appear. At first you will only be able to respond to these letters when you see them, but as the experiment continues, you might be able to anticipate when they are going to be presented. If you think you know when either X or Z are going to appear, you can press them BEFORE they are presented.

Please try to respond as quickly as you can when you think you know when X or Z are going to appear. If you have no questions, please have your

fingers ready over the X and the Z, and then press the space bar to begin the experiment.”

Stage 1

In Stage 1, participants were exposed to 6 blocks of 12 trials (72 trials in total) in which a trial constituted a pairing of two stimuli. Trial order was block randomised, and there was no break between blocks. In each block participants received 3 pairings of P and X, 1 pairing each of F, G, and W with Z; and 1 pairing each of F, G, and W with S and 1 pairing each of F, G and W with H. Each trial lasted 4 seconds. On a target trial, for example, a cue would be presented (e.g. P) for 1 second, followed by an inter-stimulus interval of 1 second, an outcome (e.g. X) of 1 second, and finally a 1 second interval before the next cue was presented (see Figure 2). In total, Stage 1 lasted 4.8 min. Once all 72 trials were complete, Stage 2 began. There was no break between the two stages.

Stage 2

In Stage 2 participants were exposed to 18 blocks of 16 trials, the order of which was randomised in each block. In this stage the trials which established the predictability of the outcomes were carried over from Stage 1 in order to make the two stages more similar. Thus, in each block participants received 3 pairings of P and X, and 1 pairing each of F, G, and W with Z. In addition, H and S were each paired with the predicted and unpredicted outcomes (e.g. X and Z) respectively, for two trials; and F, G and W were paired with each other to form 6 non-target trials for each block. In total Stage 2 lasted 19.2 min.

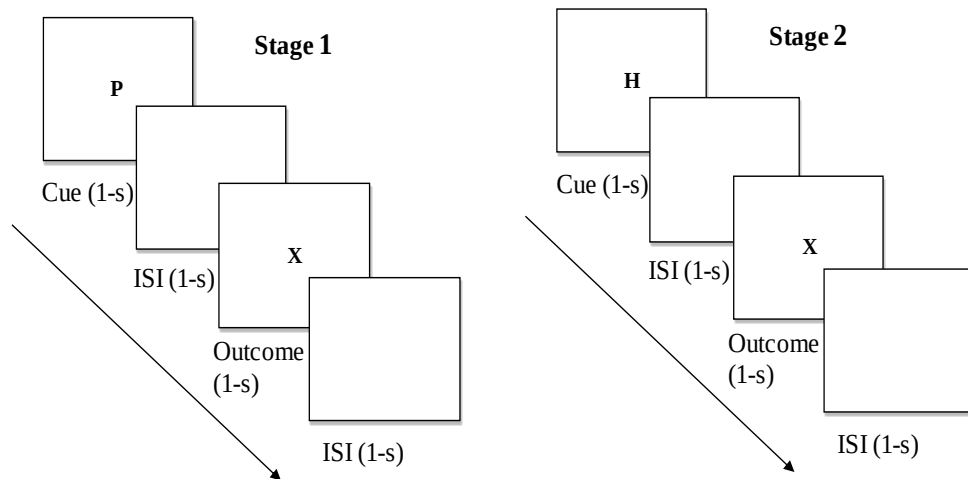


Figure 2. Example (target) trials for both Stage 1 and Stage 2. Each trial lasted 4 seconds. A stimulus would first be presented, in this case a cue, followed by an inter-stimulus interval, an outcome, and finally a second interval. Participants were required to respond to the outcomes (X or Z) by pressing the respective keys, either when the target outcome appeared on screen or when they thought it would appear, if they could anticipate its onset.

2.1.1.4 Scoring and Error Classification

Response times could range from 0 – 3 seconds. Responses made between 0 – 1 seconds indicated that participants were responding during the presentation of the cue letter (e.g. the letter P in a P–X pairing). Responses between 1 – 2 seconds indicated that participants were responding during the inter-stimulus interval following the presentation of the cue. Responses between 2–3 seconds indicated that participants were responding to the outcomes as they were presented (X or Z). Response times were only analysed for P-X and F/G/W-Z trials in Stage 1, and on H-X and S-Z trials in Stage 2. If participants omitted a response (i.e. they failed to press the respective outcome key [e.g. X or Z] within the 0–3 second window) they were allocated a response time of 3 seconds for the trial. This type of error was classed as an

‘omission’. If participants made an erroneous prediction (e.g. by selecting the letter X when in fact Z was the outcome), they would also be allocated a response time of 3 seconds. This type of error was classed as a ‘commission’.

2.1.2 Results and Discussion

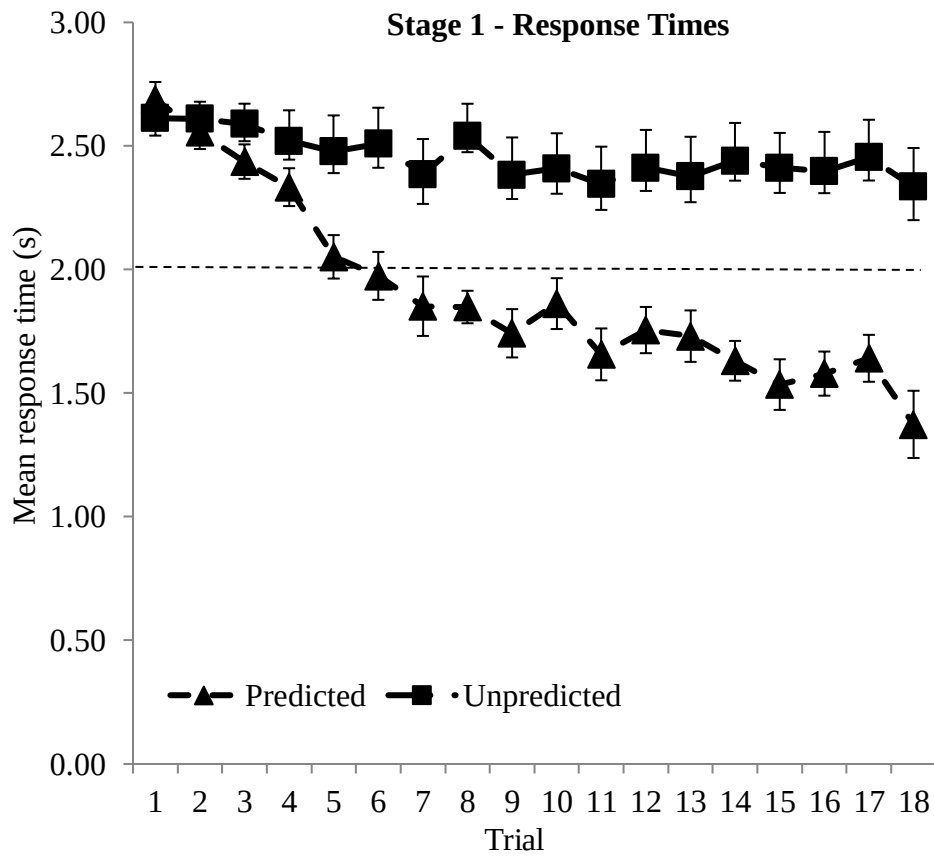


Figure 3. Mean response times to the predicted (X) and unpredicted (Z) outcomes in Stage 1. Response times within the 2 – 3 second window indicate participants were responding to the outcomes as they were presented. If participants made a response before 2 seconds (signalled by the dotted line), they were responding prior to the presentation of the outcome. Errors bars show SEM.

2.1.2.1 Response Time Data

Figure 3 shows the mean response time to the predicted and unpredicted outcomes in Stage 1 across the 18 trials. The predicted outcome refers to the outcome (e.g. X) which was reliably paired with one cue (e.g. P); the unpredicted outcome refers to the outcome (e.g. Z) which was preceded by

three letters (F/G/W). The response times for these data were collated by looking at responses to the target stimuli within the 0-3 second window beginning immediately upon the presentation of the cue that preceded the targets. Inspection of the data reveals that as training progressed, response times for the predicted outcome shortened. On average, response times to the unpredicted outcome were greater than or equal to 2 seconds, whilst response times to the predicted outcome were (with training) less than 2 seconds. For instance, from Trial 6 onwards participants were, on average, responding to the outcome prior to its presentation. In contrast, response times to the unpredicted outcome stayed relatively constant across all 18 trials, with participants, on average, responding to the outcome only during its presentation.

This impression was confirmed with a 2 x 18 repeated measures analysis of variance (ANOVA) of individual response times with the variables of outcome (Unpredicted vs Predicted) and trial (1 – 18); in light of sphericity violations for the “trial” variable, $\chi^2 (152) = 259.55, p < .001$ and the interaction between trial and outcome, $\chi^2 (152) = 215.50, p < .001$, Greenhouse-Geisser values are reported.⁵ The ANOVA revealed significant main effects of outcome, $F (1, 31) = 20.71, MSE = 4.26, p < .001, \eta_p^2 = .40$, and trial, $F (8, 253) = 13.48, MSE = .50, p < .001, \eta_p^2 = .30$, and a significant interaction between the two variables, $F (8, 259) = 6.12; MSE = .47, p < .001, \eta_p^2 = .17$. Simple main-effects analyses revealed no significant differences between the outcomes for trials 1 - 4, smallest, $F (1, 31) = 2.27$,

⁵ Greenhouse-Geisser values are reported for all appropriate analyses where epsilon is less than .75 (see Field, 2009). However, for ease of readability, corrected degrees of freedom are rounded to nearest whole number.

$MSE = .04$, $p = .14$, but a significant difference for all other trials, smallest F (1, 31) = 5.68, $MSE = .51$, $p = .02$.

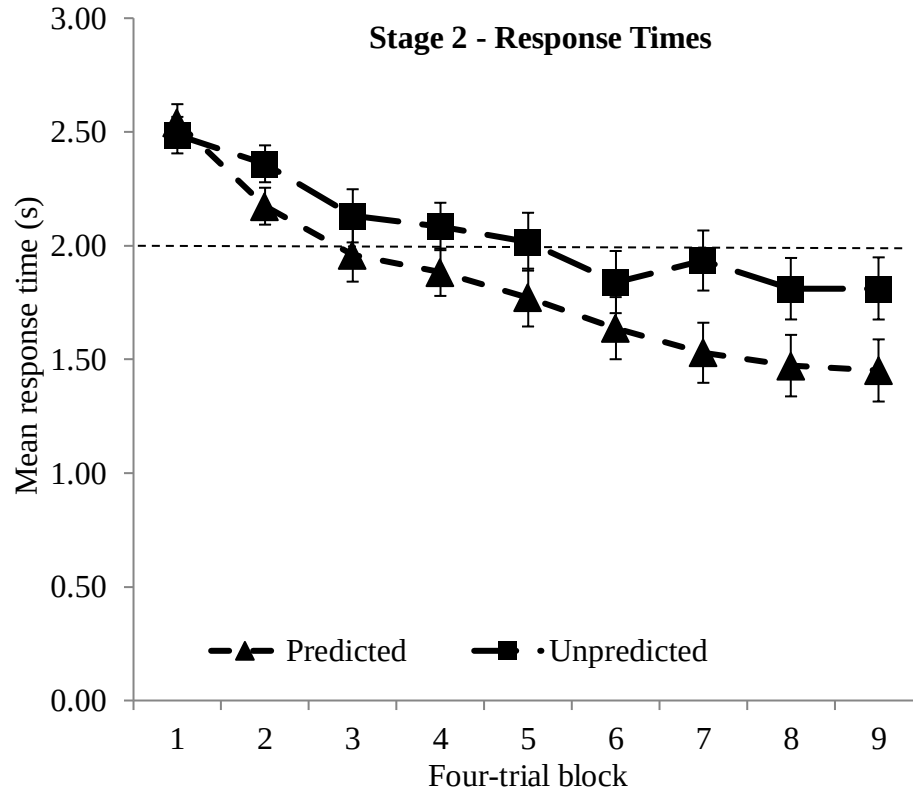


Figure 4. Mean response times to the predicted (X) and unpredicted (Z) outcomes in Stage 2. Response times within the 2 – 3 second window indicate participants were responding to the outcomes as they were presented. If participants made a response before 2 seconds (signalled by the dotted line), they were responding prior to the presentation of the outcome. Errors bars show SEM.

Figure 4 shows the mean response times to the predicted and unpredicted outcomes in Stage 2, on trials when the outcomes were reliably, and equivalently, predicted by cues H and S. The data in Figure 4 have been collapsed into 9 blocks, each comprising 4 trials. In keeping with Stage 1, the response time data for these analyses were collated by looking at responses within the 0-3 second window following the onset of cue H or S. Figure 4

demonstrates that despite both outcomes initially producing similar mean response times on the first block, response times to the previously predicted outcome became faster than the previously unpredicted outcome. These impressions were confirmed with a 2 x 9 repeated measures ANOVA of individual response times with the variables of outcome (unpredicted vs predicted) and block (1 – 9), which revealed significant main effects of outcome, $F(1, 31) = 4.86$, $MSE = 1.54$, $p < .05$, $\eta_p^2 = .14$, and block, $F(4, 105) = 30.60$, $MSE = .46$, $p < .001$, $\eta_p^2 = .50$. The interaction between these variables was not significant, $F(4, 115) = 2.05$, $MSE = .31$, $p = .09$, $\eta_p^2 = .06$. Overall, these results demonstrate that the previously predicted outcome was learned about at a faster rate than the unpredicted outcome.

2.1.2.2 Error Data

The mean percentage of errors for the predicted and the unpredicted outcome was calculated for both Stage 1 and Stage 2. The Stage 1 error counts were calculated by analysing the number of errors committed on the P–X trials for the predicted outcome and the F/G/W–Z trials for the unpredicted outcome. Inspection of these data revealed that a relatively low percentage of errors were committed for both the predicted outcome ($M = 3.81\%$, $SEM = 1.65$) and the unpredicted outcome ($M = 3.65\%$, $SEM = 1.05$). Participants also made comparable percentages of omissions ($M = 3.21\%$, $SEM = .88$) and commissions ($M = 4.25\%$, $SEM = 1.82$). A 2×2 repeated measures ANOVA of individual percent errors with the variables of outcome (predicted vs unpredicted) and error type (omission vs. commission) revealed no main effect of outcome, $F(1, 31) = .01$, $MSE = 70.66$, $p = .91$, $\eta_p^2 = .00$, no effect of Error

Type, $F(1, 31) = .46$, $MSE = 75.54$, $p = .50$, $\eta_p^2 = .02$, and no interaction between these two variables, $F(1, 31) = .88$, $MSE = 69.69$, $p = .35$, $\eta_p^2 = .03$.

The Stage 2 error counts were calculated by analysing the number of errors committed on the H – X and S – Z trials. Inspection of these data revealed that participants, again, made relatively few errors for both the predicted ($M = 5.56\%$, $SEM = 1.87$) and unpredicted outcome ($M = 4.03\%$, $SEM = 1.12$) in this stage. As in Stage 1, slightly more commissions ($M = 5.60\%$, $SEM = 2.09$) were committed than omissions ($M = 3.99\%$, $SEM = .89$). An identical ANOVA to that conducted upon the percentage error data from Stage 1 was conducted upon the data from Stage 2. This analysis revealed no main effect of outcome, $F(1, 31) = 1.77$, $MSE = 41.82$, $p = .19$, $\eta_p^2 = .05$, no effect of error type, $F(1, 31) = .55$, $MSE = 151.04$, $p = .47$, $\eta_p^2 = .02$, and no interaction between these two variables, $F(1, 31) = 1.22$, $MSE = 35.87$, $p = .28$, $\eta_p^2 = .04$.

Experiment 1 demonstrated that establishing outcomes as differentially predicted, altered the extent to which these outcomes subsequently supported learning to new cues. Following learning about P–X and F/G/W–Z in Stage 1, participants displayed faster overall learning to the predicted outcome (H–X) in Stage 2 than the unpredicted outcome (S–Z). These results pose a challenge to theories of learning which do not grant an independence of processing to outcomes (Esber & Haselgrove, 2011; Le Pelley, 2004; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972). These models assume that when a previously learned-about outcome is separated from the cue which established the predictability of the outcome, the associative history of the outcome should not influence subsequent learning.

However, the results of Experiment 1 appear to contradict this assumption and conceptually replicate the findings from Griffiths et al. (2015) and Thorwart et al. (2017).

2.2 Experiment 2

Although the results of Experiment 1 are consistent with the suggestion that outcomes can come to acquire differential associability as a consequence of being either well or poorly predicted in the past, it is possible to develop an alternative explanation for these results from the perspective of standard associative models of learning. During Stage 1, participants were given trials in which F, G and W were each paired with Z, which could result in the establishment of excitatory associations between each of these stimuli and Z. On other trials, F, G and W, are paired with H, and on other trials with S. Consequently, on these F/G/W-H or F/G/W-S trials there may have been an expectation of the presentation of Z which, importantly, was always unfulfilled. These circumstances are precisely those which standard models of associative learning (e.g. Rescorla & Wagner, 1972) predict will result in the establishment of an inhibitory association between H and Z, and S and Z. This inhibition will then have to be overcome on trials in which S and Z are paired in Stage 2, before the acquisition of excitatory associative learning between these stimuli can be expressed, thus the S-Z training will appear to be less successful than the H-X training, which was precisely the result observed in Experiment 1. Experiment 2 sought to rule out this alternative explanation for the results of Experiment 1, by arranging the stimuli in a manner which would

circumvent any differential inhibitory association being formed between the stimuli which served as the test cues in Stage 2 (H and S) and the unpredicted outcome during Stage 1.

The strategy in Experiment 2 was to therefore employ a design where H and S were never preceded by F, G or W. Thus, there was no opportunity for an unfulfilled expectation of Z to be present during trials featuring H and S, and consequently the establishment of inhibition. However, to ensure participants were still familiar with H and S before they were established as cues for Z and X, trials in which these stimuli were paired with each other were presented in Stage 1 (i.e.: S–H and H–S). As in Experiment 1, non-target trials were again formed by pairing the F, G and W stimuli with each other (as in Stage 2 of Experiment 1). The full design of Experiment 2 can be seen in Table 6.

Table 6. Design of Experiment 2

<i>Stage 1</i>			<i>Stage 2</i>		
P – X,	P – X,	P – X	P – X,	P – X,	P – X
F – Z,	G – Z,	W – Z	F – Z,	G – Z,	W – Z
H – S,	S – H		H – X,	S – Z	
F – G,	G – F		F – G,	G – F	
F – W,	W – F		F – W,	W – F	
G – W,	W – G		G – W,	W – G	

Note. X refers to the predicted outcome. Z refers to the unpredicted outcome. P refers to the cue which reliably predicts the outcome X. F/G/W refer to the cues which partially predict the less well-predicted outcome Z, and also serve as non-target trials when paired with each other. H and S refer to the cues of the outcomes X and Z in Stage 2.

If the effect observed in Experiment 1 is still present using this design, this would suggest that any impaired performance involving the unpredicted outcome in Stage 2 was not solely underpinned by an inhibitory association

between the previously unpredicted outcome (e.g. Z) and its novel cue (e.g. S) during Stage 1 training. In addition, following completion of Stage 2, a set of rating scales were presented which required participants to rate how predictive each stimulus was of the target stimuli (X and Z). If these scales are consistent with the response time data in Experiment 1, it would be expected that the cue (S) which was paired with the previously unpredicted outcome (Z), would be rated as being less predictive than the cue (H) which was paired with the previously predicted outcome (X).

2.2.1 Method

2.2.1.1 Participants

Thirty-two participants (29 females; 3 males) were recruited from the University of Nottingham's School of Psychology. Participants ranged from 18 – 46 years of age ($M = 20.2$; $SEM = 0.89$). Participants received course credit for their participation. A cash prize of £10 was also awarded to the participant who completed the experiment with the fastest overall reaction time on trials where the target stimuli were present. All other details were the same as in Experiment 1.

2.2.1.2 Apparatus and Stimuli

The apparatus and stimuli for Experiment 2 were identical to Experiment 1. However once Stage 2 had been completed participants were presented with two ratings screens, one for each of the target stimuli (X and Z). These screens requested participants provide a rating for each stimulus's (i.e. 'F' 'G' 'H' 'P' 'S' and 'W') ability to predict the absence or presence of the

target stimuli on a scale ranging from ‘-100’ to ‘+100’ [‘-100’ = ‘Predicts absence of target stimulus’; ‘0’ = ‘Don’t Know’ ‘+100’ = ‘Predicts presence of target stimulus’]. Six rating scales appeared on the same screen at the same time, for each rating screen. Each of the six scales was placed alongside each of the respective letters.

2.2.1.3 Procedure

The procedure for Experiment 2 was identical to Experiment 1, with the exception that the letters (H and S) which would serve as cues for X and Z in Stage 2 were paired with each other in Stage 1 twice per block (i.e. H–S, S–H). Thus, these pairings were presented a total of 12 times across Stage 1. In Stage 1, the stimuli F, G and W, were also paired with each other to form non-target trials (e.g. F–G, F–W, G–F, G–W, W–F and W–G), as in Stage 2 of Experiment 1. These pairings were each presented once in each block, thus being presented a total of 36 times across Stage 1. In total Stage 1 lasted 5.6 mins. The procedure for Stage 2 was identical to Stage 2 of Experiment 1.

Once participants had completed Stage 2 they were presented with the following text: *“Thank you. The experiment is nearly complete. Before finishing, however, there are a few questions to complete. Please press the ‘SPACEBAR’ to proceed.”* Participants were then presented with a rating screen for each of the target stimuli (X and Z). Participants then provided a score ranging from -100 to + 100 regarding how predictive of the absence or presence of X or Z each other letter was. Participants provided their score by adjusting a tick maker on a scale. The numerical value of the tick marker position was provided in a grey box beneath the scale. The order of the rating screens was counterbalanced between participants.

2.2.2 Results and Discussion

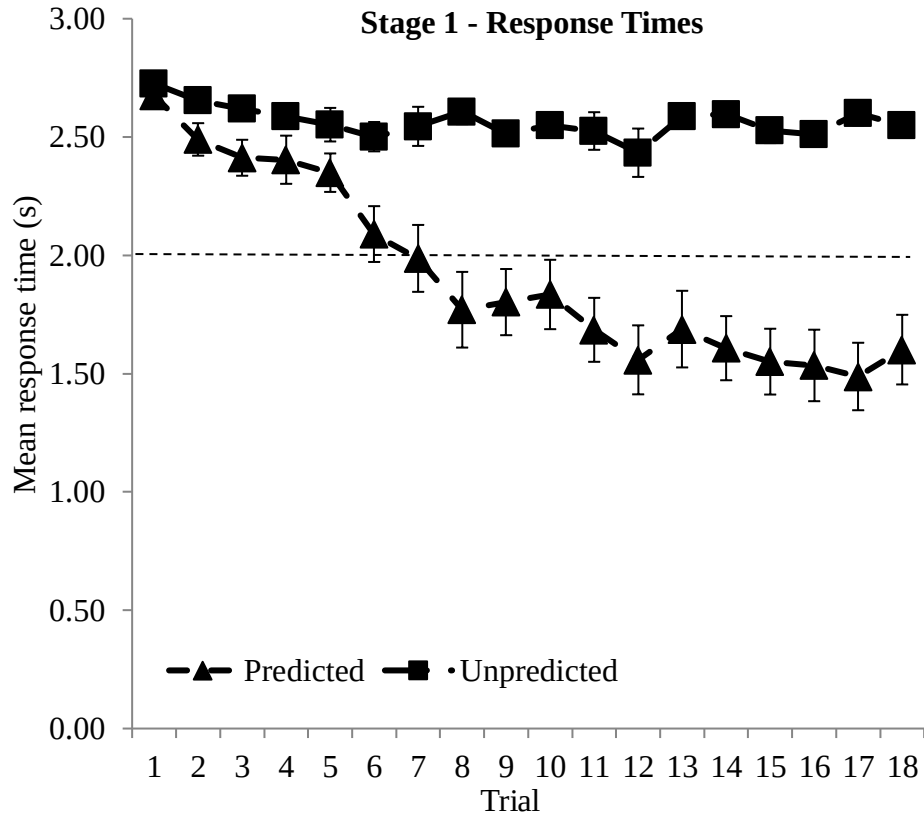


Figure 5. Mean response times to the predicted (X) and unpredicted (Z) outcomes in Stage 1. Response times within the 2 – 3 second window indicate participants were responding to the outcomes as they were presented. If participants made a response before 2 seconds (signalled by the dotted line), they were responding prior to the presentation of the outcome. Errors bars show SEM.

2.2.2.1 Response Time Data

Figure 5 shows the mean response time to the predicted and unpredicted outcomes in Stage 1 across the 18 trials. In keeping with Experiment 1, the response times for these data were collated by looking at responses within the 0 – 3 second window beginning with the presentation of the cue. Inspection of the data presented in Figure 5 reveals that as training

progressed, response times for the predicted outcome shortened. Furthermore, Figure 5 demonstrates that from Trial 7 onward participants were, on average, responding to the predicted outcome before its presentation. In contrast, response times to the unpredicted outcome stayed relatively constant across all 18 trials, with participants, on average, responding to the outcome only when the outcome had been presented.

A 2 x 18 repeated measures ANOVA was performed on these data with the variables of outcome (Unpredicted vs Predicted) and trial (1 – 18). The ANOVA revealed significant main effects of outcome, $F(1, 31) = 55.23$, $MSE = 2.19$, $p < .001$, $\eta_p^2 = .64$, and trial, $F(8, 247) = 15.5$, $MSE = .40$, $p < .001$, $\eta_p^2 = .34$, and a significant interaction between these variables, $F(8, 242) = 10.74$, $MSE = .40$, $p < .001$, $\eta_p^2 = .26$. Simple main-effects analyses revealed significant differences between the outcomes ($ps < .05$) from Trial 2 onward, excluding Trial 4, $F(1, 31) = 3.45$, $p = .07$ and Trial 5, $F(1, 31) = 3.63$, $p = .06$.

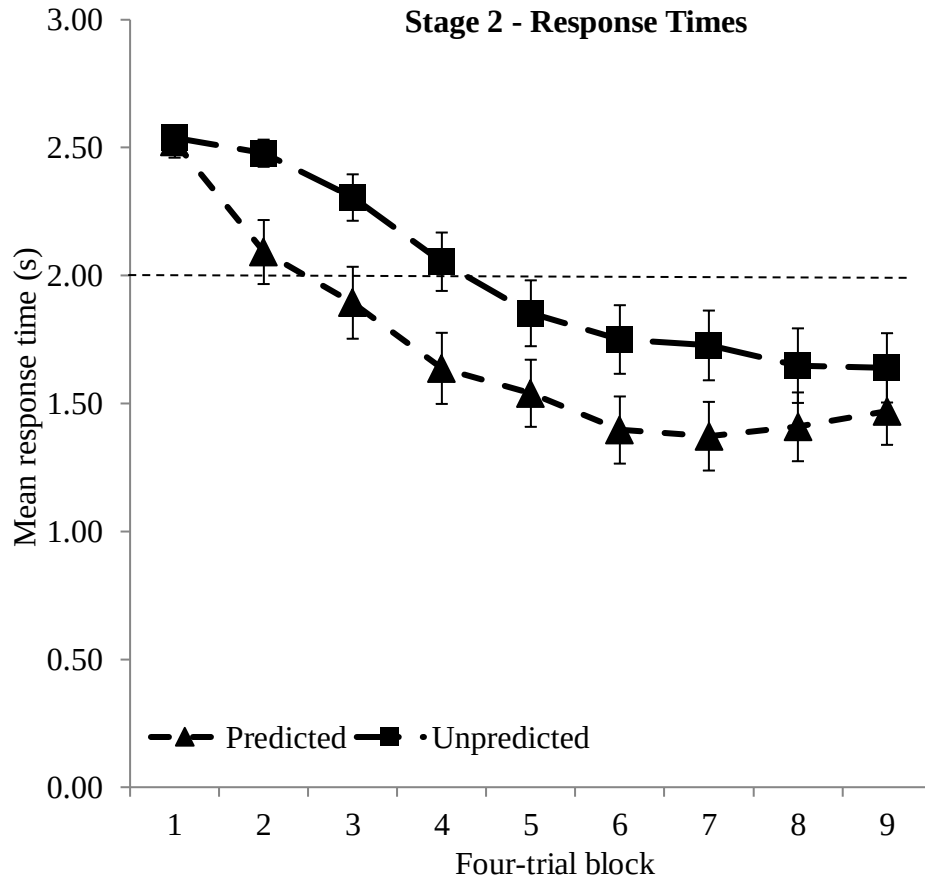


Figure 6. Mean response times to the predicted (X) and unpredicted (Z) Outcomes in Stage 2. Response times within the 2 – 3 second window indicate participants were responding to the outcomes as they were presented. If participants made a response before 2 seconds (signalled by the dotted line), they were responding prior to the presentation of the outcome. Errors bars show SEM.

Figure 6 shows, over 9 four-trial blocks, the mean response times to the predicted and unpredicted outcomes in Stage 2, on trials when the outcomes from Stage 1 were both reliably and equivalently predicted by cues H and S. In keeping with Experiment 1, response times were similar for both the previously predicted and unpredicted outcomes at the outset of Stage 2. However, overall, response times to the previously predicted outcome from Stage 1 were faster than the unpredicted outcome from Stage 1.

A 2 x 9 repeated measures ANOVA was performed on these data with the variables of outcome (Unpredicted vs Predicted) and block (1 – 9). The ANOVA revealed significant main effects of outcome, $F(1, 31) = 8.08$, $MSE = 1.56$, $p < .01$, $\eta_p^2 = .21$, and block, $F(3, 92) = 41.54$, $MSE = .57$, $p < .001$, $\eta_p^2 = .57$. The interaction effect between the two factors approached significance, $F(3, 106) = 2.34$, $MSE = .28$, $p = .07$, $\eta_p^2 = .07$. For the curious reader, simple main effects analyses were conducted to identify the source of the interaction which approached significance. These analyses revealed significant differences between the outcomes ($ps < .05$) on Blocks 2 - 7. There were no significant differences on blocks 1, 8 or 9, largest $F(1, 31) = 2.79$, $p = .11$.

2.2.2.2 Error Data

As in Experiment 1, the mean percentage of errors for the predicted and the unpredicted outcome was calculated for both Stage 1 and Stage 2 data. Inspection of the data from Stage 1 revealed that participants made more errors for the unpredicted outcome ($M = 5.82$, $SEM = 1.31$) than the predicted outcome ($M = 3.56$, $SEM = 1.16$) and more commissions ($M = 6.16$, $SEM = 1.49$) than omissions ($M = 3.21$, $SEM = .99$). These impressions were largely confirmed with a 2×2 repeated measures ANOVA of percentage of errors with the variables of outcome (unpredicted vs predicted) and error type (commission vs omission). The ANOVA revealed a main effect of outcome $F(1, 31) = 6.2$, $MSE = 24.61$, $p = .02$, $\eta_p^2 = .18$. The main effect of error type approached significance, $F(1, 31) = 3.42$, $MSE = 81.62$, $p = .07$, $\eta_p^2 = .10$.

There was no significant interaction between these two factors, $F(1, 31) = .118$; $MSE = 32.73$, $p = .73$, $\eta_p^2 = .00$.

Inspection of the Stage 2 data revealed that participants made a comparable number of errors for the predicted outcome ($M = 5.43$, $SEM = 1.45$) and the unpredicted outcome ($M = 4.69$, $SEM = 1.68$), with more commissions ($M = 6.90$, $SEM = 1.62$) than omissions ($M = 3.21$, $SEM = 1.51$). An identical ANOVA to that conducted on the error data from Stage 1 was conducted on the Stage 2 data which revealed a main effect of error type, $F(1, 31) = 4.83$, $MSE = 90.17$, $p = .04$, $\eta_p^2 = .14$, but no effect of outcome $F(1, 31) = 1.31$, $MSE = 13.31$, $p = .26$, $\eta_p^2 = .04$, and no interaction, $F(1, 31) = .08$, $MSE = 61.75$, $p = .78$, $\eta_p^2 = .00$.

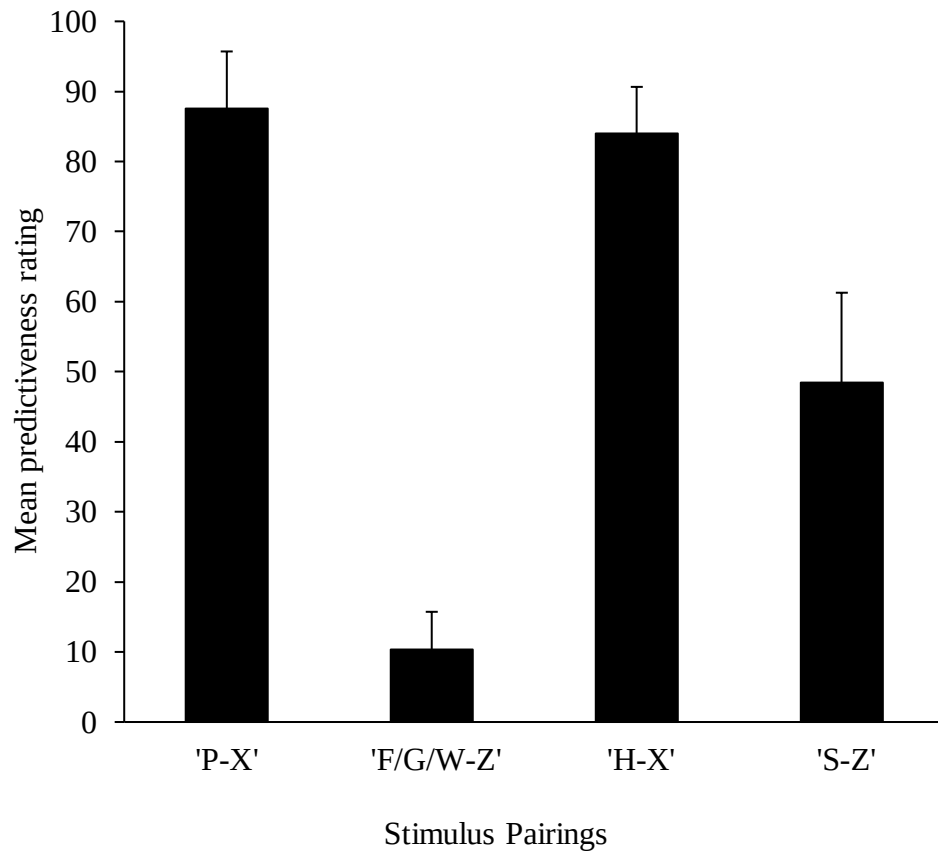


Figure 7. Mean predictiveness rating for the cues at the end of Stage 2. Errors bars show SEM.

2.2.2.3 Ratings data

Figure 7 shows participants' ratings of how predictive they deemed each of the cues to be for the target outcomes (X or Z). The ratings for the three stimuli paired with the unpredicted outcome in Stage 1 (e.g. 'F', 'G', 'W') were collapsed by averaging the values for these stimuli. Four participants failed to provide ratings, either omitting some or all of the rating scales; an additional participant was removed as their ratings were more than 3 standard deviations above the mean (Z score > 3.27); closer inspection of their ratings indicated they used the scale in the opposite manner in which it was

intended. Thus, a total of 5 participants were omitted from the subsequent analyses ($n = 27$).

As can be seen in Figure 7, the ‘P–X’ ratings and ‘H–X’ ratings were comparably high, and both higher than the ‘S–Z’ rating. The three stimuli (‘F/G/W’) paired with the unpredicted outcome (‘Z’) received only a mildly positive rating. A repeated measures one-way ANOVA of individual ratings with four levels (P–X, F/G/W–Z, H–X, and S–Z) was significant, $F(2, 56) = 41.62$, $MSE = 2074.94$, $p < .001$, $\eta_p^2 = .62$. Planned comparisons using Bonferonni corrected t-tests (adjusted $p = .025$) revealed a significant difference between the predictiveness ratings for the stimulus pairings P–X and F/G/W – Z, $t(26) = 9.01$, $p < .001$, and a significant difference between H–X and S–Z, $t(26) = 2.86$, $p < .01$. These results corroborate the response time data from Stage 2: participants were biased towards learning about the cue paired with the previously predicted outcome (H–X) relative to the cue paired with the previously unpredicted outcome (S–Z).

Experiment 2 replicated and extended the findings of Experiment 1, again demonstrating that the previous predictability of an outcome influences the rate at which it is subsequently learned about when signaled by a cue which it has not previously been paired with. These results were obtained using an experimental design that ruled out the contribution of inhibitory learning that may be employed to explain the results of Experiment 1. In keeping with Experiment 1, the Stage 2 error data indicates the difference in the response time data is not undermined by a speed-accuracy trade off as both outcomes produced a comparable (and relatively low) number of errors, thus indicating that whilst participants are able to learn the association involving the

previously unpredicted outcome, the previously predicted outcome enters into associations more readily. The rating data corroborate this. Although both cues for the outcomes from Stage 2 received positive ratings, participants provided higher ratings to the cue paired with the previously predicted outcome than the cue that was paired with the previously unpredicted outcome.

Together the results of Experiments 1 and 2 appear to provide evidence that an effect comparable to learned predictiveness can be observed in outcomes. That is, the associative history of an outcome can bias subsequent learning about an outcome, with the bias occurring in favour of a previously well-predicted outcome. However, in Experiments 1 and 2 of this chapter, the Stage 1 training trials which rendered the outcomes as predictable or unpredictable were carried over to Stage 2, which is different to learned predictiveness experiments. In a standard learned predictiveness experiment (i.e., Le Pelley & McLaren, 2003), the Stage 1 training trials are omitted during Stage 2 training. In light of this discrepancy between the two experimental paradigms, Experiment 3 sought to assess whether the presence of Stage 1 target trials, during Stage 2, were necessary for the effect to be observed.

2.3 Experiment 3

The purpose of Experiment 3 was to assess the extent to which the target trials which first established the predictability of the outcomes in Stage 1 were necessary in Stage 2. In Experiments 1 and 2, the trials in Stage 1 that established X and Z as differentially predictive (i.e., P – X, G/F/W – Z) were

carried over to Stage 2. The inclusion of these trials in Stage 2, however, is at odds with similar procedures which have revealed comparable effects in cues (i.e., learned predictiveness experiments). In a learned predictiveness experiment (Le Pelley & McLaren, 2003), Stage 1 training provides participants with training which renders some cues as good predictors and some cues as poor predictors of certain outcomes. In Stage 2 these cues are again presented this time being paired with novel outcomes. This time, however, during Stage 2 training, the training trials from Stage 1 which established the cues as reliable or unreliable predictors are omitted. As such, there is a discrepancy between the design of Experiments 1 and 2 in this chapter and the design typically employed in a learned predictiveness experiment. This experiment therefore sought to assess whether the effect observed in Experiments 1 and 2 is preserved when the target trials in Stage 1 are omitted from Stage 2 and the design is more comparable to a learned predictiveness experiment. As such, the design of Experiment 3 was the same as Experiment 2, however, this time in Stage 2 the trials which established the predictability of the outcomes in Stage 1 were omitted (P-X, G/F/W/ - Z). The design of Experiment 3 can be seen in Table 7. If the target trials from Stage 1 are necessary in order to produce the effect, it would be expected that the omission of the target trials from Stage 2, would result in no difference between the outcomes. If the trials are not necessary in order to produce the effect, their omission should have little bearing on the outcome predictability effect observed in Experiments 1 and 2.

Table 7. Design of Experiment 3.

<i>Stage 1</i>			<i>Stage 2</i>	
P – X,	P – X,	P – X	-	
F – Z,	G – Z,	W – Z	-	
H – S,	S – H		H – X,	S – Z
F – G,	G – F		F – G,	G – F
F – W,	W – F		F – W,	W – F
G – W,	W – G		G – W,	W – G

Note. X refers to the predicted outcome. Z refers to the unpredicted outcome. P refers to the cue which reliably predicts the outcome X. F/G/W refer to the cues which partially predict the less well-predicted outcome Z, and also serve as non-target trials when paired with each other. H and S refer to the cues of the outcomes X and Z in Stage 2.

2.3.1 Method

2.3.1.1 Participants

Thirty-two participants (26 females; 6 males) were recruited from the University of Nottingham’s School of Psychology. Participants ranged from 18 – 29 years of age ($M = 19.78$; $SEM = 0.39$). Participants received course credit for their participation. A cash prize of £10 was awarded to the participant who completed the experiment with the fastest overall reaction time on trials where the target stimuli (‘X’ or ‘Z’) were present.

2.3.1.2 Apparatus and Stimuli

As in the previous experiments the stimuli were presented (and responses recorded) in the graphical experimental software package PsychoPy2, v1.82.02 (see [Peirce, 2007](#); [Peirce 2009](#)), running on Windows 7 on a standard desktop computer. The stimuli components were the same as

Experiment 1 (See Experiment 1, Method). The only difference between the two experiments was the omission of the Stage 1 target training trials in Stage 2.

2.3.1.3 Procedure

The procedure for Experiment 3 was identical to Experiment 2, except that in this experiment participants did not receive the Stage 1 target trials in Stage 2, thus in total Stage 2 comprised 18 blocks of 10 trials (again trials were block randomised). Each block was identical (i.e. the same stimuli pairings were presented). This stage lasted for 9 mins. The H – X and S – Z pairings were presented a total of 36 times. The filler trials from Stage 1 continued to ensure that the two stages were comparable. In total there were 108 filler trials.

2.3.2 Results and Discussion

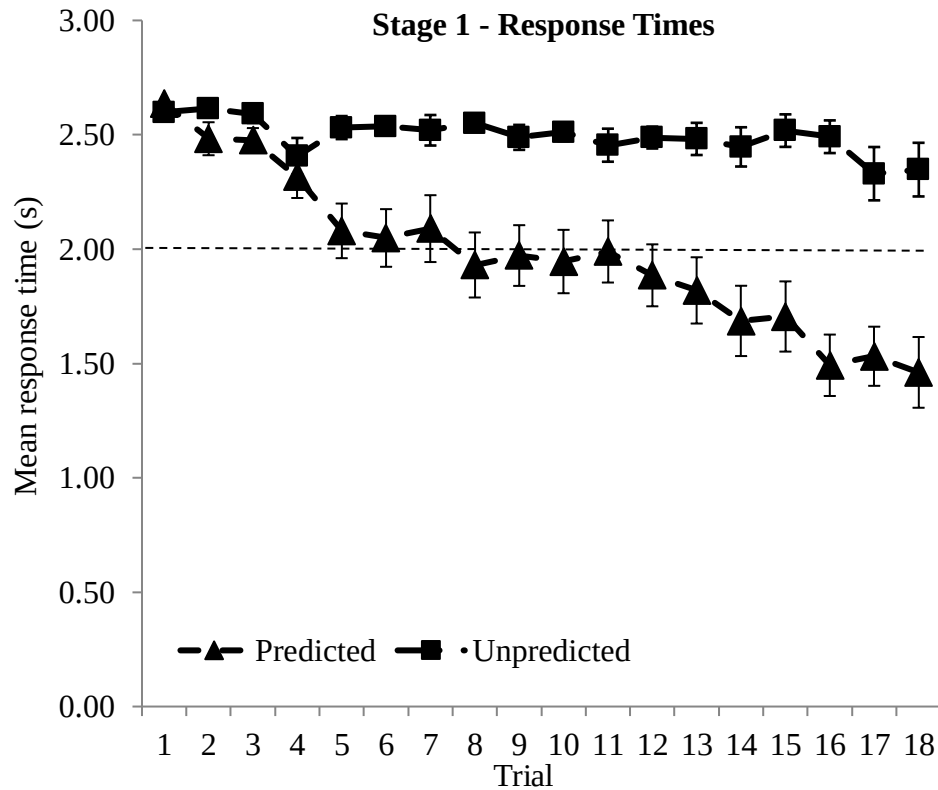


Figure 8. Mean response times to the predicted (X) and unpredicted (Z) outcomes in Stage 1. Response times within the 2 – 3 second window indicate participants were responding to the outcomes as they were presented. If participants made a response before 2 seconds (signalled by the dotted line), they were responding prior to the presentation of the outcome. Errors bars show SEM.

2.3.2.1 Stage 1 – Response Time

Figure 8 shows the average response time to the predicted and unpredicted outcomes in Stage 1 across the 18 trials. Figure 8 demonstrates that as a function of training response times for the predicted outcome decreased. Moreover, from Trial 7 onward participants were correctly predicting the outcome prior to its onset. However, response times to the

unpredicted outcome remained relatively constant across all trials, with participants unable to predict the onset of this outcome.

A 2 x 18 repeated measures ANOVA was performed on these data. The two factors were 'Outcome' ['Unpredicted' and 'Predicted'] and 'Trials' [Trials 1 - 18]. The ANOVA revealed significant main effects for the factors 'Outcome', $F(1, 31) = 40.25$, $MSE = 1.93$, $p < .001$, $\eta_p^2 = .57$, and 'Trials' $F(17, 218.4) = 11.13$, $MSE = .55$, $p < .001$, $\eta_p^2 = .26$, and a significant interaction between the two factors, $F(8.38, 259.85) = 8.13$; $MSE = .34$, $p < .001$, $\eta_p^2 = .20$. Simple main effects analysis revealed significant differences between the outcomes, $p < .01$ for Trials 5 – 18. These data reveal that the two outcomes were learned about differently as a result of their predictability, with the predicted outcome producing faster response times than the unpredicted outcome. The effect increased as a function of the number of trials.

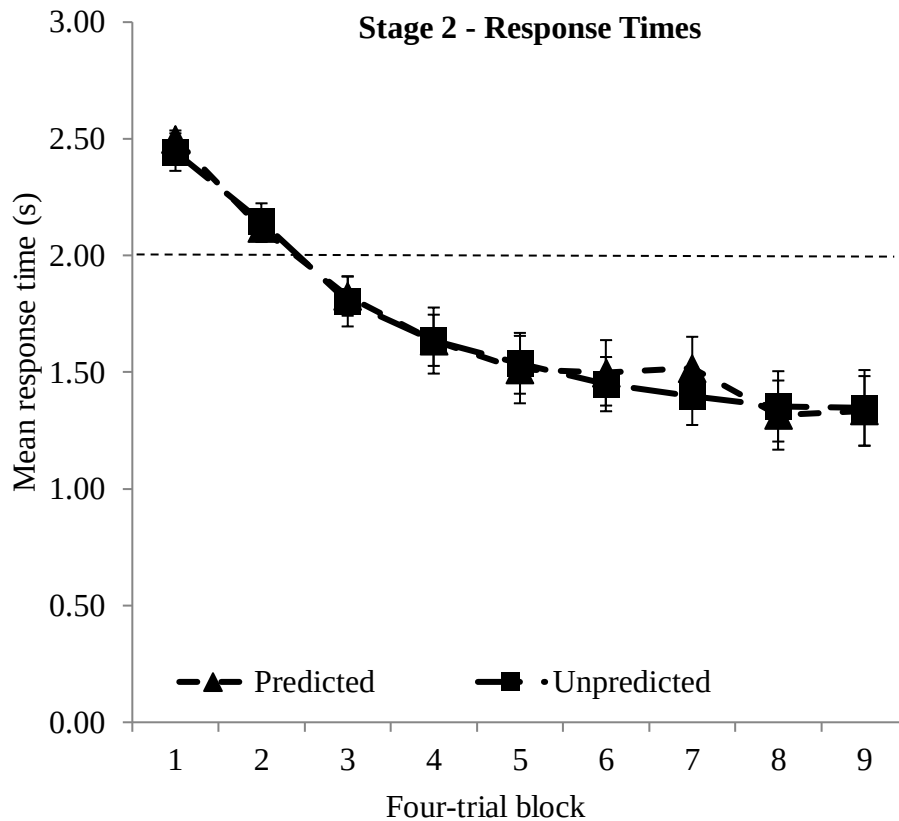


Figure 9. Mean response times to the predicted (X) and unpredicted (Z) outcomes in Stage 2. Response times within the 2 – 3 second window indicate participants were responding to the outcomes as they were presented. If participants made a response before 2 seconds (signalled by the dotted line), they were responding prior to the presentation of the outcome. Errors bars show SEM.

2.3.2.2 Stage 2 – Response Time

Figure 9 demonstrates the mean response time to the previously predicted and unpredicted outcomes in Stage 2. The predicted stimulus refers to the H – X pairings, whilst the unpredicted stimulus refers to the S – Z pairings. As in Experiments 1-2 the data were collapsed into 4 trial blocks. Inspection of the data demonstrates that, unlike Experiments 1 and 2, response times were comparable for both outcomes, with both outcomes producing faster response times as training progressed. From Block 3 onward participants

were responding to both outcomes before they were presented, thus revealing that participants learned the relationship between the cues and the outcomes.

A 2×9 repeated measures ANOVA revealed a non-significant difference between the outcomes, $F(1, 31) = .09$, $MSE = .36$, $p = .76$, $\eta_p^2 = .00$. There was a main effect of block, however, $F(2.35, 72.84) = 34.42$, $MSE = .95$, $p < .001$, $\eta_p^2 = .53$. The interaction between these effects was non-significant, $F(8, 248) = .57$, $MSE = .07$, $p = .80$, $\eta_p^2 = .02$. Given the null effect of outcome obtained in this experiment additional Bayesian analysis was conducted. This was done to determine if the similarity in response times between the outcomes supports the null hypothesis (Dienes, 2011). The Bayesian analysis was performed as outlined by Dienes (2014). This analysis produces a Bayes Factor, where values of less than .33 provide support for the null hypothesis; values > 3 provide evidence for the alternative hypothesis, whilst values ranging from .33 – 3 are inconclusive, providing support to neither hypothesis. In order to conduct this analysis an estimated effect size is required. To provide this value the mean difference ($M = .26$) and Standard Deviation (.58) between the predicted and unpredicted outcomes for Experiments 1 and 2 were calculated. By calculating the mean difference ($= .02$) and standard error of the mean difference ($= .02$) for Experiment 3, along with the estimated effect size of Experiments 1 and 2, this provided a Bayes factor of .05, providing support for the null hypothesis. Bayesian analysis further revealed that the likelihood of the null being true is 12.09, whilst the likelihood of the alternative hypothesis being true is .63. These results therefore reveal that participants learned about both outcomes at comparable

rates, with both outcomes producing faster responses times as a function of the number of trials.

2.3.2.3 Error Data

The mean percentage of errors for the predicted and the unpredicted outcome was calculated for both Stage 1 and Stage 2 data. Inspection of the data from Stage 1 revealed that participants made more errors for the unpredicted outcome ($M = 4.68$, $SEM = 1.25$) than the predicted outcome ($M = 3.56$, $SEM = .87$) and more commissions ($M = 5.64$, $SEM = 1.10$) than omissions ($M = 2.60$, $SEM = 1.02$). A 2×2 repeated measures ANOVA of percentage of errors with the variables of outcome (unpredicted vs predicted) and error type (commission vs omission) was performed on these data. The ANOVA revealed a main effect of error type, $F(1, 31) = 6.71$, $MSE = 43.98$, $p < .05$, $\eta_p^2 = .18$, with participants committing more commissions than omissions. However, there was no main effect of outcome $F(1, 31) = 1.30$, $MSE = 31.29$, $p = .26$, $\eta_p^2 = .04$, and no significant interaction between these two factors, $F(1, 31) = .00$; $MSE = 27.62$, $p = .93$, $\eta_p^2 = .00$.

Inspection of the Stage 2 data revealed that participants made a comparable number of errors for the predicted outcome ($M = 4.25$, $SEM = 1.12$) and the unpredicted outcome ($M = 3.86$, $SEM = .96$), with more commissions ($M = 5.77$, $SEM = 1.15$) than omissions ($M = 2.34$, $SEM = .93$). An identical ANOVA to that conducted on the error data from Stage 1 was conducted on the Stage 2 data which revealed a main effect of error type, with participants producing more commissions than omissions, $F(1, 31) = 6.73$, $MSE = 55.87$, $p < .05$, $\eta_p^2 = .18$, but no effect of outcome $F(1, 31) = .29$, MSE

$= 17.08$, $p = .60$, $\eta_p^2 = .01$, and no interaction, $F(1, 31) = .00$, $MSE = 18.22$, $p = .96$, $\eta_p^2 = .00$.

Experiment 3 explored whether the Stage 1 trials which established the predictability of the outcomes were necessary in Stage 2, in order to produce the outcome predictability effect observed in Experiments 1 and 2. The data presented here suggests that the presence of Stage 1 target trials in Stage 2 are necessary in order to obtain the outcome predictability effect. The omission of these trials resulted in both outcomes producing comparable response times and error rates. One explanation of this experiment is that the effect observed in Experiments 1 and 2 is sensitive to a context change, which in this case is brought about by the absence of the Stage 1 target trials which participants have been trained to attend to. The sudden omission of Stage 1 target trials, which arguably constitute the most salient part of the task for participants, may result in a significant shift in the task context and result in either a lack of transfer, or the resetting, of stimulus associability.

Alternatively, the absence of an effect in the current experiment could be explained in terms of McLaren and Mackintosh's (2000) proposition that the associative properties of a stimulus decay with the passage of time. According to McLaren and Mackintosh, learning will take place on each trial in which two stimuli are paired together. Following this pairing, however, the associative properties of a stimulus will decay (unless updated on a subsequent trial) to a level where the asymptotic value is that of the value prior to the conditioning trial, plus a relatively small share of the increase on the trial in which the stimuli were paired together. This decay mechanism allows transient relationships to be distinguished from stable relationships as the associative

properties of a stimulus which has entered into a transient relationship will decay relative to a stimulus which has entered into a stable relationship and is subsequently updated. Given this notion of decay postulated by McLaren and Mackintosh, it is conceivable that the absence of the Stage 1 target trials in Stage 2 resulted in the associability of the outcomes resetting to a comparable level. If this is the mechanism responsible for the absence of the effect, this would suggest that the associability of an outcome decays rapidly – perhaps too rapidly to be of use.

However, in light of the above results another explanation of the effect observed in Experiments 1 and 2 could also be postulated. In the absence of the Stage 1 target trials during Stage 2, the effect is no longer observed. It is therefore possible then that the effect observed in Experiments 1 and 2 was solely driven by these trials and that Stage 1 training was unnecessary. For instance, it could be that the effect will be present as long as the target trials (P – X, F/G/W – Z) establishing the predictability of the outcomes are presented alongside the outcomes which are equally well predicted by different cues (H – X and S – Z), yet Stage 1 training featuring these target trials (P – X, F/G/W – Z) is unnecessary. In this sense, it is possible that rather than prior predictability of an outcome influencing new learning about this outcome, the effect is driven by the concurrent predictability of the outcomes in Stage 2. Experiment 4 therefore sought to assess this possibility by examining whether Stage 1 training was necessary for the effect to be observed, by presenting the Stage 2 training employed in Experiment 2 in the absence of Stage 1 training.

2.4 Experiment 4

Experiment 3 demonstrated that in the absence of Stage 1 target trials (i.e., $P - X$, $G/F/W - Z$) during Stage 2, the effect observed in Experiments 1 and 2 was absent. That is, the prior predictability of an outcome did not influence subsequent learning featuring this outcome. This experiment therefore raises a question concerning whether Stage 1 is necessary in order to produce the effect observed in Experiments 1 and 2. To further clarify, it could be the case that the target trials which establish the predictability of the outcomes (i.e., $P - X$, $F/G/W - Z$) are necessary during Stage 2, alongside trials where each of the outcomes are predicted reliably by different cues (i.e., $H - X$, $S - Z$). However, Stage 1 training where the outcomes are rendered as predictable or unpredictable is unnecessary. If this is the case, this would appear to further differentiate the effect observed in Experiments 1 and 2 from learned predictiveness experiments where Stage 1 training is crucial to the effect, and suggest an additional mechanism is producing the effect. That is, the prior predictability of an outcome is of little relevance, and instead the concurrent predictability of outcomes is key to producing the effect. Experiment 4 therefore sought to assess the extent to which Stage 1 training was indeed necessary to produce this effect. The design of Experiment 4 was therefore the same as Stage 2 of Experiment 2. However, importantly, this time no Stage 1 training was provided before Stage 2. The design of Experiment 4 can be seen in Table 8.

Table 8. Design of Experiment 4.

<i>Stage 1</i>	<i>Stage 2</i>		
-	P – X,	P – X,	P – X
-	F – Z,	G – Z,	W – Z
-	H – X,	S – Z	
-	F – G,	G – F	
-	F – W,	W – F	
-	G – W,	W – G	

Note. X refers to the predicted outcome. Z refers to the unpredicted outcome. P refers to the cue which reliably predicts the outcome X. F/G/W refer to the cues which partially predict the less well-predicted outcome Z, and also serve as non-target trials when paired with each other. H and S refer to the cues of the outcomes X and Z in Stage 2. In Stage 1 '-' denotes there was no training provided.

2.4.1 Method

2.4.1.1 Participants

Thirty-two participants (27 females; 5 males) were recruited from the University of Nottingham's School of Psychology. Participants ranged from 18 – 23 years of age ($M = 19.34$; $SEM = 0.24$). Participants received course credit for their participation or a cash payment of £3.

2.4.1.2 Apparatus and Stimuli

As in Experiments 1 and 2 the stimuli were presented (and responses recorded) in the graphical experimental software package PsychoPy2, v1.82.02 (see [Peirce, 2007](#); [Peirce 2009](#)), running on Windows 7 on a standard desktop computer. The stimuli components were exactly the same as Experiment 1 (See Experiment 1, Method).

2.4.1.3 Procedure

The procedure for Experiment 4 was identical to Experiment 2, except that in this experiment participants did not complete the Stage 1 training. Thus, in total participants were exposed to 18 blocks of 16 trials (again trials were block randomised). Each block was identical (i.e., the same stimuli pairings were presented). This experiment lasted for 14.4 mins. The H – X and S – Z pairings were presented a total of 36 times. The predicted outcome (e.g., X) followed the predictive cue (e.g. P) a total of 54 times; the unpredicted outcome (Z) was also presented 54 times, being preceded an equal number of times (i.e. 18) by the same three letters (e.g., F, G and W). The letters preceding the unpredicted outcome also predicted one another (e.g. F – G, W – G), with these combinations forming the 108 filler trials present in this stage.

2.4.2 Results and Discussion

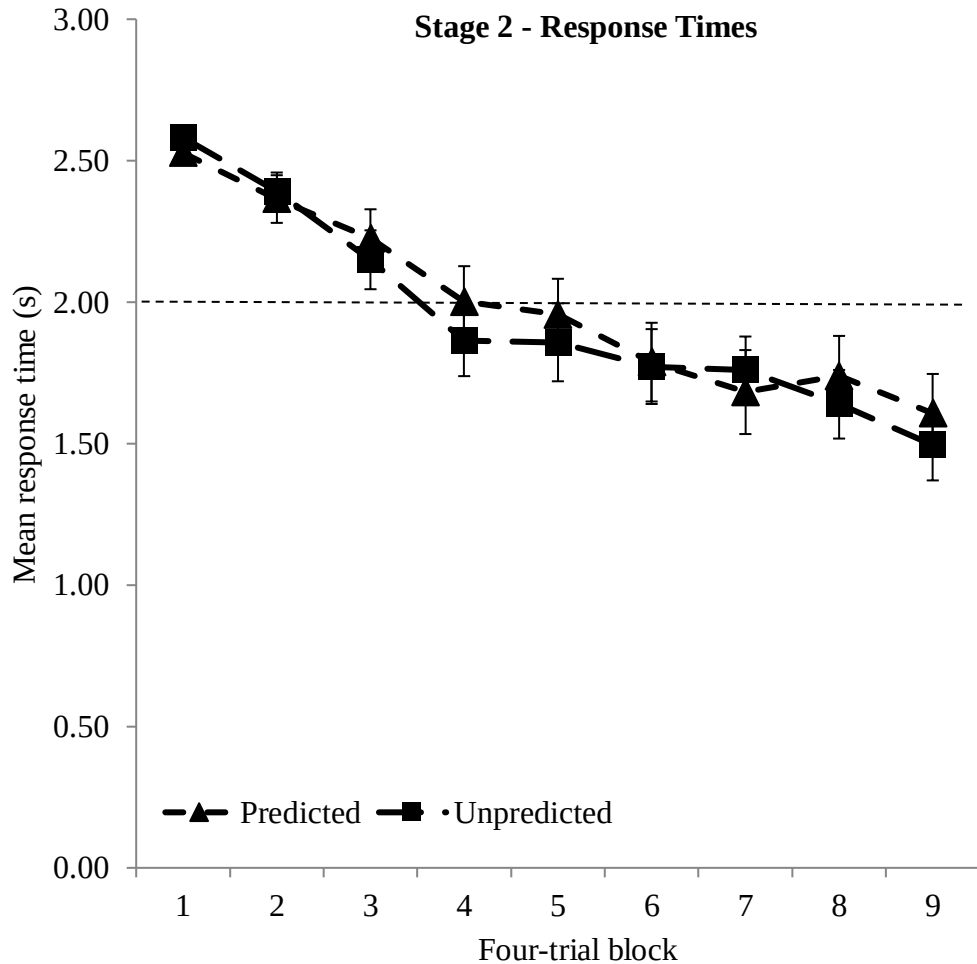


Figure 10. Mean response times to the predicted (X) and unpredicted (Z) outcomes in Stage 2. Response times within the 2 – 3 second window indicate participants were responding to the outcomes as they were presented. If participants made a response before 2 seconds (signalled by the dotted line), they are responding prior to the presentation of the outcome.

2.4.2.1 Stage 2 – Response Times

Figure 10 illustrates the mean response times to the ‘predicted’ and ‘unpredicted’ outcomes. The predicted stimulus refers to the ‘H – X’ pairings, whilst the unpredicted stimulus refers to the ‘S – Z’ pairings. As in

Experiments 1 and 2, the data were collapsed into 4 trial blocks. Inspection of the data demonstrates that, unlike Experiments 1 and 2, response times were largely the same for both outcomes, with both decreasing as a function of the number of trials. Typically, from Block 4 onward participants were responding to both outcomes before they were presented, indicating that participants learn the relationship between the cues and the outcomes.

A 2 [Outcome: ‘Unpredicted’, ‘Predicted’] $\times$ 9 [Block: 1- 9] repeated measures ANOVA revealed that there was a non-significant difference between the outcomes $F(1, 31) = .31, MSE = .82, p = .58, \eta_p^2 = .01$. The interaction between these factors was also non-significant, $F(4.14, 128.43) = 1.04, MSE = .19, p = .39, \eta_p^2 = .03$. However, the ANOVA revealed that there was a main effect of the factor block, $F(3.19, 98.92) = 31.11, MSE = .59, p < .001, \eta_p^2 = .50$. Given the null effect of outcome obtained in this experiment additional Bayesian analysis was again conducted in accordance with Dienes (2011). This was done to determine if the similarity in response times between the outcomes supports the null hypothesis. By calculating the mean difference ($M = -0.02$) and standard error of the mean difference ($SEM = -0.03$) for Experiment 4, along with the estimated effect size of Experiments 1 and 2 this provided a Bayes factor of .02, providing support for the null hypothesis. Bayesian analysis further revealed that the likelihood of the null being true is 10.65 whilst the likelihood the alternative hypothesis being true is .23. As such, these results provide evidence of an absence in a difference in learning about the outcomes when only Stage 2 training is provided.

2.4.2.2 Error Data

The mean percentage of errors for the predicted and the unpredicted outcome was calculated. Inspection of the data revealed that participants made a comparable amount of errors for both the predicted ($M = 4.73$, $SEM = .101$) and the unpredicted outcome ($M = 4.86$, $SEM = .103$). However, participants appeared to make more errors of commission ($M = 5.64$, $SEM = 1.02$) rather than omission ($M = 3.95$, $SEM = 1.02$), particularly for the predicted outcome ($M = 6.42$, $SEM = 1.09$) relative to the unpredicted outcome ($M = 4.86$, $SEM = .94$). These impressions were confirmed with a 2×2 repeated measures ANOVA of percentage of errors with the variables of outcome (unpredicted vs predicted) and error type (commission vs omission). The ANOVA revealed no effect of outcome $F(1, 31) = .03$, $MSE = 20.46$, $p = .87$, $\eta_p^2 = .00$, or error type $F(1, 31) = 2.24$, $MSE = 40.91$, $p = .14$, $\eta_p^2 = .07$. There was, however, a significant interaction between these two factors, $F(1, 31) = .442$; $MSE = 20.75$, $p = .04$, $\eta_p^2 = .12$. Simple main effects analysis was conducted to explore the source of the interaction. These analyses revealed significant differences between the type of error committed for the predicted outcome, with more errors of commission than omission, $F(1, 31) = 6.21$, $p = .02$, $\eta_p^2 = .17$. However, there was no difference in error types committed on the unpredicted outcome, $F(1, 31) = .00$, $p = 1.00$, $\eta_p^2 = .00$.

Experiment 4 assessed whether Stage 1 training was necessary in order to produce the effect observed in Experiments 1 and 2. Given the result of Experiment 3 it could be postulated that Stage 1 is unnecessary to produce the effect observed in Experiments 1 and 2. These results, however, appear to indicate this is not the case. In this experiment both outcomes were learned

about at the same rate in the absence of the Stage 1 training. Both outcomes also produced a comparable amount of errors. However, there was a difference in the type of errors committed for the predicted outcome with more errors of commission. This could possibly reflect the fact that participants were less certain about the predicted outcome due to the lack of Stage 1 training (where the outcome is rendered predictable). The reduced certainty given the lack of Stage 1 training is also likely to influence learning with the unpredicted outcome; however, given the greater number of associates in this Stage with the unpredicted outcome (F/G/W/S - Z), it is possible participants were more likely to respond with the response required for the unpredicted outcome than the predicted outcome. In summary, this experiment demonstrates that the Stage 1 training provided in Experiments 1 - 3 is necessary to produce the effect observed in Experiments 1 and 2. This therefore suggests that the previous predictability of the outcomes governs, at least partially, the differential learning rates of the outcomes in Stage 2.

2.5 General Discussion

Four experiments investigated whether an outcome's prior associative history influences subsequent learning involving that outcome, when separated from the cue with which the outcome was first paired with. In Experiments 1 and 2 participants displayed a bias in learning towards a previously predicted outcome, with the predicted outcome entering into associations with novel cues more readily than the previously unpredicted outcome. Although Experiment 1 could be accounted for by appealing to the role of inhibitory learning established in Stage 1, the design of Experiment 2 ruled out this

explanation and replicated the effect observed in Experiment 1 both with response time measures and, in addition, self-report ratings.

Experiment 3 assessed whether the Stage 1 target trials which first established the predictability of the outcomes were required to also be present in Stage 2 in order to obtain the effect. In this experiment the omission of these trials in Stage 2 resulted in both outcomes producing comparable response times and error rates. These results could possibly demonstrate that the effect is context sensitive. The omission of the Stage 1 target trials would certainly appear to constitute a change to the context, which could potentially disrupt a transfer of learning between the two stages. Alternatively, the absence of the effect, could be accounted for by assuming that the associability of the outcomes decay at a rapid rate in the absence of the trials which first established the predictive nature of the outcomes.

Experiment 4 assessed whether Stage 1 training was necessary to produce the effect observed in Experiments 1 and 2. In light of the results of Experiment 3, it is possible that the presence of the Stage 1 target trials during Stage 2 was driving the effect, and thus Stage 1 training was unnecessary during Experiments 1 and 2. As such, in Experiment 4 participants completed only Stage 2 training of Experiment 2, in the absence of any Stage 1 training. This experiment demonstrated that participants did not display any bias in learning toward the outcomes on the trials of interest (H-X, S-Z), in the absence of Stage 1 training. These results therefore demonstrate that the previous training involving the outcomes is crucial to obtaining the effect.

The results of these studies are broadly consistent with the findings of Griffiths et al. (2015) and Thorwart et al. (2017) supporting their notion of a

“outcome predictability” effect. In their studies participants completed the food allergist task with the reactions a fictitious patient experienced serving as the predicted or unpredicted outcomes. These authors observed the outcome predictability effect when participants were required to provide a rating regarding the strength of the association between the novel cues and the outcomes or to make a prediction about which outcome might follow a cue. In Experiments 1 and 2, reported here, however, the effect was observed during the learning phase when a different measure was used (i.e., reaction time). This suggests that the effect influences how participants form associations with these outcomes, in addition to how they judge the strength of these associations at a test stage (which can be biased by rational processes). Additionally, in the experiments reported here single cues and outcomes were presented to participants as opposed to compounds of cues and outcomes, which previous studies have employed. The fact that comparable findings have also been observed when different designs have been employed would appear to strengthen the reliability of this effect.

Comparable findings have also been recently reported by Griffiths, Erlinger, Beesley and Le Pelley (2018). These authors employed a visual search task in order to assess whether the prior predictability of an outcome could influence ‘online’ learning about these outcomes. In this task participants were presented with a cue stimulus (e.g., an image of a plane), followed by a search display containing 8 coloured shapes positioned around the image. One of these shapes would contain a target arrow (facing left or right) which indicated the required response from participants. In Stage 1 two cue stimuli would predict the location of the target arrow 100% of the time

(thus rendering the locations predictable), whilst two remaining stimuli predicted two possible locations in a random fashion (thus rendering these locations unpredictable). The orientation of the target arrow was never predicted by any stimuli. In Stage 2 participants were presented with novel cues all of which reliably predicted the location of the target arrow. During this stage participants demonstrated faster responses to the target location when it was predictable in Stage 1, compared to when it was unpredictable, thus replicating the results of Experiments 1 and 2 in this chapter, in terms of the measure they employed (i.e., reaction time).

The experiments reported in this chapter, together with others (Griffiths et al., 2015; Griffiths et al. 2018; Thorwart et al. 2017) suggest that a revision may be required to learning models which stipulate that β is a fixed learning-rate parameter. It is thus tempting to suggest that changes in attention/associability to both cues (α) and outcomes (β) are underpinned by comparable rules, although their functional role must remain the same (see Thorwart et al.) which suggests that learning can change the alpha or beta of a stimulus – but this does not transfer between the two. For example, the algorithms suggested by Mackintosh (1975; see also: Esber & Haselgrove, 2011; George & Pearce, 2012; Le Pelley, 2004) concerning changes in cue associability could be applied to outcome associability. Such a suggestion immediately poses a conundrum, however. As noted in the introduction it is relatively straightforward to see the adaptive purpose of a mechanism that varies the amount of attention that is paid to what are essentially neutral cues on the basis of their predictive significance. This mechanism permits organisms to tune out stimuli that are irrelevant predictors of important

outcomes – be it food to a rat, or the goal of a computer-based task in a human – thus facilitating the acquisition of anticipatory behavior towards them.

However, it seems less clear what the adaptive purpose would be of a mechanism that changes the attention paid to an outcome based on whether or not it has been well-predicted, for by the time the outcome is presented, it is too late.

One way of resolving this conundrum is to suggest that events that share a predictive relationship, be they at the start or the end of a predictive chain, are inherently important. If organisms are conceived of as information foragers (Pirolli & Card, 1999; Pirolli, 2007; also see Melara & Algom, 2003) then an outcome that is well predicted provides more information about the structure of the world than an unpredicted outcome does, and therefore has more value - potentially increasing the attention it captures (e.g. Le Pelley, Mitchell & Johnson, 2013; Le Pelley, Pearson, Griffiths & Beesley, 2015; Pearson, Donkin, Tran, Most & Le Pelley, 2015). In addition, learning to shift attention between well-predicted and less well-predicted outcomes might permit organisms to better modify their behavior when the predictive structure of the world changes. For example, consider the case of the parents of a child who have learned that whenever their child drinks milk she suffers with a stomach upset. If the child should subsequently have the same stomach upset despite not drinking milk, then it may stimulate the parents to search for potential causes of stomach upset in order to make sense of what was thought to be a well-understood world, but which has now become uncertain. In this sense, the mechanism is sensitive to uncertainty in a way that is, conceptually, similar to that proposed by Pearce and Hall (1980).

A parsimonious explanation of the results obtained in the experiments reported here appeals to context-blocking (Baker & Mercier, 1982; Randich & LoLordo, 1979). According to this account, an association is assumed to form between the unpredicted outcome (Z) and the experimental context as a result of Stage 1 training trials (F/G/W – Z) featuring the unpredicted outcome. Thus, in Stage 2, when Z is reliably paired with a novel cue, the context-Z association blocks (or competes with) learning about the cue-Z association. In contrast, however, the predicted outcome (X) was always preceded by cue P during Stage 1, which was a more reliable predictor than the context, therefore reducing the likelihood of a context-X association forming, and obscuring learning when X is paired with a novel cue in Stage 2.

One issue with this analysis, however, is the difficulty in ascertaining what constitutes the experimental context in this task. There could be a variety of features which form the experimental context (such as the background colour of the task), and the salience of these features are likely to differ between participants. This makes it difficult to determine with any degree of certainty what constitutes the context in this experiment. That said, we can take some measure of the background expectation of Z by examining behaviour towards the stimuli, F, G and W. These stimuli are the most exposed stimuli in the experiment as they are paired with each other to form non-target trials; thus, participants must constantly monitor these stimuli in order to detect the target stimuli. It is, therefore, conceivable that these stimuli are the most salient feature of the context (or may overshadow any alternative feature of the context).

Given this logic, if learning about F/G/W – Z pairings is taken to be a measure of the context – Z association, we can gauge the strength of the context – Z association by examining participants' ratings for F/G/W – Z. For instance, high ratings regarding these stimuli and Z might be taken to reflect a strong context – Z association, whilst low ratings to these stimuli represent a weak context – Z association. Importantly, if a strong context – Z association was formed in Stage 1 this could block learning about the S – Z association from proceeding at the same rate as H – X in Stage 2. The outcome X is less likely to enter into an association with the context as it is reliably predicted in the first stage of training and therefore this association is likely to overshadow any context – X association forming (i.e., there is not a background expectation of X to same extent as Z). As such, there is no competing context – X association blocking learning about H – X in Stage 2. Alternatively, however, it could also be the case that the context – Z association could result in enhanced contextual conditioning to S – Z as a result of excitation by the context, therefore facilitating learning about S – Z as opposed to blocking learning about this association (e.g., Taylor, Joseph, Balsam, & Bitterman, 2008).

If we look at the ratings provided for F/G/W - Z and correlate these with the size of the effect for each participant by subtracting the ratings provided for the S – Z association from the H – X association, we can assess the extent to which this feature of the context was correlated with the effect of interest. According to a context blocking explanation it would be expected that greater learning about the F/G/W – Z association would result in poorer learning about the S – Z association. Alternatively, if the context – Z

association facilitated learning about the S – Z association, better learning about S – Z learning would be seen. This analysis provides little support for either account, however, $r_s(27) = -.03$, $p = .43$. Moreover, participants' ratings of F, G and W, were relatively low and although demonstrating a trend toward significance, did not differ from zero, $t(27) = 1.92$, $p = .07$ (Griffiths et al., 2015 and Thorwart et al., 2017 have also refuted a context blocking account of the effect).

The presence of the F/G/W - Z trials could enable an alternative explanation, however, given the results of Experiment 3. The absence of the effect in this experiment, when the Stage 1 trials target trials are omitted in Stage 2, could indicate that rather than demonstrating some transferred learned predictability, the effect is dependent on differential concurrent predictability in Stage 2. That is, an outcome must be currently well-predicted by other stimuli in order for this outcome to enter into novel associations faster than a less-well predicted outcome. However, this would still appear to suggest that the outcome's associability is susceptible to change, and moreover this line of argument would appear to be undermined by subsequent studies which have studied the outcome predictability effect.

Another account of the effect could also be postulated based on the number of associates each outcome has in the experiments. For instance, in Stage 2 of these experiments, participants were presented with both the novel predictors of the previously presented outcomes and those stimuli which established the predictability of the outcomes in Stage 2 (excluding Experiment 3). As such, there is an inconsistency in how these outcomes are treated in this stage. The predicted outcome (X) receives consistent training

with two associates (i.e., X is always well-predicted by two stimuli), however, the unpredicted outcome (Z) receives inconsistent training: it is reliably predicted on some trials (S-Z) and partially predicted by an additional number of other stimuli on other trials (G/F/W – Z). It is therefore possible to provide an explanation of the effect in Experiments 1 and 2 based on a form of interference, given the unpredicted outcome's greater number of associates, as opposed to the outcome's prior predictability. However, Experiment 4 would appear to rule out this possibility: when participants completed Stage 2 training in the absence of Stage 1, they did not demonstrate any bias in learning toward the outcomes on the trials of interest (H-X, S-Z).

An additional alternative explanation for our results can also be developed if the contingencies in this task are viewed as instrumental as opposed to Pavlovian. For example, withholding a response is required to stimuli F, G and W, on two-thirds of the trials with these stimuli, in which case it is possible to view the task employed here as a Go/No-Go (Donders, 1969). The inhibition acquired on the No-Go trials with F, G and W may become associated with the outcome Z when it and F, G and W are paired – inhibition which will need to be overcome in Stage 2 in order for the S-Z association to be expressed. Previous studies have demonstrated that stimuli can become associated with instrumental inhibition (see Bowditch, Verbruggen & McLaren, 2016; Verbruggen & Logan, 2008). Although this analysis provides a potential explanation for the response time data recorded in Experiments 1 and 2, it seems less clear how this analysis would explain the differences in participant's causal-ratings provided in Experiment 2 where a judgement about

the relationship between H and X, and S and Z was required, rather than the performance of a speeded response.

In summary, the experiments in this chapter examined how the prior predictability of an outcome influences subsequent learning involving that outcome. An assumption common to many associative models of learning is that an outcome's associative history should have no subsequent influence on learning involving that outcome when it is separated from the cue with which it was trained. The current studies question this assumption and provide further insight into the nature of the outcome predictability effect and its implications for learning.

Chapter 3: Outcome Facilitation and Outcome Blocking

As noted in Chapter 1, when two or more cues are presented in compound this can impact upon the way in which these cues are learned about relative to when they are presented alone (i.e., Honey & Hall, 1988; Jones & Haselgrove, 2013; Kamin, 1968, 1969; Le Pelley, Oakeshott & McLaren, 2005). This is demonstrated by the blocking effect first reported by Kamin (1968, 1969). In Kamin's experiments, pretraining with an element (A+) of a compound stimulus (AX+) and a US, blocked learning about the association between the additional element (X) of the compound and the same US. This effect has subsequently been reported in a diverse range of species and a variety of theoretical accounts have been provided to account for its ubiquitous nature. According to key associative learning models this effect occurs as a cue's prior associative history not only influences learning about that cue but also about other cues presented in compound with that cue (see Chapter 1 for further discussion).

Although a vast array of studies have demonstrated this effect amongst cues, relatively little research has explored whether this effect extends to outcomes. For example, relatively few studies have explored whether prior learning about a cue and an element (e.g., shock) of a compound outcome (e.g., shock + loud noise), blocks learning about the association between the cue and an additional element (e.g., loud noise) of the outcome compound. This lack of investigation could be partly attributable to key associative learning models. As noted in Chapter 1, and expanded upon in Chapter 2, these models afford cues and outcomes a differential status with cues provided with more freedom

to interact with one another and vary in their associability than outcomes.

According to associative learning models there is little reason for the associability of an outcome to change based on its past associability, or the past associability of accompanying outcomes. Once the outcome has been presented – whether in elemental form or in compound – the event has already happened, and as such there is little benefit to altering behaviour in the same manner as cues. Nevertheless, as seen in Chapter 1, there is evidence from animal studies (e.g., Miller & Matute, 1998; Rescorla, 1980) to suggest that outcomes do indeed interact to produce effects comparable to interaction effects observed in cues.

For instance, both Rescorla (1980) and Miller and Matute (1998) demonstrated that learning about the association between a cue and an element (A – O1) of an outcome compound (O1 O2) impaired learning about the association between the cue and the additional element of the outcome compound (A – O2). Rescorla's interpretation of his results appealed to the idea that the pigeons in his experiment were distracted by the presence of the outcome element which appeared both on its own and in compound, as a result of using a second order conditioning procedure which ensured the element of the outcome acquired biological significance. This explanation does not pose a theoretical problem to associative models as it suggests that rather than associative competition between the elements of the compound influencing the results, the animal is simply distracted by a highly-valued stimulus. Miller and Matute (1998) overcame these issues to some extent by immersing an outcome blocking design in a sensory preconditioning paradigm, thus ensuring the outcomes did not come to acquire biological significance until after the

outcome compound training stage (see Chapter 1 for a more detailed discussion of these studies). Nonetheless, in both these studies indirect conditioning procedures have been employed to produce comparable effects, potentially suggesting such effects are constrained to indirect conditioning methods as noted by Burger, Mallemat and Miller (2000).

Similar studies have also been conducted in humans using methods comparable to more standard methods of conditioning (e.g., first order), however, these studies have produced heterogeneous results. For example, Cobos, Lopez, Caño, Almaraz and Shanks (2002) employed a causal judgement task to assess the impact of causal structure (whether causes or effects are presented first) and cue structure (whether multiple cues are presented prior to a single event or whether a single cue is followed by multiple events). In a series of experiments participants were required to take on the role of workers in a chemical plant which leaked dangerous substances. The substances resulted in different indicator lights being triggered and participants were required to learn the association between the substances and the indicator lights.

In the first two experiments the causal structure was manipulated with participants either completing a cause and effect version of the task, where leaked substances were presented first and indicator lights were triggered, or an effect and cause version of the task, where participants were presented with the indicator light which was triggered and then the leaked substance. In these two experiments participants were always presented with multiple cues as opposed to multiple outcomes. In both versions of the task a control cue which exclusively predicted an outcome was learned about better than a cue presented

in compound with a cue which reliably predicted the outcome on a series of separate elemental trials. In Experiment 3 both the causal structure and the cue structure was manipulated. Participants completing the condition relating to the causal structure (whether causes or events were presented first) again demonstrated the same effect in both versions of the task as in the previous two experiments. However, participants in the cue structure condition who were presented with multiple outcomes and single cues (as opposed to multiple cues and single outcomes) demonstrated no difference in their ratings of the outcomes. More specifically, participants' ratings were comparable for the control outcome which was exclusively predicted by a cue, and an outcome presented in compound with another outcome which was reliably predicted on a series of separate elemental trials. These results demonstrate the absence of an outcome interaction effect, therefore being inconsistent with the findings of Miller and Matute (1998) and Rescorla (1980); comparable findings have also been reported by Price and Yates (1995).

In contrast, however, a study by Flach, Osman, Dickinson and Heyes (2006) found a different set of results with humans. In their study these authors assessed the role of associative learning mechanisms in response priming. In Experiment 1 the authors sought to observe the response priming effect first observed by Elsner and Hommel (2001). That is, in a training stage, participants were asked to make a response by pressing one of two keys. Upon selecting a key, a brief visual effect in the form of a red or pink screen background would be presented. In a subsequent Test Stage, participants were presented with the visual effect stimuli and were asked to make a specified response as quickly as possible. For one group of participants (Group

Congruent), the response was the same as the training stage. For another group the response was the opposite to the training stage (Group Incongruent). In this experiment participants' responses were shorter in Group Congruent than Group Incongruent, therefore replicating the response priming effect reported by Elsner and Hommel (2001).

In Experiment 2 of this study participants received the same Stage 1 training as in Experiment 1, however, in a second stage of training an auditory stimulus was presented alongside a visual stimulus, thus forming an outcome compound. In one condition (condition pretrained) the response for the visual stimulus in Stage 2 was the same as in Stage 1. For the control condition, however, the visual and auditory stimuli were both novel. At a test stage the question of interest was how quickly participants were able to provide the required response to the auditory stimuli that were presented in compound with the visual stimuli in Stage 2. In this experiment participants provided shorter response times to the auditory element of the outcome compound in the control condition as opposed to the pretrained condition. That is, when the auditory element was paired with a visual stimulus that a response had been primed to previously, responses were slower than when the auditory stimulus was paired with a novel visual stimulus. As such, these results demonstrate an outcome blocking effect comparable to that reported in animals (e.g., Miller and Matute, 1998; Rescorla, 1980). In Experiment 3, however, the auditory stimuli were presented in the initial training stage (as opposed to the visual stimuli) prior to being paired with visual stimuli and the opposite effect was observed. That is, pretraining with an element of an outcome compound enabled better learning

about the added visual stimulus, therefore producing what can be referred to as an outcome facilitation effect.

Taken together the results of the above studies would appear to suggest that interaction effects are not unique to cues. These results are therefore inconsistent with the assumptions of key associative learning models, which go to great lengths to explain how cues may interact to modulate the strength of conditioned responding, but which are, largely, silent about comparable situations in outcomes. Indeed, Miller and Matute (1998) note that demonstrations of outcome blocking appear to challenge the differential treatments of cues and outcomes in these models. As can be seen from the above studies, however, outcome blocking is not the only effect to be observed, with the opposite effect, outcome facilitation, also being observed (e.g., Flach et al., 2006) and null effects also being reported (e.g., Cobos et al., 2002; Price & Yates, 1995), the latter of which are consistent with the predictions of key associative learning models. Given the heterogeneous results of the above studies and the discrepancy between theory and data, it would appear important to further explore this effect. Aside from the theoretical conundrum that the presence of this effect may pose to theories of associative learning, it would seem prudent to explore the parameters of blocking in outcomes given the relative scarcity of research on outcome blocking in humans, and the prevalence and importance of this effect - in cues - to associative learning.

In light of the above, the experiments in this chapter sought to explore the generality of outcome interaction effects in humans using a food allergist paradigm, a commonly used associative learning task, similar to that of Le

Pelley, Oakeshott and McLaren (2005). If previous learning about an element of a compound outcome can block learning between a cue and an additional outcome in a manner consistent with Miller and Matute (1998) and Rescorla (1980), it would be expected that a ‘blocked’ outcome would be learned about less well than a control outcome which is surprising. If, however, the elements of an outcome do not compete for associative strength as a result of their associative history – in the same manner as cues – it would be expected that the outcomes will be learned about to the same extent, as is suggested by some studies (Cobos et al., 2002; Price & Yates, 1995) and models of associative learning.

3.1. Experiment 5

Experiment 5 comprised a variant of the causal judgement task employed by Le Pelley et al. (2005), in which participants were asked to imagine themselves as a food allergist tasked with identifying which foods caused certain allergic reactions. Participants were informed about the foods which a fictitious patient (‘Mr. X’) had consumed and were then presented with information about the reaction the patient received. The foods (A – D) served as cues, whilst the reactions (O1-O4) served as outcomes. The design of Experiment 5 can be seen in Table 9.

In Stage 1, four cues (A – D) were presented, each of which reliably predicted one of two outcomes (O1 and O3) and participants were tasked with learning which cues were paired with which outcomes. Prior to Stage 2, participants were informed that Mr. X had started to display some additional reactions. Once Stage 2 commenced the same four cues were presented. On

blocking trials, cues A and B continued to predict the same outcomes as in Stage 1 (O1 and O3 respectively) and an additional novel outcome (O2 and O4 respectively). For the control trials, participants were presented with cues C and D, which now predicted outcomes that these cues had not previously been paired with in Stage 1 (i.e. O1 and O3 respectively), and a novel outcome (i.e. O2 and O4). In a subsequent test stage, participants were presented with a screen which required them to provide ratings of how predictive the cues (A - D) were of each of the outcomes. If the Stage 1 training with A – O1 / B - O3 blocked learning about A - O2 / B - O4, it would be expected that participants would provide lower test ratings concerning the relationship between A - O2 / B- O4 relative to ratings for C - O2 / D - O4.

Table 9. Design of Experiment 5

	<i>Stage 1</i>	<i>Stage 2</i>	<i>Test</i>
Blocking	A – O1	A – O1 O2	A: O1- O4
	B – O3	B – O3 O4	B: O1- O4
Control	C – O3	C – O1 O2	C: O1- O4
	D – O1	D – O3 O4	D: O1- O4

Note. A – D refer to foods (i.e. cues), whilst O1 – O4 refer to outcomes. A and B cues represent the blocking stimuli, whilst C and D represent the control stimuli.

3.1.1 Method

3.1.1.1 Participants

Thirty-two participants (24 females; 8 males) were recruited from the University of Nottingham’s School of Psychology. Participants ranged from 18 – 30 years of age ($M = 21.7$; $SEM = 0.61$). All participants had normal or

corrected-to-normal vision. Participants received course credit for their participation or a £3 inconvenience allowance. The study was conducted in accordance with the British Psychological Society (2009) guidelines, receiving institutional ethical approval from the University of Nottingham's Psychology ethics committee.

3.1.1.2 Apparatus and Stimuli

All stimuli were presented (and responses recorded) in the graphical experimental software package PsychoPy2, v1.83.01 (see Peirce, 2007; Peirce 2009), running on Windows 7 on a standard desktop computer (screen size: 27cm × 46cm; h × w). The foods Mr. X had eaten served as the cue stimuli. The reactions Mr. X experienced served as the outcomes. Each of the foods and reactions were assigned to the letters and outcomes in Table 9 using a Latin-Squared counterbalancing technique with each food serving as cues A – D, and each reaction serving as outcomes O1 – O4, across the course of the experiment.



Figure 11. The four foods (i.e., cues) used in this experiment were broccoli, onion, cabbage and potato

There were four pictures of foods: broccoli, cabbage, onion and potato (see Figure 11). Each food picture (9mm × 8mm; h × w) was presented in the centre of the screen against a grey background. The reactions Mr. X experienced served as the outcome stimuli. There were four different types of

reactions Mr. X experienced: diarrhoea, fever, skin rash and vomiting. In Stage 1 participants were only presented with two of these reactions (e.g. vomiting and diarrhoea). In Stage 2 participants were presented with all four of the reactions, with each of the reactions for Stage 2 being presented alongside the Stage 1 reactions to form a compound outcome stimulus (e.g. vomiting was paired with fever to form the compound stimulus ‘VOMITING AND FEVER’, whilst diarrhoea was paired with skin rash to form the compound stimulus ‘DIARRHOEA AND SKIN RASH’). Each of these reactions was presented in capitalised white Arial text (size: 32) at the bottom of the screen. The reactions were positioned beneath a white scale (length: 35cm). Participants could select their choice of reaction by positioning a cursor on the scale above the reaction they wished to select (see Figure 11 for example). Participants could only select the ends of the scale and thus had to make a binary choice. The order of the reactions beneath the scale was counterbalanced between participants.

At test each of the cues were presented individually (positioned on the left of the screen) and participants were asked to rate how predictive each cue was of each of the reactions. Each of the reactions were positioned, alongside the food, on the right of the screen (see Figure 11 for example). Participants made their ratings for each reaction by moving a cursor on a Likert scale which ranged from 0 – 100 [‘0’ = ‘Very unpredictable’, ‘100’ = ‘Very predictive’].

3.1.1.3 Procedure

Participants were tested individually or in small groups of two to three. Once participants had provided their consent they read the following instructions on a computer monitor before proceeding to Stage 1:

“Thank you for participating in this experiment.”

In this experiment we would like you to imagine that you are an allergist (i.e. someone who tries to discover the cause of allergic reactions). You have just been presented with a new patient who suffers from different types of allergic reactions as a result of eating certain foods. In an attempt to discover which foods cause the different types of allergic reaction in Mr. X, you arrange for him to eat a number of different foods on their own, and observe the type of allergic reaction he suffers. On the following screens, you will be shown the foods Mr. X has eaten, and you will be asked to predict what type of allergic reaction he will suffer as result of eating each meal. Each allergic reaction will be presented at the bottom of the screen. Make your prediction by selecting one of the allergic reactions below each of the foods. You will then be provided with feedback about what reaction Mr. X experienced. You will have to guess at first, but with the aid of the feedback your predictions should soon start to become more accurate. The participant with the highest overall accuracy will receive a £10 cash prize. Please press the SPACE BAR to begin.”

Stage 1

In Stage 1 of the experiment participants were exposed to 6 blocks of 16 trials (96 trials in total). Trial order was block randomised. There was no break between blocks. In each block participants received 4 pairings of each cue – outcome pairing (A – O1, B – O3, C – O3 and D – O1). On each trial, participants were presented with the following text at the top of the screen: “Mr X. has eaten the following food (below), which reaction did he experience?”. This text was presented above the food which was being presented (only one food would ever be presented on each trial, see Figure 12).

To proceed to the next trial, participants needed to make a response by selecting a reaction beneath the food. Once they had made a response, by selecting a reaction out of the binary choices available to them they were presented with feedback informing them about the reaction Mr. X had experienced. On this screen participants were presented with the following text: “*Mr X. experienced: [reaction name here]*”. The reaction Mr. X experienced was presented in the centre of the screen. Participants could take as much time as they needed when selecting which reaction they thought Mr. X had experienced. The feedback screen was presented for 1.5 seconds, before the next trial commenced automatically. Once all 96 trials were complete, Stage 2 began.

Stage 2

Prior to commencing Stage 2 participants were presented with the following text: “*Mr X has started to display additional reactions when eating these foods. The reactions will be listed at the bottom of the screen as before. Please press the SPACE BAR to continue.*”

In Stage 2, participants were exposed to 2 blocks of 16 trials (in line with Stage 1, trials were block randomised) with each trial type being presented a total of 4 times per block. As in Stage 1 participants were presented with the text: “*Mr X. has eaten the following food (below), which reaction did he experience?*” which was positioned above the food being presented. However, this time there were additional reactions presented alongside the Stage 1 reactions (e.g. ‘VOMITING’ from Stage 1, would now be presented with ‘FEVER’ in Stage 2 to form ‘VOMITING AND FEVER’, see Figure 12). The order in which the Stage 2 reactions were presented

alongside the Stage 1 reactions was counterbalanced across experimental conditions. For half of the participants the Stage 2 reaction was presented on the right of the reaction from Stage 1 which it was paired with (e.g. O1 O2); for the other half the Stage 2 reaction was positioned on the left of Stage 1 reaction which it was paired with (e.g. O2 O1). Once participants completed all 32 trials of Stage 2, they proceeded to the test stage.

Test

Prior to commencing the test stage, participants were presented with the following text: *“You will now be asked to rate how predictive each food was of each of the reactions. Please press the SPACE BAR to continue”*. Each of the foods were then presented at test, one per screen (see Figure 12). On each screen participants were presented with the text: *“Please rate how predictive this food was of each of the outcomes [0 = very unpredictable; 100 = very predictive]. Make your rating and then press the spacebar to proceed”*. Participants were then required to make a rating for each of the reactions using the Likert scales positioned to the right of the reactions. Once they had rated each food’s ability to predict each reaction, they were informed that the experiment was complete.

After eating this food, what allergic reaction did Mr X experience?



DIARRHOEA VOMITING

DIARRHOEA

After eating this food, did Mr X have an allergic reaction?



DIARRHOEA and SKIN RASH VOMITING and FEVER

DIARRHOEA and SKIN RASH

Please rate how predictive this food was of each of the outcomes.
Make your rating and then click on the grey box beneath the scale to confirm your rating. 0 = very unpredictable; 100 = very predictive



VOMITING	very unpredictable	very predictive	79
DIARRHOEA	very unpredictable	very predictive	CLICK HERE
FEVER	very unpredictable	very predictive	CLICK HERE
SKIN RASH	very unpredictable	very predictive	CLICK HERE

Please Press the space bar to move onto the next screen

Figure 12. An example screen that participants would be presented with during Stage 1 (top), Stage 2 (middle) and at test (bottom).

3.1.2 Results and Discussion

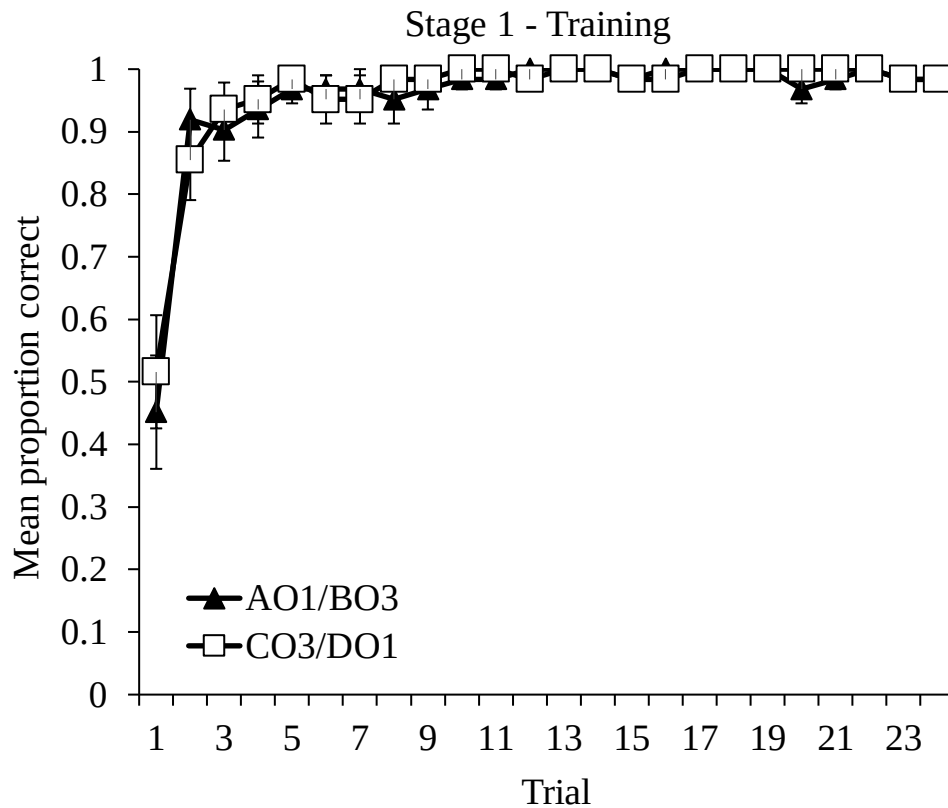


Figure 13. Mean proportion of correct responses for A – D: O1 and O3. Error bars represent SEM.

3.1.2.1 Stage 1

Figure 13 shows the mean percentage of correct responses for each cue – outcome pairing (A-O1, B-O3, C-O3 and D-O1) across the 24 trials of Stage 1. These data were calculated by averaging participants' correct responses for each trial. Participants' proportion of correct responses for blocking stimuli A and B, and control stimuli C and D were collapsed together for each trial. As can be seen in Figure 13, learning progressed rapidly for both the A-O1/B-O3 trials and the C-O3/D-O1 trials, with both cue-outcome associations being learned at a comparable rate. Furthermore, from Trial 4 onward participants were approaching asymptote for both A-O1/B-O3 and C-

O3/D-O1 with participants producing a mean correct response rate of > 90%, and this high accuracy was maintained across the course of Stage 1.

A 2×24 repeated measures ANOVA was performed on these data with the factors of Outcome (A-O1/B-O3 vs C-O3/D-O1) and Trial (1 – 24). The ANOVA revealed a significant main effect for the factor Trial, $F(23, 690) = 40.73$, $MSE = .02$, $p < .001$, $\eta_p^2 = .58$. There was no significant effect of outcome, $F(1, 30) = .67$, $MSE = .02$, $p = .42$, $\eta_p^2 = .02$, and no interaction between outcome and trial, $F(23, 690) = .58$, $MSE = .02$, $p = .95$, $\eta_p^2 = .02$. The ANOVA therefore confirmed the impressions of the graph, revealing that participants learned about both A-O1/BO3 and C-O3/D-O1 at comparable rates, with most learning occurring on the first two trials before reaching asymptote.

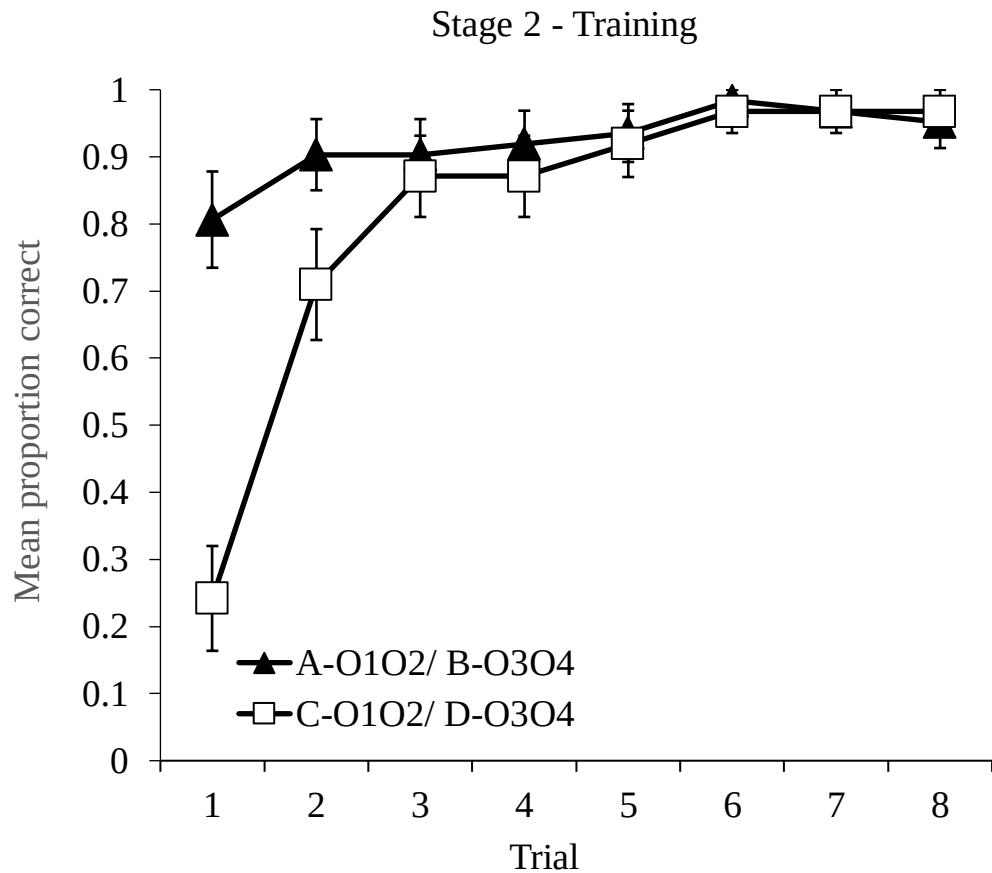


Figure 14. Mean proportion of correct responses for A – D for O1O2 and O3O4. Error bars represent SEM.

3.1.2.2 Stage 2

Figure 14 demonstrates the mean percentage of correct responses for each cue – outcome compound (A-O1O2, B-O3O4 and C-O1O2, D-O3O4) across the 8 trials of Stage 2. These data were collated by averaging participants' responses for each trial type (blocking [A and B] vs control [C and D]). As can be seen in Figure 14 participants initially demonstrated a higher percentage of correct responses for the blocking trials than the control trials, suggesting they selected the outcome compound featuring the outcome which the cue was previously paired with from Stage 1 (whilst this would yield

correct responses for the blocking trials, this would produce incorrect responses for the control trials as participants were expecting a different outcome on the basis of Stage 1 training). From Trial 3 onward, however, both trials produced a comparable percentage of correct responses. Furthermore, participants appeared to reach asymptote from Trial 6 onward, with a mean percentage of correct responses > 90%.

A 2×8 repeated measures ANOVA was performed on these data with the factors of Outcome type (A-O1O2/B-O3O4 vs C-O1O2/D-O3O4) and Trial (1 – 8). In light of sphericity violations for the factor ‘Trials’ $\chi^2(27) = 124.58, p < .001$ and the interaction between the factors, $\chi^2(27) = 89.89, p < .001$, Greenhouse-Geisser values are reported. The ANOVA revealed a main effect for the factor of Outcome type, $F(1, 30) = 19.36, MSE = .07, p < .001, \eta_p^2 = .39$, a main effect for the factor trial, $F(7, 210) = 51.27, MSE = .05, p < .001, \eta_p^2 = .63$, and a significant interaction between these two factors, $F(7, 210) = 14.55, MSE = .08, p < .001, \eta_p^2 = .33$. To explore the source of the interaction, simple main effects analyses were conducted. These analyses revealed significant differences between the outcomes on Trial 1, $F(1, 30) = 55.18, MSE = .02, p < .001, \eta_p^2 = .65$, and Trial 2, $F(1, 30) = 5.09, MSE = .04, p < .05, \eta_p^2 = .15$, all other trials $p > .05$ (smallest $p = .37$). Overall, these results demonstrate that participants’ learning from Stage 1 transferred to Stage 2. Participants initially produced more correct responses for the blocking trials

than the control trials, however, by Trial 3 onward participants responses were comparable.

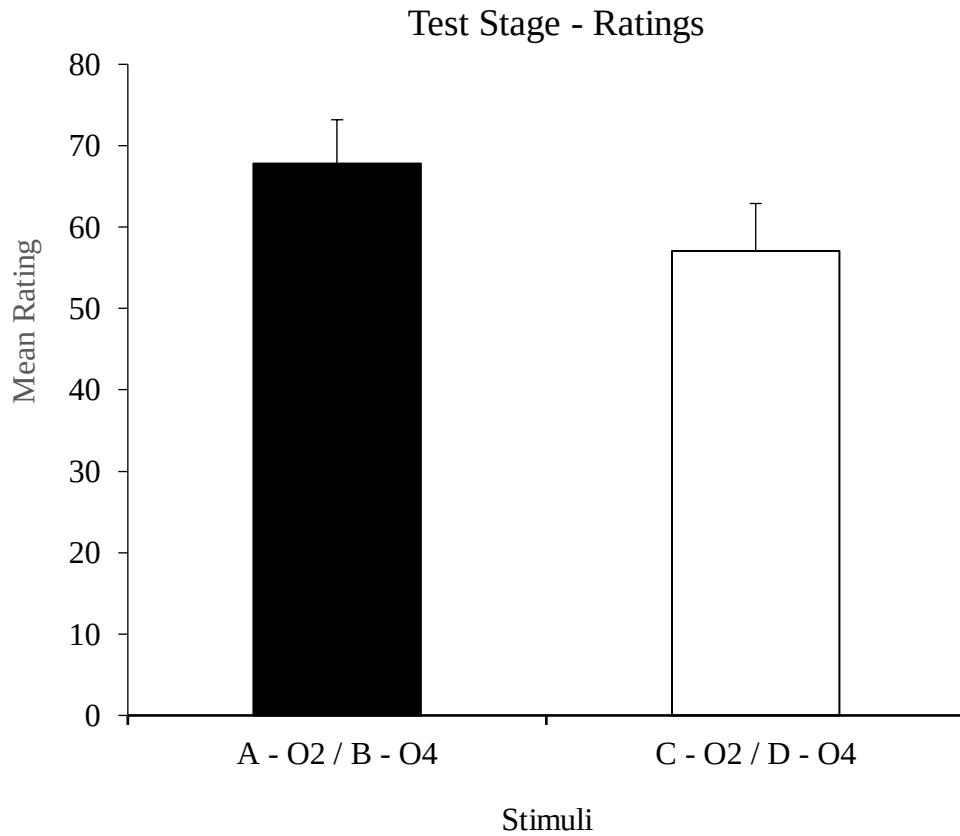


Figure 15. Mean of participants' raw predictiveness ratings for A/B - O2/O4 and C/D - O2/O4 at test. 'Blocked' outcomes (A/B) ratings are greater than the control outcomes (C/D). Error bars represent SEM.

3.1.2.3 Test Data - Ratings

Figure 15 displays participants' raw ratings at test of how predictive cues A – D were for O2 and O4 (the novel outcomes in Stage 2). For the 'blocked' outcomes participants' ratings for A – O2 and B – O4 were calculated by averaging the value of their ratings for these stimulus pairings to form the variable AB – O2/O4. For the control trials, participants' ratings of C – O2 and D – O4 were calculated by averaging the values of their ratings for

these stimulus pairings to form the variable CD – O2/O4. Surprisingly, as can be seen in Figure 15, participants rated the AB – O2/O4 associations higher than the CD – O2/O4 association. A paired samples t-test was performed on these data, which revealed a significant difference between the ratings for these stimulus pairings, $t(31) = 2.60$, $p = .02$, $d = .39$, thus confirming that participants learned more about the A/B – O2/O4 association than the C/D – O2/O4 association. These results therefore conflict with the results of previous studies in animals which have demonstrated blocking amongst elements of an outcome.

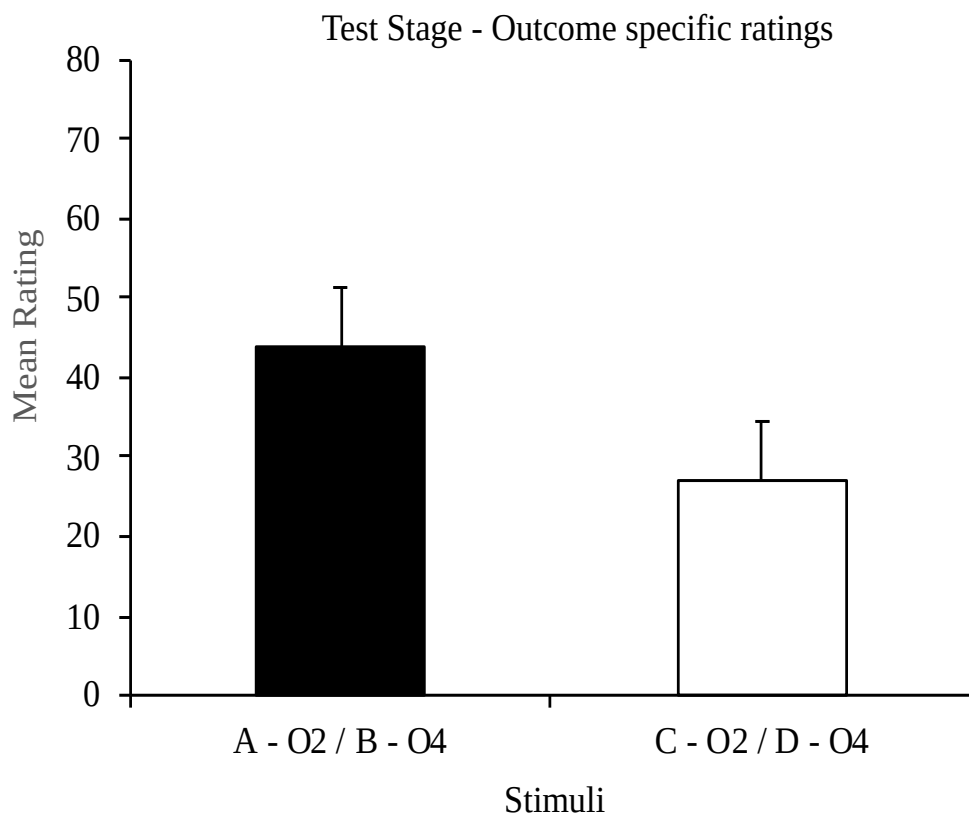


Figure 16. Mean of participant's difference ratings for A/B – O2/O4 and C/D – O2/O4 at test. Difference ratings were calculated by subtracting ratings given to the added outcome which was not paired with the cue, from the added outcome which was paired with the cue (in accordance with Le Pelley and McLaren, 2003). Error bars represent SEM.

3.1.2.4 Test Data (Difference Ratings)

Figure 16 illustrates participants' outcome-specific ratings at test. These were calculated in order to assess whether participants' ratings were higher or lower for the specific outcomes of interest (i.e. O2/O4). It could be, for instance, that participants rated stimuli A/B as being more predictive for any outcome, as opposed to the higher A/B – O2/O4 rating being exclusive to the outcomes of interest. To calculate an outcome specific measure of participants' ratings for cues A – D and O2 and O4, participants' ratings for the novel outcome in Stage 2 which was not paired with the cue of interest, was subtracted from the ratings provided for the novel outcome which the cue was paired with in Stage 2 (in accordance with a method outlined for calculating difference scores employed by Le Pelley and McLaren, 2003). For instance, participants' ratings of A – O4 (the novel outcome in Stage 2 which A was not paired with) were subtracted from participants' ratings of A – O2 (the novel outcome in Stage 2 which A was paired with). As can be seen in Figure 16 whilst the mean ratings for A/B – O2/O4 and C/D – O2/O4 are both lower than Figure 15 (which represents participants' raw ratings), the general trend remained, with participants' providing higher ratings of the AB - O2/O4 association than the CD – O2/O4. A paired samples t-test revealed there was a significant difference between these ratings, $t(31) = 2.06$, $p = .04$, $d = .36$. These results confirm that the effect was outcome specific (e.g. the A – O2 was association was rated highly as opposed to A being rated highly for both novel outcomes in Stage 2).

In summary, Experiment 5 demonstrated that a 'blocked' outcome was learned about better than a control outcome (e.g. an outcome that had been

surprising). When participants experienced Stage 1 training featuring A–O1, B–O3, C–O3 and D–O1, followed by Stage 2 training with A–O1O2, B–O3O4, C–O1O2 and D–O3O4, participants displayed greater learning about A–O2 and B–O4, than to C–O2 and D–O4. These results are inconsistent with those reported by Miller and Matute (1998) and Rescorla (1980), which both report an outcome blocking effect, and those reported by Cobos et al. (2002) and Price and Yates (1995) which reported null effects. In contrast, the results from the current experiment appear to demonstrate a facilitation effect comparable to that reported by Flach et al. (2006). That is, a novel outcome, presented in compound with a previously well learned about outcome, appears to receive enhancements in its associability.

3.2 Experiment 6

Experiment 5 demonstrated that a ‘blocked’ outcome was learned about better than a control outcome which had been surprising. That is, when a novel outcome was presented in compound with a previously well-predicted outcome, participants demonstrated better learning about the novel outcome, relative to a control condition where a novel outcome was presented alongside an outcome that a cue had not been paired with before. These results suggest that pairing a novel element of an outcome with a previously well learned about outcome element facilitates learning about the novel element as opposed to hindering learning. Despite being inconsistent with the results of previous studies – which have displayed outcome blocking (e.g., Miller & Matute, 1998; Rescorla, 1980) – this finding would still appear to pose a problem to key associative learning models. These results indicate that the associative

history of an outcome's components can modify its subsequent associability (although in a different direction to some previous studies), which key associative learning models do not posit. Given the relative novelty of this result, however, it is important to assess the reproducibility of this effect. As such, Experiment 6 sought to provide a replication of Experiment 5. Two notable changes were made, however.

First, in Experiment 6 the nature of the task differed between participants. Half of the participants completed the same version of the task as in Experiment 5 (i.e. a causal judgement task), where the foods served as cues and the reactions served as outcomes. For the other half of the participants, however, the causal order of the events was reversed. That is, the reactions served as cues and the foods served as the outcomes. As such, participants in this group were required to make a diagnostic judgement, where they made a decision about the food Mr. X consumed on the basis of the reaction he received. This was done to assess whether the effect was sensitive to the type of judgement participants were required to make. As noted previously Cobos et al. (2002) tested whether the type of judgement participants made impacted upon the interaction effects observed in their experiments. In their study the type of task participants completed had little influence on the effect observed. However, similar studies have reported conflicting results, for example, a study by Waldmann (2000) found the type of judgement task completed did indeed influence whether or not blocking would be demonstrated amongst cues (also see Van Hamme, Kao & Wasserman, 1993). In Waldmann's experiments participants demonstrated blocking amongst cues when a causal judgment task

was employed, however, when participants completed a diagnostic judgment task a blocking effect was not observed.

In order to account for the absence of blocking when a diagnostic judgement task was employed in these studies, these authors appealed to causal-model theory (see Waldmann & Holyoak, 1992). According to this theory participants are sensitive to the status assigned to events they are presented with first. That is, whether the events are causes of an effect, requiring them to make causal judgements, or whether they are presented with effects which have causes, requiring them to make diagnostic judgements. Causal-model theory predicts that blocking will be observed when participants complete a causal judgement task as participants are exclusively informed that the stimulus they are first presented with (e.g., the cause) can produce this outcome (e.g., the effect) on its own. Furthermore, the relationship between the additional cause and the effect is undefined (i.e., the additional cause is never presented on its own and therefore participants should remain uncertain about its effects).

However, according to causal-model, theory the nature of a diagnostic task means that participants are more likely to consider that there are multiple potential causes of an effect which has been presented first, therefore preventing stimulus competition. Importantly, according to causal-model theory, when presented with an effect first (e.g., sickness) participants are likely to consider many potential causes of the effect. In fact, the task itself may often prompt participants to do precisely this (e.g. ‘what caused x?’). The theory also suggests that when presenting the effect first and then revealing the causes of the effect subsequently, this allows the participants to form

independent relationships about the causal role of the stimuli paired with the event, which they are prevented from doing so with a causal judgement task (see Waldmann & Holyoak, 1992; Waldmann 2000). Furthermore, regardless of whether the effect is provoked by the cause they have conceived of, participants will be cognisant of the fact there are multiple potential causes of the effect and are therefore more receptive to the idea of additional potential causes being presented to them in a subsequent stage.

In contrast to causal-model theory, associative learning theories suggest that the nature of the events (e.g., cause – effect) is of little importance as to whether interaction effects will be observed. Instead, these theories state that the temporal order of events is an important factor which determines whether interaction effects are observed. Interaction effects will be observed in blocking paradigms where multiple cues are presented first, whether they are causes or effects, but not when multiple outcomes (events which are presented following a cue) are presented, regardless of whether they are framed as causes or effects. Given the different assumptions of these theories, Experiment 6 sought to assess whether the type of judgment task completed will have any impact. For instance, if the effect observed in Experiment 5 reflects an associative process between the events, where cues are merely the stimuli which are presented prior to outcomes - regardless of whether the cue was the cause (e.g. food) or the effect (e.g. the reaction) - the same effect should be observed. If, however, the type of judgement participants are required to make (e.g. causal or diagnostic) influences the effect observed, as is suggested by some cognitive models of causality (e.g., Waldmann, 2000), it might be

suggested that differences will be observed dependent on the type of task participants have completed (e.g., Waldmann & Holyoak, 1992).

In addition, a second change to this task was made. In this experiment a greater range of outcomes were provided. In Experiment 5, participants were provided with two outcomes in Stage 1 (O1, O3) and two outcomes in Stage 2 (O2, O4). In Experiment 6, however, participants were presented with four outcomes (O1 – O4) which were each paired with the cues A - D in Stage 1. In Stage 2, a further four additional outcomes (O5 – O8) were presented, with O5 and O6 serving as blocked outcomes, and O7 and O8 serving as control outcomes. This ensured that each cue was paired with a unique outcome. This was done to assess whether using the same outcomes for blocking and control trials in Experiment 5 artificially influenced the effect observed. It is possible, for instance, that when one food predicted one reaction, participants may have deemed that another food was less likely to predict the same reaction (i.e., if A predicts O2, it is less likely that C also predicts O2). It could therefore be argued that participants' ratings are not independent. By separating the trial types with different outcomes, however, we can overcome this issue and gain a clearer picture of the effect size. The design of Experiment 6 can be seen in Table 10.

Table 10. Design of Experiment 6

	<i>Stage 1</i>	<i>Stage 2</i>	<i>Test</i>
Blocking	A – O1	A – O1 O5	A – O5
	B – O2	B – O2 O6	B – O6
Control	C – O3	C – O4 O7	C – O7
	D – O4	D – O3 O8	D – O8

Note. A – D refer to cues, whilst O1 – O8 refer to outcomes. A and B trials represent the blocking trials, whilst C and D trials represent the control trials. A-D were foods for the causal task, with O1-O8 serving as reactions. In the causal task, A-D were reactions and O1-O8 were foods.

3.2.1 Method

3.2.1.1 Participants

Thirty-two participants (28 females; 4 males) were recruited from the University of Nottingham’s School of Psychology. Participants ranged from 18 – 30 years of age ($M = 20.22$; $SEM = 0.52$). All other details are the same as the previous experiment. Participants were randomly allocated (on the condition that there were 16 participants in each group) to either complete either the causal judgment task ($n = 16$) or the diagnostic judgement task ($n = 16$).

3.2.1.2 Apparatus and Stimuli

The same stimuli were used as in Experiment 5 (see Experiment 5, Method). However, four additional foods and reactions were incorporated in this experiment due to the greater number of stimuli employed. The four additional foods were: lettuce, mushroom, pepper and tomato. The four additional reactions were: headache, nausea, no appetite and stomach cramps.



Figure 17. The four additional foods included were pepper, mushroom, tomato and lettuce

The causal judgement version of the task was identical to that employed in Experiment 5 with the inclusion of the additional reactions, thus, unless otherwise stated, the following refers exclusively to the diagnostic judgement task. For this task the reaction – this time serving as the cue – that Mr. X experienced was presented in centre of the screen in capitalized white Arial font (text size: .02), and the foods (which served as outcomes in this task) were presented above a white scale (length: 45cm), which was positioned beneath the reaction. The scale was longer than that in Experiment 5 to due to the additional stimuli. At test each of the reactions were presented individually on the left of the screen, on the right each of the foods were presented. Due to the greater number of stimuli in this experiment the test screen had two stages (this was the same for both the causal and diagnostic judgement task), where participants would rate the respective cue’s ability to predict four of the eight outcomes first before moving onto the other four outcomes. The order of these screens was counterbalanced. The Likert scale displayed to participants when making their rating was identical to that used in Experiment 5.

3.2.1.3 Procedure

As in Experiment 5 both groups experienced 6 blocks of 16 trials in Stage 1, and 2 blocks of 16 trials with each trial type being presented a total of 4 times per block, before proceeding to a test stage. The procedure for Experiment 6 was identical to Experiment 5 for those completing the causal judgment version of the task, except that in this experiment there were a greater number of cue-outcome associations, thus the following refers exclusively to the diagnostic judgement task (unless otherwise stated).

For those completing the diagnostic judgement task, participants were presented with the following alternative cover text:

“In this experiment we would like you to imagine that you are an allergist (i.e. someone who investigates reactions to foods). You have just been presented with a new patient “Mr. X”, who suffers from different types of reactions as a result of eating certain foods. In an attempt to discover the relationship between the reactions Mr. X experiences and the different types of foods he has eaten, you observe the type of reactions he has experienced and try and workout which foods he had eaten.

On the following screens, you will be shown the REACTIONS Mr. X has experienced, and you will be asked to predict which food he had been eaten. Each FOOD will be presented at the bottom of the screen. Make your prediction by selecting one of the FOODS below each of the reactions. You will then be provided with feedback about what FOOD Mr. X had eaten. You will have to guess at first, but with the aid of the feedback your predictions should soon start to become more accurate. The participant with the highest overall accuracy will receive a £10 cash prize.

Please press the SPACE BAR to begin”

Stage 1

Once Stage 1 commenced participants were then presented with the following text at the top of the screen: “Mr. X experienced the below reaction, which food did he eat?” This text was presented above the reaction which was being presented (see Figure 18). To proceed to the next trial participants needed to make a response by selecting a food beneath the reaction (the order of the foods was counterbalanced). Once they had made a response, they were then presented with feedback informing them about the food Mr. X had eaten (see Figure 18).

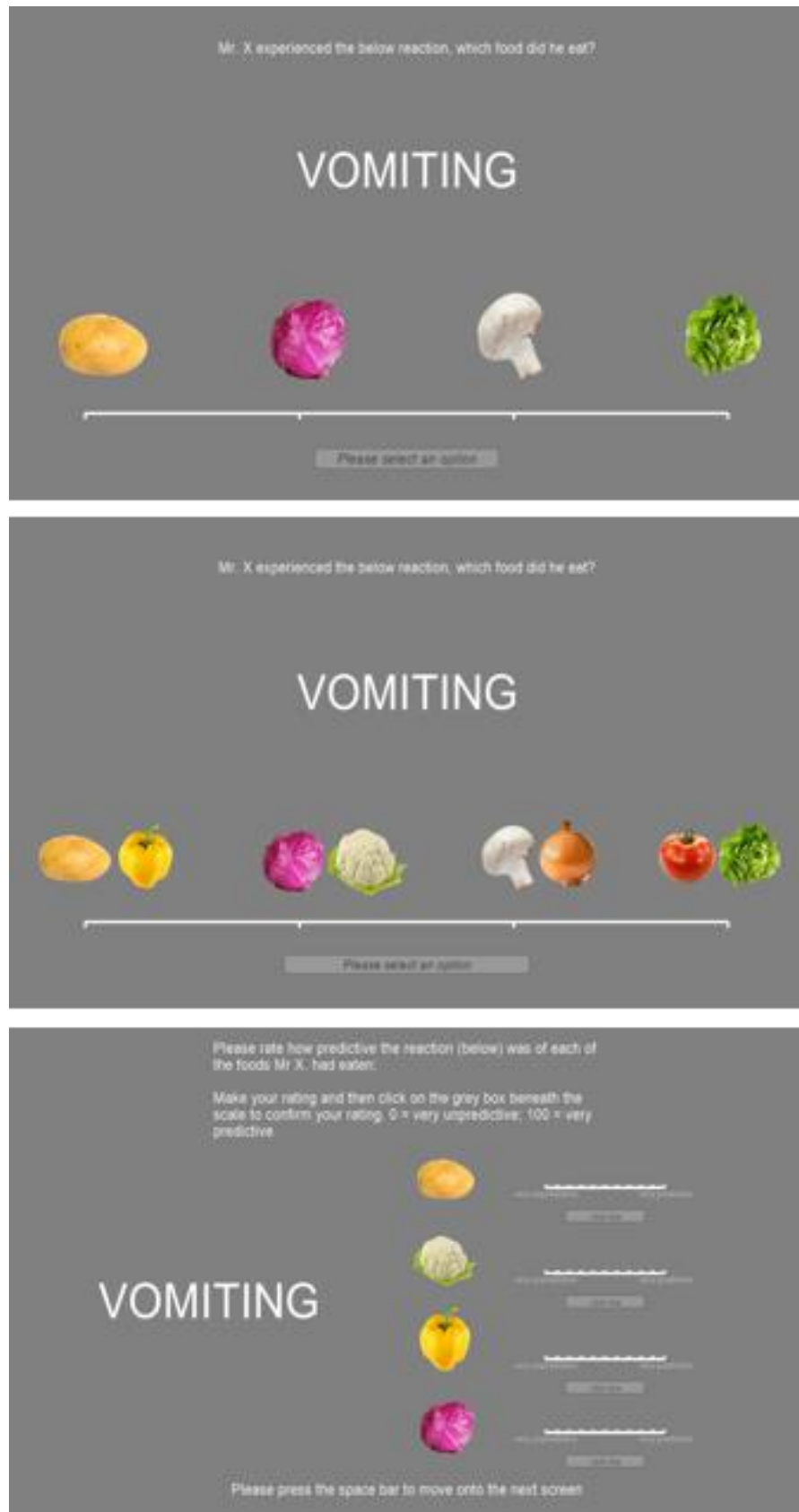


Figure 18. Example screens participants were presented with during Stage 1 (left), Stage 2 (middle) and at test (right) when completing the diagnostic iteration of the task.

Stage 2

Prior to commencing Stage 2 participants were presented with the following text:

“Mr X has started to eat additional foods.

The foods he has eaten will be presented at the bottom of the screen as before.

Please press the SPACE BAR to continue.”

On each trial participants were presented with the text: “*Mr. X experienced the reaction (below), which foods did he eat?*” which was positioned above the reaction which was being presented. However, this time there were additional foods which were presented alongside the Stage 1 foods (e.g. ‘Potato’ from Stage 1, would now be presented with ‘Pepper’ in Stage 2 see Figure 18). The order in which the Stage 2 foods were presented alongside the Stage 1 reactions was counterbalanced across experimental conditions.

Test

Prior to commencing the test stage participants were presented with the following text: “*Please rate how predictive the reaction (below) was of each of the foods Mr X. had eaten. Make your rating and then click on the grey box beneath the scale to confirm your rating. 0 = very unpredictable; 100 = very predictive*”. Each of the reactions were then presented at test, one per screen (see Figure 18).

3.2.2. Results and Discussion

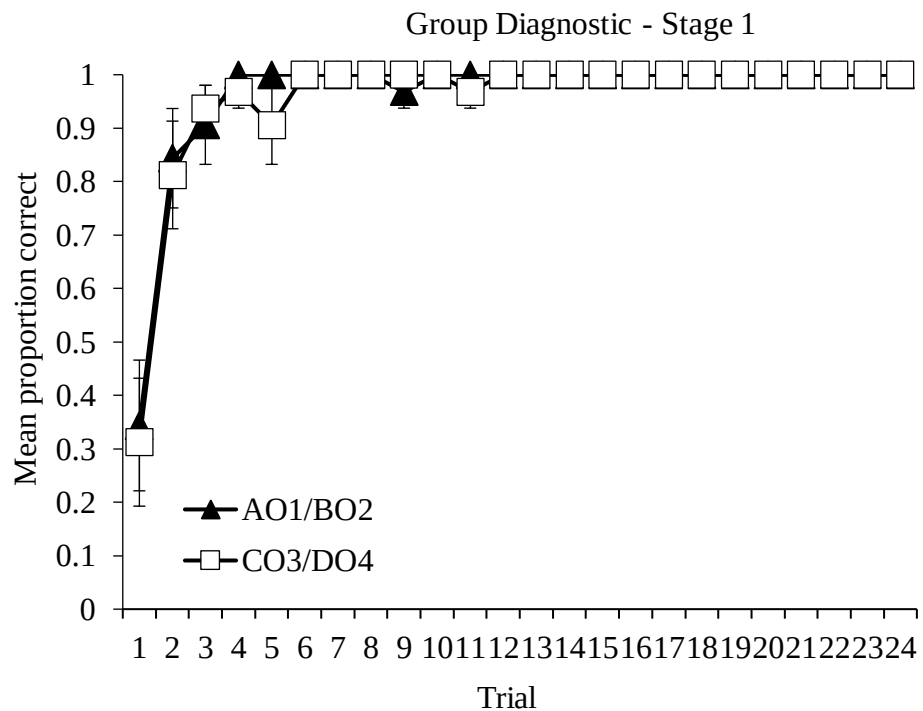
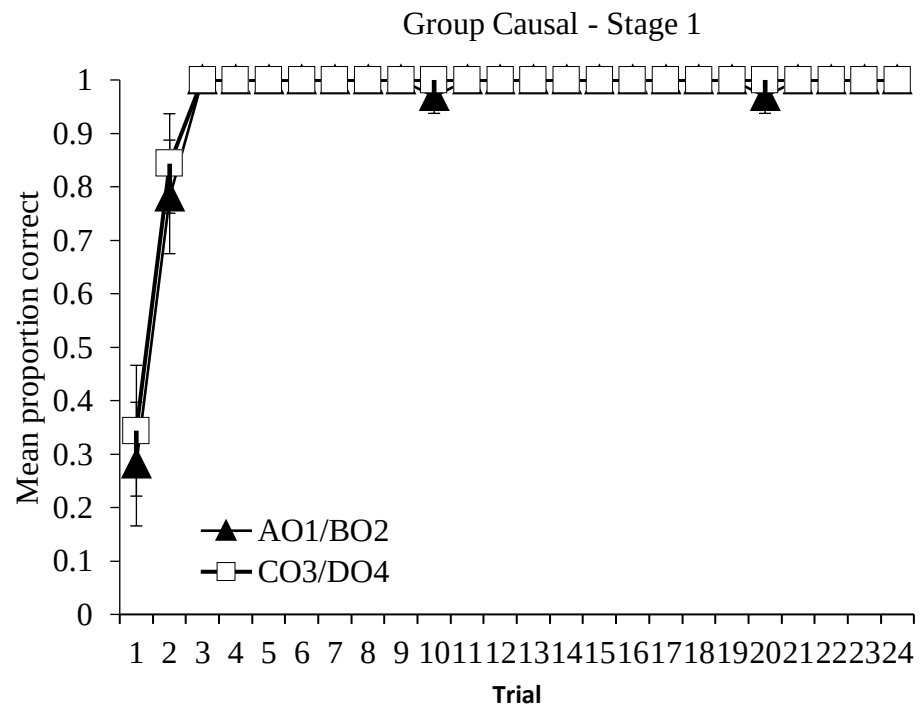


Figure 19. Mean proportion of correct responses for A-O1/B-O2 and C-O3/D-O4. Error bars represent SEM.

Figure 20. Mean proportion of correct responses A-O1/B-O2 and C-O3/D-O4. Error bars represent SEM.

3.2.2.1 Stage 1

Figures 19 and 20 show the mean proportion of correct responses for each cue-outcome pairing (A-O1, B-O2, C-O3 and D-O4) across Stage 1 training. As can be seen regardless of the task completed (e.g., causal/diagnostic) participants displayed the same pattern of results, with participants rapidly learning the relationships between the cue-outcome associations and reaching asymptote from Trial 3 onward. A three-way mixed-model ANOVA was performed with a between-subject factors of task type (Causal vs Diagnostic), and within-subjects factor of outcome (O1/O2 vs O3/O4) and Trial (1 – 24). The ANOVA revealed a main effect of trial, $F(23, 690) = 116.02$, $MSE = .01$, $p < .001$, $\eta_p^2 = .80$, demonstrating that participants learned the associations as training progressed. There was no effect of task type $F(1, 30) = .69$, $MSE = .01$, $p = .41$, $\eta_p^2 = .02$, and no effect of outcome $F(1, 30) = .02$, $MSE = .01$, $p = .90$, $\eta_p^2 = .00$. All interactions were non-significant, largest $F(1, 30) = 1.85$, $MSE = .01$, $p = .18$, $\eta_p^2 = .05$.

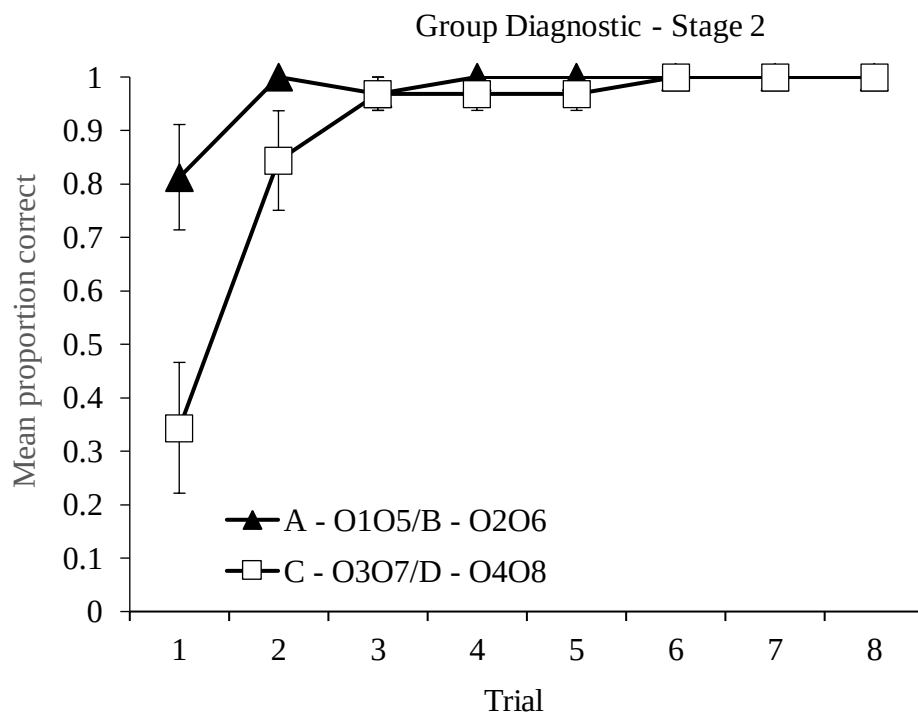
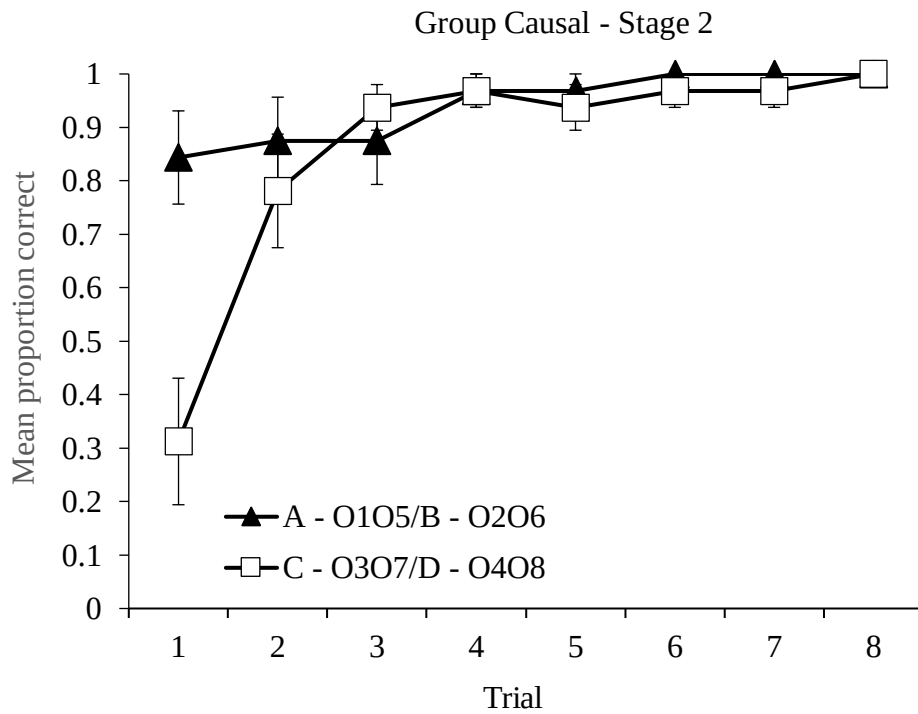


Figure 21. Mean proportion of correct responses for A-O1O5/B-O2O6 and C-O4O7/D-O3O8. Error bars represent SEM.

Figure 22. Mean proportion of correct responses for A-O1O5/B-O2O6 and C-O4O7/D-O3O8. Error bars represent SEM.

3.2.2.2 Stage 2

Figures 21 and 22 demonstrate the mean proportion of correct responses for the Blocking trials (A-O1O5/B-O2O6) and the Control trials (C-O4O7/D-O3O8) across Stage 2 training. These figures demonstrate that - regardless of task type - participant performance was comparable, with participants' learning from Stage 1 transferring to Stage 2. That is, at the outset of Stage 2, participants maintained a high proportion of correct responses for the Blocking trials, however, participants' responses for the control trials were initially low (as expected given that these trials were now paired with different outcomes), before drawing level with the blocking trials by Trial 3.

A three-way mixed-model ANOVA was performed on these data with a between-subjects factor of task type (causal vs diagnostic), and within-subjects factors of outcome, blocked (O1O5/O2O6) vs control (O4O7/O3O8) and Trial (1 – 8). The ANOVA revealed a main effect of the factors Outcome, $F(1, 30) = 31.41$, $MSE = .03$, $p < .001$, $\eta_p^2 = .51$ and Trial, $F(7, 210) = 55.68$, $MSE = .03$, $p < .001$, $\eta_p^2 = .65$ and a significant interaction between these factors, $F(7, 210) = 23.05$, $MSE = .02$, $p < .001$, $\eta_p^2 = .44$. There was no main effect of the factor task type $F(1, 30) = 1.86$, $MSE = .06$, $p = .18$, $\eta_p^2 = .05$, and all other interactions were non-significant, largest $F(7, 210) = .74$, $MSE = .02$, $p = .63$, $\eta_p^2 = .02$. Simple main effects revealed the cause of the Outcome $\times$ Trial interaction. There was an effect of outcome on Trial 1, $F(1, 30) = 48.30$, $MSE = .04$, $p < .001$, $\eta_p^2 = .62$ and Trial 2, $F(1, 30) = 6.08$,

$MSE = .02$, $p = < .05$, $\eta_p^2 = .17$. All other trials were non-significant, smallest $p = .33$. These results demonstrated that at the outset of Stage 2 participants continued to select the outcome compound featuring the outcome which a cue had previously been paired with. However, over the course of training participants corrected their responses to the control trials in light of the new outcomes these stimuli were paired with.

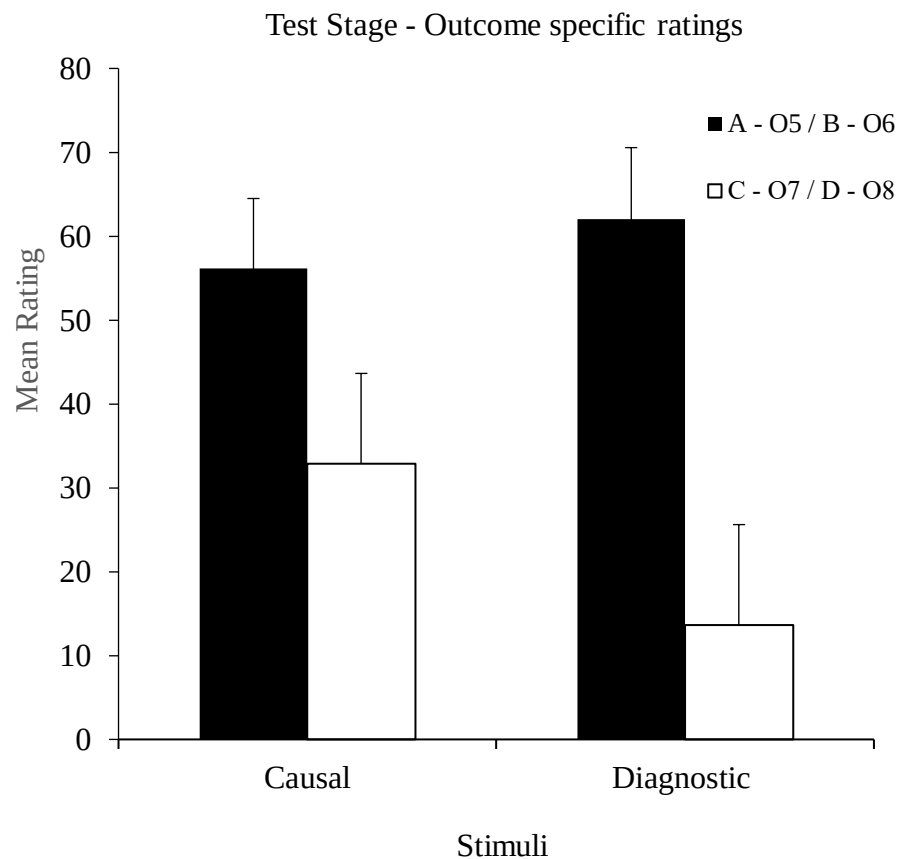


Figure 23. Mean of participants' difference ratings for A/B – O5/O6 and C/D – O7/O8 at test. Difference ratings were calculated by subtracting ratings given to the added outcome which was not paired with the cue, from the added outcome which was paired with the cue (in accordance with Le Pelley and McLaren, 2003). Error bars represent SEM.

3.2.2.3 Test Data

Figure 23 demonstrates participants' outcome specific ratings at test regarding the predictiveness of cues A – D for O5 – O8. Henceforth, outcome specific or difference ratings are used as these are deemed a more stringent test and are the typical measure employed in such tasks (e.g., Le Pelley & McLaren, 2003). To calculate an outcome specific measure of participants' ratings for this task, Le Pelley and McLaren's (2003) method for calculating differences scores was, again, used. That is, the ratings received by the novel outcome which was not paired with the cue but was also a novel outcome which was treated in a comparable manner, was subtracted from the novel outcome which the cue was paired with. For instance, to calculate an outcome specific measure of A-O5 (a blocked outcome), participants' ratings for A-O6 (also a blocked outcome but not paired with A) was subtracted from participants' ratings of A-O5. To further elucidate, to calculate an outcome specific measure of C-O7 (an unblocked outcome), participants' ratings for C-O8 (also an unblocked outcome not paired with C) were subtracted from participants' ratings of C-O7. As can be seen in Figure 23, the pattern of results were in the same direction as Experiment 5, with the differences between the outcomes appearing larger, particularly in the diagnostic group. A 2×2 mixed model ANOVA with a between subjects factor of task type (Causal vs Diagnosis) and a within-subjects factor of outcome (O5/O6 vs O7/O8), revealed a main effect of outcome with participants demonstrating greater learning about the blocked outcome, $F(1, 29) = 18.17$, $MSE = 1096.67$, $p < .001$, $\eta_p^2 = .39$. There was no main effect of task type, $F(1, 29) = .35$,

$MSE = 1995.04$, $p = .56$, $\eta_p^2 = .01$, and no interaction, $F(1, 29) = 2.27$, $MSE = 1096.67$, $p = .15$, $\eta_p^2 = .07$.

In summary, Experiment 6 replicated the results of Experiment 5, providing further evidence that a “blocked” outcome can be learned about better than an unblocked outcome. These results were obtained using a design which differed from Experiment 5. Each cue was paired with a unique outcome in Stage 1 in this experiment, and the nature of the task differed between participants. For half the participants the foods served as cues, whilst the reactions served as outcomes. For others the reactions served as cues whilst the foods served as outcomes. Despite the difference in the nature of the task completed by participants this appeared to have no effect on participants’ performance. That is, participants continued to show a bias in learning toward the blocked outcome regardless of task type, which some studies have found to be an important factor influencing whether interaction effects are observed in cue-blocking paradigms (e.g., Waldmann & Holyoak, 1992; Waldmann, 2000). Furthermore, the results of this experiment also demonstrate that the effect is not dependent on the outcomes being shared between blocking and control trials. In Experiment 5 the same outcomes were used for both blocking and control trials, thus it is possible that participants did not provide independent ratings (i.e., if A predicted O2, they may have deemed C to be less likely to predict O2). However, in this experiment different outcomes were used for blocking and control trials and the same effect was observed.

3.3 Experiment 7

Experiments 5 and 6 demonstrated an outcome facilitation effect. That is, prior learning about an element of an outcome compound resulted in better learning about the added outcome, relative to a control where both outcome elements were surprising. Experiment 6 demonstrated that this effect occurs when causal and diagnostic tasks are completed, suggesting the effect is not sensitive to the type of judgment made. This effect is inconsistent with some previous studies, such as those by Miller and Matute (1998), Rescorla (1980) and Experiment 2 of Flach et al. (2006) which have demonstrated outcome blocking (also Cobos et al., 2002, which demonstrated no effect of an outcome blocking procedure).

One possible explanation of these results could be to do with the type of control employed in the experiments reported here. Notably, a variety of control procedures have been included in the studies mentioned above which is further complicated due to the range of conditioning procedures employed in these studies. A variety of controls have also been employed in cue-blocking studies which has sometimes resulted in different effects being produced (see Blaser, Couvillion & Bitterman, 2006; Guez & Miller, 2008). One of the most common control types though, which has reliably demonstrated blocking, is that reported by Kamin (1968, 1969). In Kamin's studies the control trials simply involved omitting any Stage 1 training and hence exposure to the cue and the outcome and presenting them only in Stage 2.

Given the relative novelty of the outcome facilitation effect, this experiment sought to assess whether outcome interaction effects could be observed with a greater range of control conditions. More specifically, the

additional control used in this experiment was the same as that employed by Kamin but applied to an outcome blocking procedure (henceforth referred to as the Kamin control). The full design of Experiment 7 can be seen in Table 11. Given the increased number of cues in this experiment (as a result of the introduction of Kamin control trials in Stage 2), the number of outcomes was reduced in this task to be consistent with Experiment 5.⁶ The outcome compounds were also the same as those presented on the other trials to ensure the outcomes were not differentially novel to those presented on other trials. In addition, in keeping with Experiment 6 both causal and diagnostic judgement tasks were employed in order to assess whether the additional control trials were sensitive to the type of judgment made.

Table 11. Design of Experiment 7

	<i>Stage 1</i>	<i>Stage 2</i>	<i>Test</i>
Blocking	A – O1	A – O1 O2	A: O1-O4
	B – O3	B – O3 O4	B: O1-O4
Control	C – O3	C – O1 O2	C: O1-O4
	D – O1	D – O3 O4	D: O1-O4
Kamin Control	--	E – O1 O2	E: O1-O4
	--	F – O3 O4	F: O1-O4

Note. A – E refer to cues, whilst O1 – O4 refer to outcomes. A and B represent the blocking trials, C and D represent the control trials, whilst E and F represent the Kamin control trials. A-E were foods for the causal task, with O1-O4 serving as reactions. In the causal task, A-E were reactions and O1-O4 were foods.

⁶ Furthermore, if the same number of outcomes were employed as in Experiment 6, where each cue was paired with a unique outcome in the first stage, this would result in additional Kamin control trials being paired with only two of the four outcome compounds presented, creating a discrepancy in the number of exposures of each outcome compound.

3.3.1 Method

3.3.1.1 Participants

Sixty-four participants (57 females; 7 males) were recruited from the University of Nottingham's School of Psychology. Participants ranged from 18 – 48 years of age ($M = 23.84$; $SEM = 0.76$). Participants were randomly allocated (on the condition that there were a comparable number of participants in each group) to complete either the causal judgment task ($n = 32$) or the diagnostic judgement task ($n = 32$).

3.3.1.2 Apparatus and Stimuli

The six food stimuli were: broccoli, cabbage, mushroom, pepper, potato and tomato. The six reaction stimuli were: diarrhoea, fever, no appetite, skin rash, stomach cramps and vomiting.



Figure 24. The foods included were broccoli, cabbage, mushroom, pepper, potato and tomato.

Both versions of the task were identical to those employed in Experiment 6, with the exception that two additional cues were included, and a reduced number of outcomes were used. As such, at test all the ratings for each were cue made on one screen as opposed to two screens. That is, for the causal judgement task foods were presented individually on the left of the screen, and

each of the six reactions were presented on the right of the screen. In the diagnostic judgement task, each reaction was presented on the left of the screen and each of the six foods were presented on the right of the screen.

3.3.1.3 Procedure

All details were identical to Experiment 6 except that in Stage 2 participants were presented with the additional Kamin control trials. As such, participants were presented with 2 blocks of 24 trials in Stage 2, before proceeding to test. In addition, as noted above, due to the reduced number of outcomes, rather than making their test ratings over two screens, participants made their ratings on one screen, thus being consistent with Experiment 5.

3.3.2. Results and Discussion

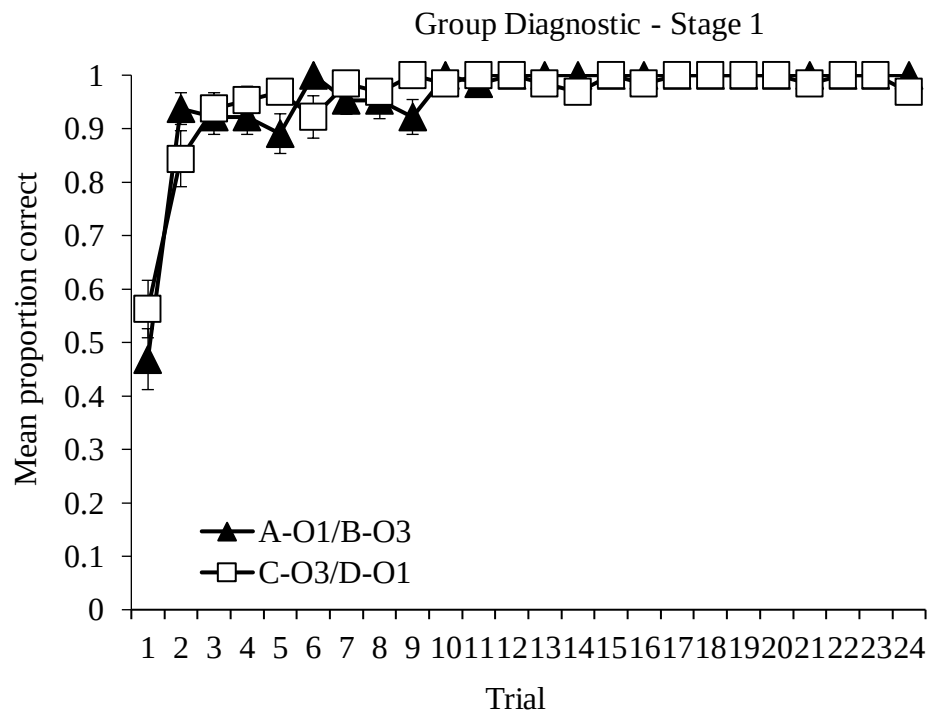
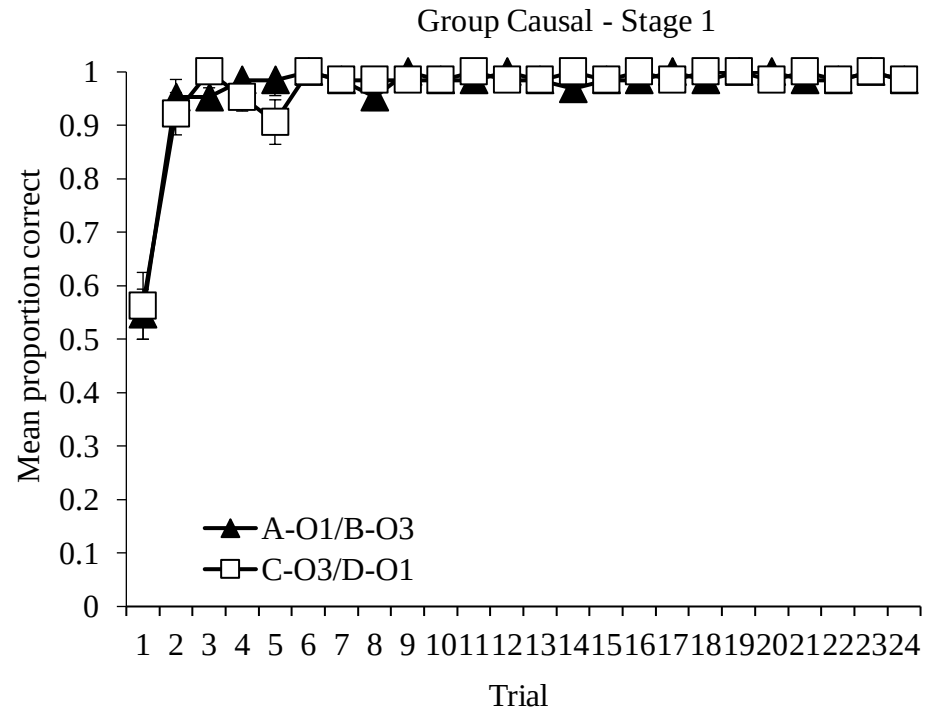


Figure 25. Mean proportion of correct responses for A-O1/B-O3 and C-O3/D-O1. Error bars represent SEM.

Figure 26. Mean proportion of correct responses for A-O1/B-O3 and C-O3/D-O1. Error bars represent SEM.

3.3.2.1 Stage 1

Figures 25 and 26 show the mean proportion of correct responses for each cue-outcome pairing (A-O1, B-O3, C-O3 and D-O1) across Stage 1 training. Again, participants generally displayed the same pattern of results regardless of the task completed (e.g. Causal/Diagnostic), with participants rapidly learning the relationships between the cue-outcome associations. A three-way mixed model ANOVA was performed with a between-subject factors of task type (Causal vs Diagnostic), and within-subjects factor of outcome (O1 vs O3) and Trial (1 – 24). The ANOVA revealed a main effect of trial $F(23, 1426) = 84.98, MSE = .05, p < .001, \eta_p^2 = .58$, demonstrating that participants learned the associations as training progressed. There was no effect of task type $F(1, 62) = .93, MSE = .06, p = .34, \eta_p^2 = .02$, and no effect of outcome $F(1, 62) = .03, MSE = .02, p = .86, \eta_p^2 = .00$. All interactions were non-significant, largest $F(23, 1426) = 1.72, MSE = .06, p = .13, \eta_p^2 = .03$.

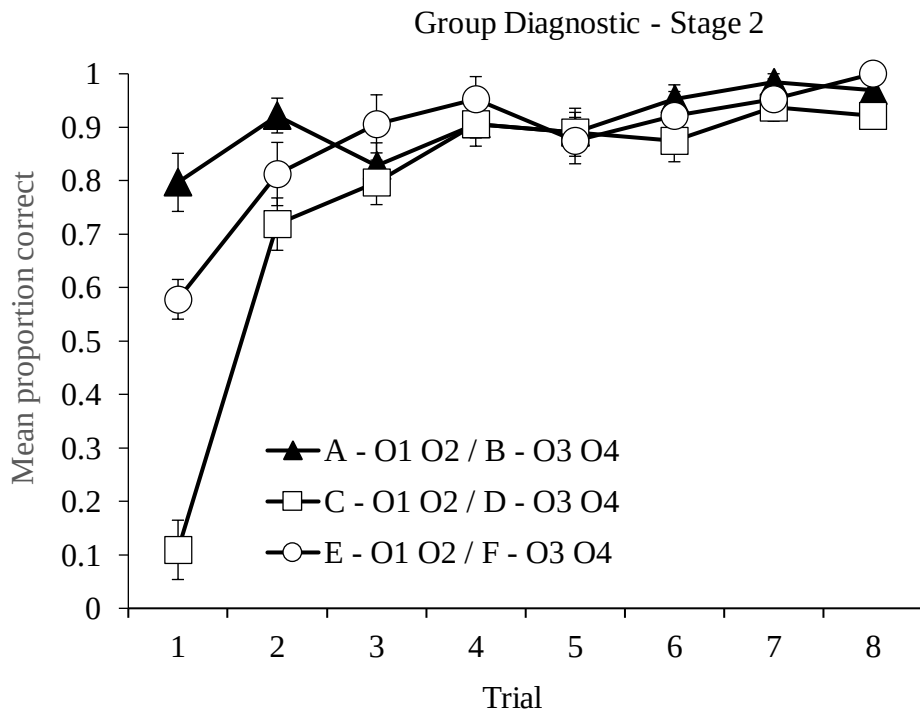
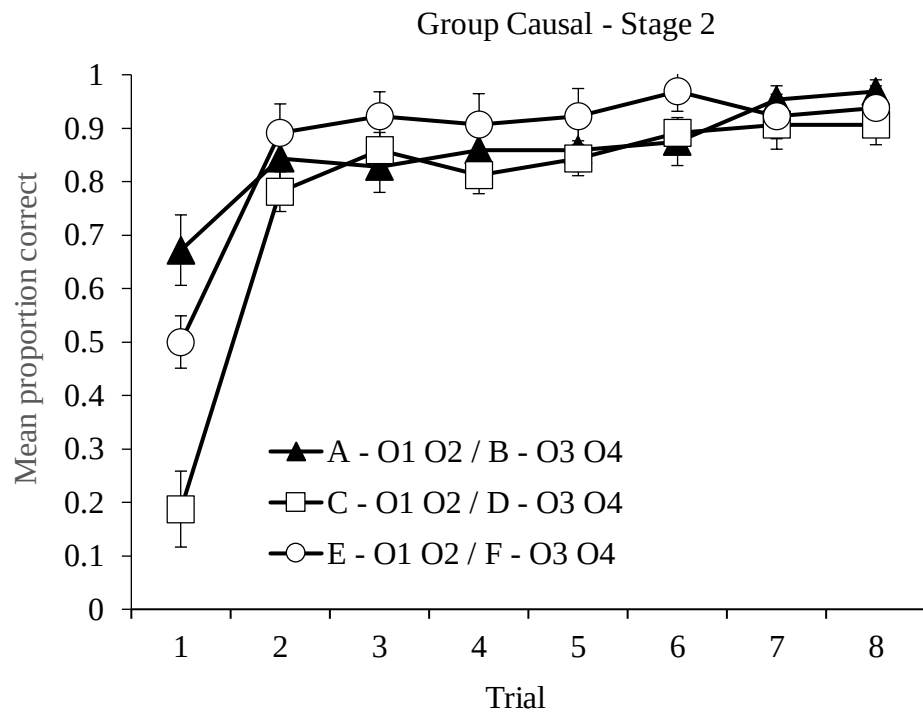


Figure 27. Mean proportion of correct responses for each trial type. Error bars represent SEM.

Figure 28. Mean proportion of correct responses for each trial type. Error bars represent SEM

Figures 27 and 28 demonstrate the mean percentage of correct responses for each cue – outcome compound (A-O1O2/B-O3O4, C-O1O2/D-O3O4, E-O1O2/F-O3O4) during Stage 2 training for both the causal and diagnostic judgment task. As can be seen, participant's performance followed a similar trend regardless of task type. When Stage 2 commenced participants primarily selected the outcome compound which contained one of the outcomes which the respective cue was previously paired with, therefore participants demonstrated a high percentage of correct responses for A and B trials (as these cues continued to be associated with the outcome they predicted in Stage 1). For the C and D trials participants' percentage of correct responses were initially low, as these stimuli were now paired with different outcomes than they had been in Stage 1, before drawing level with A and B by Trial 3. For cues E and F participants mean percentage of correct responses were approximately at chance level, as these stimuli had not been exposed before.

A three-way mixed model ANOVA was performed on these data with a between-subjects factor of task type (causal vs diagnostic), and within-subjects factors of cue-outcome pairing (A-O1O2/B-O3O4 vs C-O1O2/D-O3O4 vs E-O1O2/F-O3O4) and Trial (1 – 8). The ANOVA revealed a main effect for the cue-outcome pairings, $F(2, 124) = 21.26$, $MSE = .09$, $p < .001$, $\eta_p^2 = .26$, and Trial, $F(7, 434) = 96.42$, $MSE = .07$, $p < .001$, $\eta_p^2 = .61$, and a significant interaction between these factors, $F(14, 868) = 12.87$, $MSE = .08$, $p < .001$, $\eta_p^2 = .17$. There was no main effect of the factor of task type $F(1, 62) = .54$, $MSE = .19$, $p = .46$, $\eta_p^2 = .00$, and all other interactions were non-significant, largest $F(2, 124) = 1.15$, $MSE = .10$, $p = .32$, $\eta_p^2 = .02$. Simple

main effects identified the cause of the Cue-Outcome pairing $\times$ Trial interaction. There was an effect of outcome on Trial 1, $F(2, 61) = 48.14$, $p < .001$, $\eta_p^2 = .61$, Trial 2, $F(2, 61) = 3.24$, $p < .05$, $\eta_p^2 = .10$, and Trial 3, $F(2, 61) = 3.30$, $p < .05$, $\eta_p^2 = .10$. All other trials were non-significant, smallest $p = .08$.

To identify where the differences between the factor outcome occurred on these trials, Bonferroni adjusted paired samples t-tests were performed, adjusted $p = .005$ ($3 \text{ stimuli} \times 3 \text{ trials} = 9 \text{ comparisons}$, $.05/9 = .005$). These t-tests revealed that for Trial 1 the mean proportion of correct responses for A-O1O2/B-O3O4 was higher than C-O1O2/D-O3O4, $t(63) = 9.68$, $p < .001$, and E-O1O2/F-O3O4, $t(63) = 3.69$, $p < .001$. The mean proportion of correct responses for E-O1O2/F-O3O4 was also higher than C-O1O2/D-O3O4, $t(63) = 6.69$, $p < .001$. The t-tests comparisons on Trials 2 and 3 were non-significant when considering the Bonferroni correction, smallest $p = .01$. These results demonstrate that at the outset of Stage 2 participants continued to select the outcome compound featuring the outcome which a cue had previously been paired with (except for E/F which had not been presented in Stage 1). However, over the course of training participants corrected their responses to the C and D trials in light of the new outcomes these stimuli were paired with.

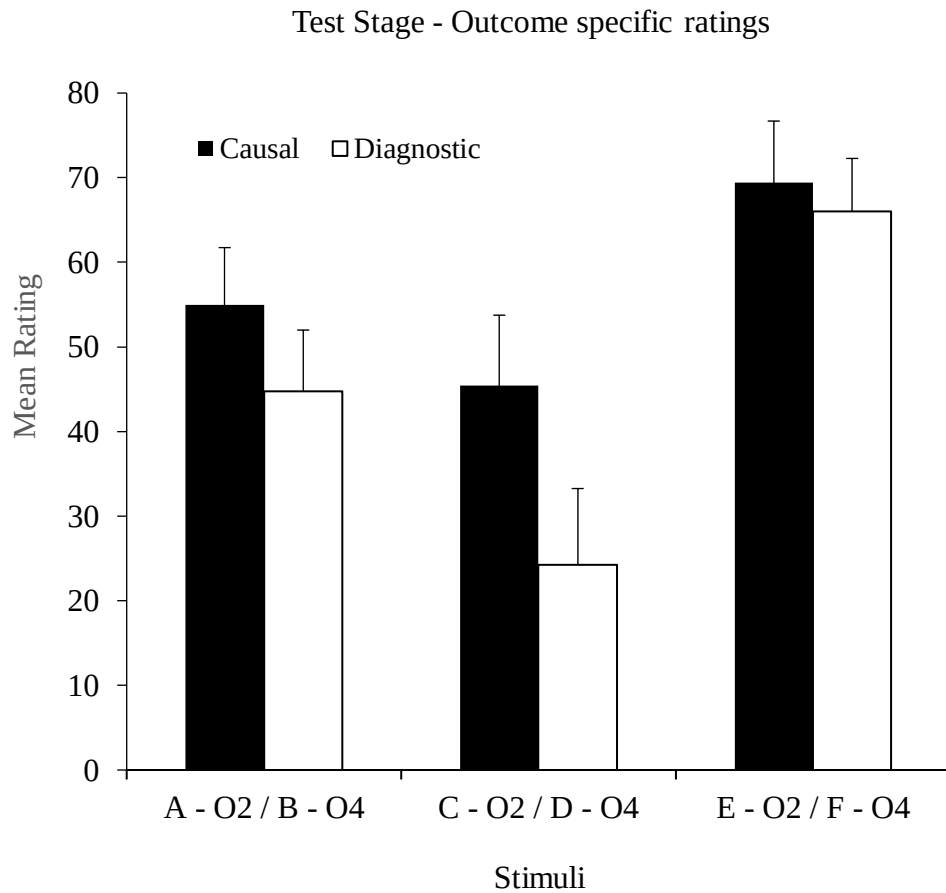


Figure 29. Mean of participants' difference ratings for the outcomes at test, for both causal and diagnostic judgment tasks. Difference ratings were calculated by subtracting ratings given to the added outcome, which was not paired with the cue, from the added outcome which was paired with the cue (in accordance with Le Pelley and McLaren, 2003). Error bars represent SEM.

3.3.2.3 Test Data

Figure 29 shows participants ratings at test of how predictive cues A – F were for outcomes O1 – O4 by task type. As can be seen, participants rated the A/B stimuli higher than the stimuli C/D stimuli, however, the E/F stimuli were rated higher than both A/B and C/D, regardless of task type. A 2×3 mixed model ANOVA with a between-subjects factor of task type (Causal vs

Diagnosis) and a within-subjects factor of cue-outcome pairing (A/B – O2/O4, C/D – O2/O4, E/F – O2/O4), revealed a significant main effect of cue-outcome pairing, $F(2, 122) = 16.08$, $MSE = 1163.43$, $p < .001$, $\eta_p^2 = .20$. There was no main effect of task type, $F(1, 61) = 1.95$, $MSE = 3266.47$, $p = .17$, $\eta_p^2 = .03$. The interaction between these two variables was also non-significant, $F(2, 122) = 1.20$, $MSE = 1163.43$, $p = .30$, $\eta_p^2 = .02$. To further explore the main effect of cue-outcome pairings, Bonferroni adjusted paired samples t-tests were performed (adjusted $p = .017$). These t-tests revealed that ratings for A/B-O2/O4 were lower than ratings for E/F-O2/O4, $t(62) = -3.67$, $p < .001$, but were higher than ratings for C/D-O2/O4, $t(62) = 2.55$, $p < .017$. Ratings for C/D-O2/O4 were also lower than E/F-O2/O4, $t(62) = -5.06$, $p < .001$.

In summary, Experiment 7 revealed that when a Kamin control was employed, an outcome blocking effect was observed. This effect is consistent with the findings of some previous studies, particularly Miller and Matute (1998) who also included a Kamin-type control; however, it is contrary to that reported with the type of control included in Experiments 5 and 6, which resulted in an outcome facilitation effect. Notably, the outcome facilitation effect still remained in the current experiment when the blocking outcomes were compared to the control outcomes employed in Experiments 5 and 6. The fact that there is evidence of both of these effects in the same experiment could suggest that an inconsistency between the two stages, regarding the control trials employed in Experiments 5 and 6, may account for the outcome facilitation effect. For example, in Experiments 5 and 6, the control stimuli reliably predict one outcome in Stage 1. In Stage 2, however, these stimuli no longer predict this outcome and instead predict two different outcomes (thus

ensuring the outcomes are surprising). Employing such a control in an outcome blocking study, however, creates an inconsistency between the two stages of training for blocking trials and control trials, which might render the control stimulus as an inconsistent predictor (P. Jones, personal communication, April 12, 2017). In contrast, on blocking trials the cue continues to be paired with the same outcome in Stage 2 that it is paired with in the first stage of training, thus rendering the cue on a blocking trial as a consistent predictor.

An account of these findings can be provided by the Mackintosh (1975) model. According to this model the control trial employed in Experiments 5 and 6 created an important inconsistency concerning the predictive nature of these trials and the blocking trials. For example, according to the Mackintosh model, in Stage 1 both sets of stimuli (A/B and C/D) have comparable associability as they are both equally predictive of the outcomes. In Stage 2, however, the model stipulates that the associability for cues on the blocking trials will remain high, as they continue to predict an outcome they were previously paired with. In contrast though, the associability for the control trials will begin to decline as these stimuli are no longer paired with the outcome they were previously paired with and instead, predict two different outcomes. As such, given the decreased associability of the control stimuli, the model predicts that these outcomes will be learned about less well, relative to the blocking stimuli.

When employing the Kamin control, however, there is no inconsistency between the two stages, therefore better equating the predictive validity of the Kamin control and blocking stimuli. Furthermore, given the novelty of the cues in Stage 2 for this type of control, it is also plausible to

suggest that the salience of these stimuli is higher than other stimuli presented, therefore accounting for the better learning observed to these cues. In addition, given that the Kamin control only ever predicts the outcome compound, as opposed to the blocking stimuli which predicted an element of the outcome compound in Stage 1, it is possible to suggest that the Kamin control is the most reliable and consistent predictor of the outcomes in Stage 2. This reasoning would further suggest that better learning would be observed about the outcomes on these trials, in keeping with the spirit of the Mackintosh model.

3.4 Experiment 8

Experiment 7 demonstrated that when including an additional control type, in the form of the Kamin control, an outcome blocking effect was observed. That is, prior learning about an element of an outcome appeared to block learning about the added outcome, relative to a control stimulus that had not been presented in Stage 1 and only ever predicted the outcome compound in Stage 2. This effect, although consistent with some previous studies, is contrary to that reported when employing the control trial used in Experiments 5 and 6 in this chapter, which produced an outcome facilitation effect. The Mackintosh model can explain these findings (as noted above), however, an additional theoretical account can also be provided by Wagner's SOP model (1981).

Recall that according to this model each stimulus consists of multiple elements which can occupy one of three different states: a primary state (A1), a secondary state (A2) or an inactive state (I). Before being presented, a stimulus

is in the inactive state where it is unable to attract attention and produce a response. Following the first presentation of a stimulus, however, its elements enter into their A1 activity state, where the stimulus can produce its strongest response, before decaying into the A2 activity state. In the A2 state the stimulus's elements produce a weaker response; during this state the stimulus's elements are considered at the periphery of an organism's focus. Importantly, the model states that stimuli elements can enter into their A2 state by the process of decay or through priming via another stimulus. Recall that the capacity of A2 is deemed to be greater than that of A1, and that stimuli are deemed to decay from the A1 to A2 state at a rapid rate, relative to the rate of decay from A2 to inactive. Furthermore, the state which the elements of a stimulus reside in is key to the type of learning which take place. When the elements of both stimuli are in their A1 state excitatory learning is assumed to take place. If, however, the elements of a cue are in the A1 state, yet the elements of an outcome are in their A2 state, then inhibitory learning is assumed to take place.

This model, like the Mackintosh model, can account for both the outcome facilitation effect and outcome blocking effect observed in the previous experiments. According to this model the facilitation effect can be accounted for by assuming that Stage 1 training with an outcome on a blocking trial (e.g., A – O1) would result in the Stage 1 outcome residing in its A2 state during Stage 2. This therefore allows the novel outcome in Stage 2 (e.g., O2) to fully occupy the A1 state, where greater learning will take place about the association between the cue and the novel outcome (e.g., A – O2). In essence, participants can fully attend to the novel outcome being presented without

devoting any resources to learning about the outcome which was paired with this cue in Stage 1. On a control trial, Stage 1 training with an outcome also ensures that this outcome (e.g., C – O3) resides in its A2 state during Stage 2. However, during Stage 2 training the novel outcome is presented alongside an outcome which the cue had not previously been presented with (e.g., C – O1 O2), therefore both outcomes enter the A1 state and compete for the limited capacity available in this state. As such, participants are required to attend to both outcomes, and therefore each outcome reciprocally prevents the other from being learned about to the same extent as the novel outcome, which is presented on a blocking trial.

The model can also accommodate the outcome blocking effect observed in Experiment 7 when the Kamin control was included. According to the model, better learning is displayed about the Kamin control as the stimuli presented in Stage 1 will, to an extent, be primed into A2 by the context, thus a form of context blocking restricts learning about these stimuli to the same extent as the Kamin control. For the control trials presented in Stage 1, they have the disadvantage of being both primed by the context and competition between the two surprising outcomes in Stage 2 as they compete to enter the A1 state, thus accounting for the lower ratings. On the blocking trials, there is an attenuation of the competition between the elements of the outcome compound, but the cues that signal the outcome do have the issue of being already primed by the context. The Kamin control does not encounter this latter issue as the stimuli are novel and therefore salient. However, they do have the issue of competition amongst the outcomes in Stage 2 as they compete to enter into their A2 state. So long as the effect of the novelty of the

stimuli is greater than the competition between the elements of the outcome, however, then Wagner's theory can provide a full account of the results of Experiment 7.

Both models are therefore able to provide an account of the findings reported in this chapter. Where these models differ, however, is in relation to their predictions as to when the effect should be present. The Mackintosh account predicts that the effects which have been observed are dependent upon the prediction error from the first trial in Stage 2 influencing the associability of cues on the next trial. According to the Mackintosh model the stimuli are fully attended to on the first trial of Stage 2 in a blocking design. On the second trial, however, the associability of the stimuli is updated. For instance, upon first presentation of the stimuli in Stage 2, the associability of the stimuli is comparable. However, following the first trial the associability of the stimuli will be altered based on the information provided in the trial. For example, the associability of the stimuli A/B will increase as these stimuli will be rendered good predictors, along with the novel stimuli E/F. The associability of the C/D stimuli will decrease though as they are rendered poor predictors. Consequently, the effect observed in Experiment 7 should not be evident if only one training trial is presented in stage 2.

The account derived from Wagner's SOP does not make the same prediction. According to this analysis one compound trial in stage two should be sufficient to obtain the effect, as the competition between the elements of the outcome compound and the priming by the context should be evident at this stage. Unlike the Mackintosh model, SOP does not require additional trials for this effect to be apparent. According to the SOP model the interaction

effects between the outcomes should be present after only one trial. For instance, after one trial in Stage 2, the A/B stimuli should prime the outcomes which were previously paired with these stimuli into their A2 state, thus allowing the additional outcomes to occupy more of their A1 state - facilitating learning about them. After one trial the competition between the two outcomes paired with the C/D trials should also be evident (as they compete to enter their A1 state), whilst the novelty of the Kamin controls should overcome any priming by the context of the stimuli exposed in Stage 1. Arguably, these effects will be at their strongest at this stage.

One way of testing these accounts of Experiment 7 is by employing a one-trial outcome blocking procedure, where only one trial is presented in Stage 2. One-trial blocking procedures have been used a number of times in the cue-blocking literature to test various attentional accounts of blocking. Interestingly, studies employing this procedure have produced heterogeneous results. For instance, blocking was not displayed after one-trial in some of these studies (e.g., Dickinson, Nicholas & Mackintosh, 1983; Mackintosh 1975b), however, it has been observed in others (e.g., Gillan & Domjan, 1977; Willner, 1978). This procedure does not appear to have been used to assess outcome interaction effects though. As such, Experiment 8 sought to assess the diverging predictions of the two accounts outlined above by employing a one-trial outcome blocking procedure. The design of this experiment was the same as Experiment 7, with the exception of the number of trials presented in Stage 2. As with the previous experiments both causal and diagnostic versions of the task were included. The full design can be seen in Table 12. According to the Mackintosh model no differences should be observed between ratings about

each of the cue-outcome pairings, as the prediction error from the first trial has not had the chance to be updated. In contrast, according to the SOP model one trial should be sufficient to obtain the effects obtained in Experiment 7.

Table 12. Design of Experiment 8

	<i>Stage 1</i>	<i>Stage 2</i>	<i>No. of Trials</i>	<i>Test</i>
Blocking	A – O1	A – O1 O2	1	A: O1-O4
	B – O3	B – O3 O4	1	B: O1-O4
Control	C – O3	C – O1 O2	1	C: O1-O4
	D – O1	D – O3 O4	1	D: O1-O4
Kamin Control	--	E – O1 O2	1	E: O1-O4
	--	F – O3 O4	1	F: O1-O4

Note. A – E refers to cues, whilst O1 – O4 signalled outcomes. A and B represent the blocking trials, C and D represent the control trials, whilst E and F represent the Kamin control trials. A-E were foods for the causal task, with O1-O4 serving as reactions. In the causal task, A-E were reactions and O1-O4 were foods.

3.4.1 Method

3.4.1.1 Participants

Sixty-four participants (46 females; 18 males) were recruited from the University of Nottingham's School of Psychology. Participants ranged from 18 – 50 years of age ($M = 22.45$; $SEM = 0.78$). Again, participants were randomly allocated (on the condition that there were a comparable number of participants in each group) to complete either the causal judgment task ($n = 32$) or the diagnostic judgement task ($n = 32$).

3.4.1.2 Apparatus, Stimuli and Procedure

The apparatus, stimuli and procedure were identical to that reported in Experiment 7, with the exception that in Stage 2 only one trial of each cue-outcome pairing was presented. As such, Stage 2 only consisted of 6 trials.

3.4.2 Results and Discussion

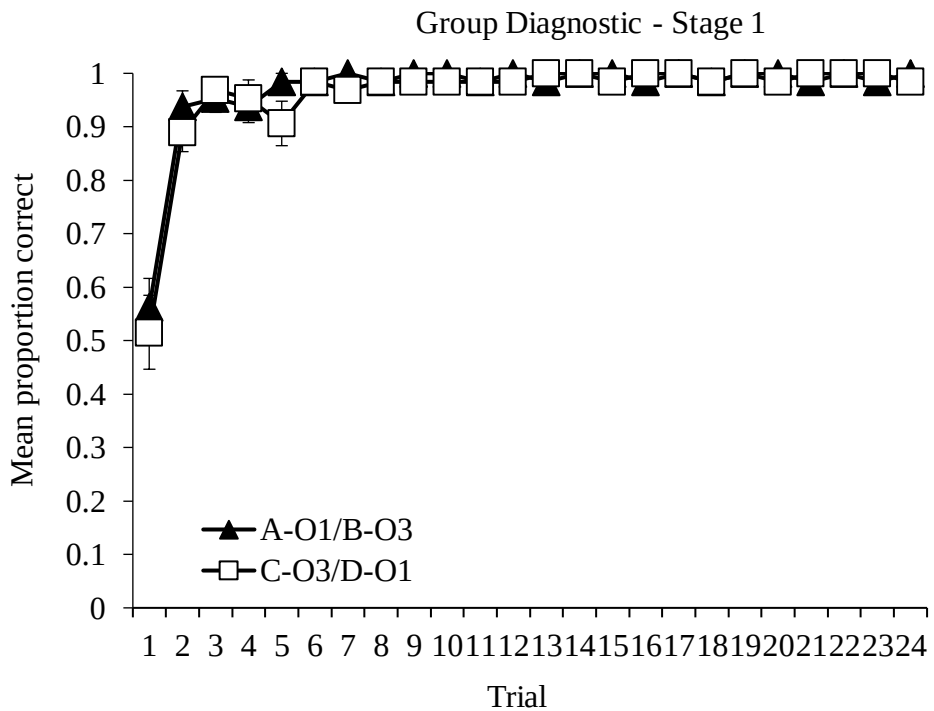
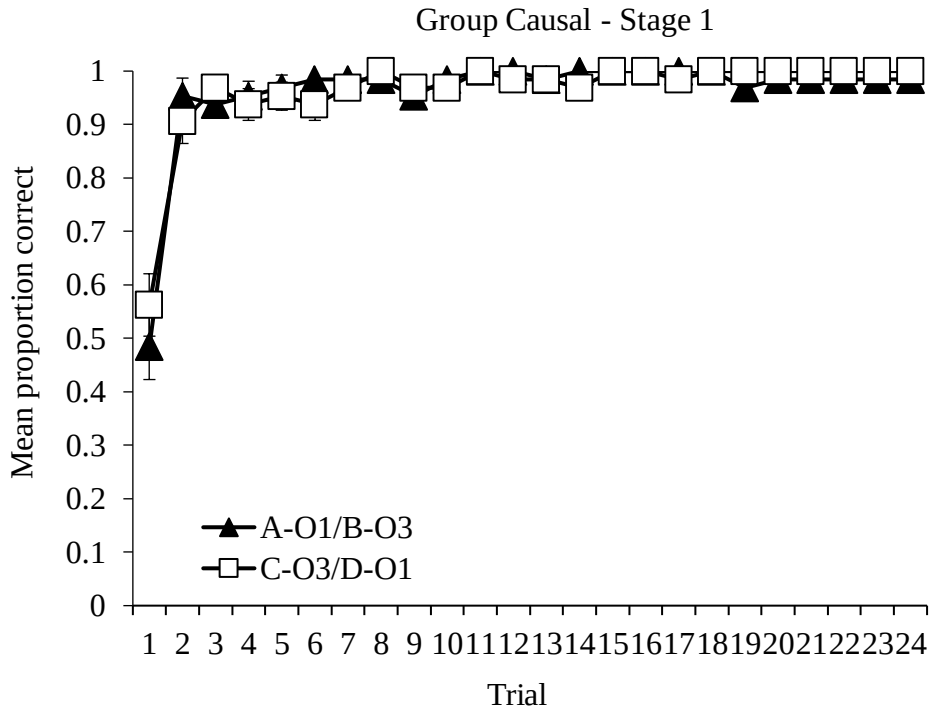


Figure 30. Mean proportion of correct responses for A-O1/B-O3 and C-O3/D-O1. Error bars represent SEM.

Figure 31. Mean proportion of correct responses A-O1/B-O3 and C-O3/D-O1. Error bars represent SEM.

3.4.2.1 Stage 1

Figures 30 and 31 show the mean proportion of correct responses for each cue-outcome pairing (A-O1, B-O3, C-O3 and D-O1) across Stage 1 training for both the causal and diagnostic judgment task. Again, participants generally displayed the same pattern of results regardless of the task completed (e.g. Causal/Diagnostic), with participants rapidly learning the relationships between the cue-outcome associations. A three-way mixed model ANOVA was performed with a between-subject factors of task type (Causal vs Diagnostic), and within-subjects factor of outcome (O1 vs O3) and Trial (1 – 24). The ANOVA revealed a main effect of trial, $F(23, 1426) = 93.26$, $MSE = .01$, $p < .001$, $\eta_p^2 = .60$, demonstrating that participants learned the associations as training progressed. There was no effect of task type $F(1, 62) = .10$, $MSE = .06$, $p = .75$, $\eta_p^2 = .00$, and no effect of outcome $F(1, 62) = .47$, $MSE = .02$, $p = .49$, $\eta_p^2 = .00$. All interactions were non-significant, largest $F(1, 62) = .87$, $MSE = .02$, $p = .35$, $\eta_p^2 = .01$.

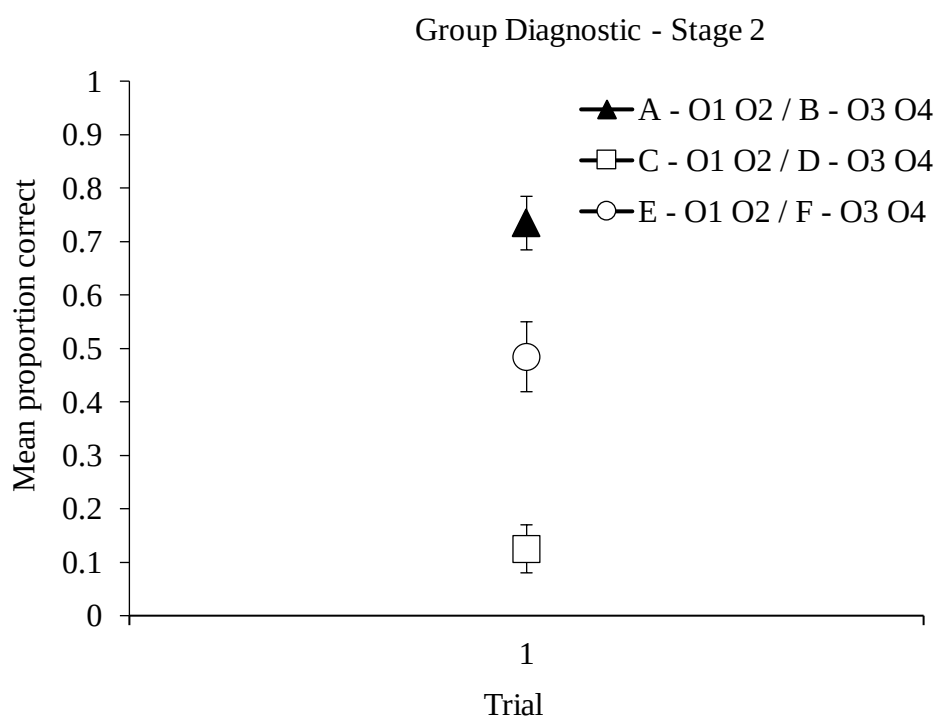
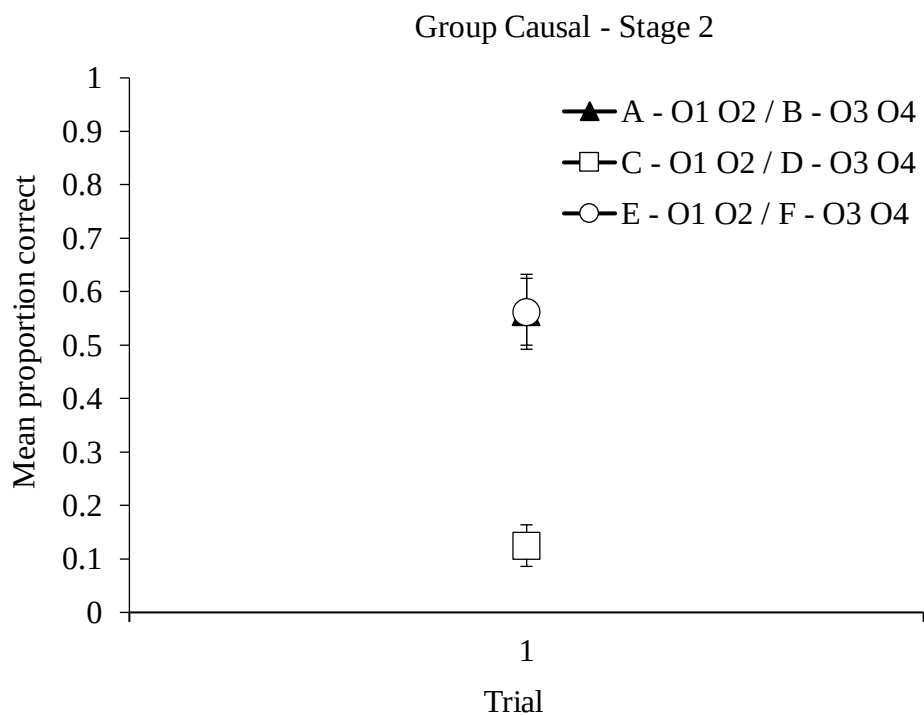


Figure 32. Mean proportion of correct responses for each trial type. Error bars represent SEM.

Figure 33. Mean proportion of correct responses for each trial type. Error bars represent SEM.

3.4.2.2 Stage 2

Figures 32 and 33 demonstrate the mean proportion of correct responses for each cue – outcome compound (A-O1O2/B-O3O4, C-O1O2/D-O3O4, E-O1O2/F-O3O4) during Stage 2 training. As can be seen in Figure 32 in the causal group, the mean proportion of correct responses for the blocking stimuli and the Kamin control were comparable, whilst the proportion of correct responses to the control stimuli which featured in Stage 1 were low. For the diagnostic group participants demonstrated a relatively high proportion of correct responses for the blocking stimuli, and a comparable level of correct responses were displayed to the causal group for the Kamin control and the control stimuli which featured in Stage 1.

A two-way mixed model ANOVA was performed on these data with a between-subjects factor of task type (causal vs diagnostic), and within-subjects factors of cue-outcome pairing (A-O1O2/B-O3O4 vs C-O1O2/D-O3O4 vs E-O1O2/F-O3O4). The ANOVA revealed a main effect for the cue-outcome pairings, $F(2, 124) = 41.63$, $MSE = .12$, $p < .001$, $\eta_p^2 = .40$. There was no main effect of the factor of task type $F(1, 62) = .61$, $MSE = .08$, $p = .44$, $\eta_p^2 = .01$, and the interaction between these factors was non-significant, $F(2, 124) = 2.28$, $MSE = .12$, $p = .11$, $\eta_p^2 = .04$. To identify the source of the main effect for the cue-outcome pairings, Bonferroni adjusted paired samples t-tests were performed (adjusted $p = .017$). These t-tests revealed that the mean proportion of correct responses for A-O1O2/B-O3O4 was higher than C-O1O2/D-O3O4, $t(63) = 8.81$, $p < .001$, but not E-O1O2/F-O3O4, $t(63) = 1.82$, $p = .07$. The mean proportion of correct responses for E-O1O2/F-O3O4 was also higher

than C-O1O2/DO3O4, $t(63) = 7.60, p = < .001$. These results demonstrate that at the outset of Stage 2 participants displayed a higher proportion of correct responses to the blocking and Kamin control trials than to the control trials which featured in Stage 1.

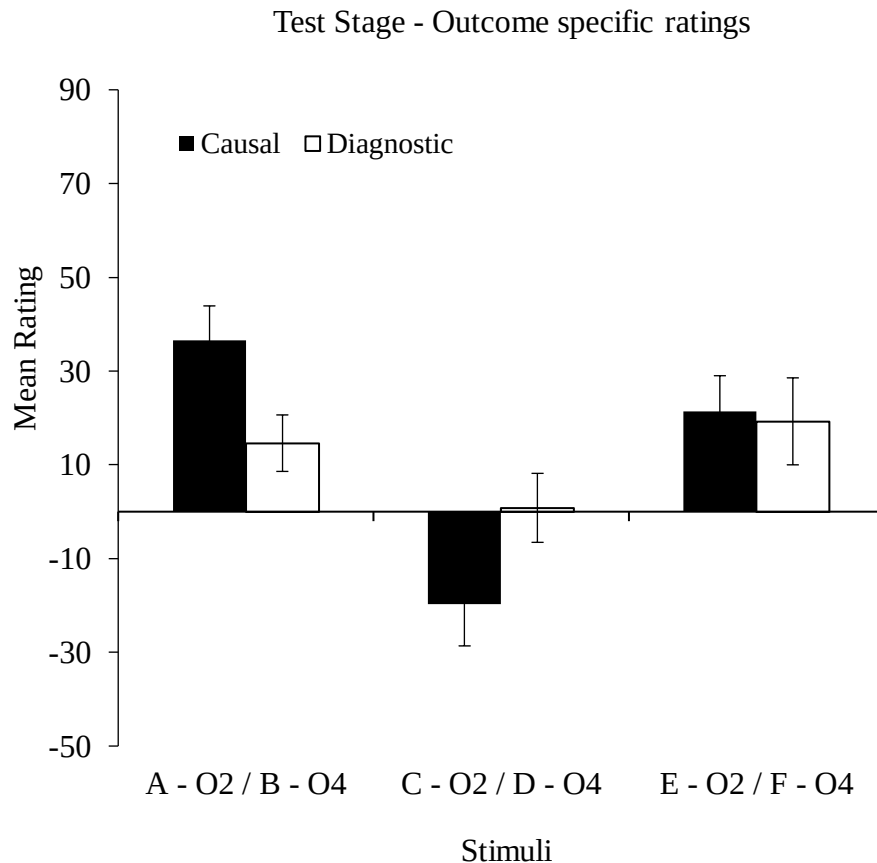


Figure 34. Mean of participants' difference ratings for the outcomes at test, for both causal and diagnostic judgment tasks. Difference ratings were calculated by subtracting ratings given to the added outcome, which was not paired with the cue, from the added outcome which was paired with the cue (in accordance with Le Pelley and McLaren, 2003). Error bars represent SEM.

3.4.2.3 Test Data

Figure 34 shows participants' difference ratings at test of how predictive cues A – F were for outcomes O1 – O4 by task type. As can be seen, in the causal group participants rated the A/B stimuli higher than the C/D

stimuli, and slightly higher than the E/F stimuli. C/D ratings were negative suggesting participants were rating these cues as predictive of the novel outcome that was presented alongside the outcome these stimuli had previously been paired with. However, a one sample t-test was performed comparing the collapsed mean ratings (across task type) of C/D – O2/O4 to 0, to assess whether the ratings were significantly lower than 0. The t-test revealed that participants ratings were not significantly below 0, $t(63) = 1.61$, $p = .11$. In the diagnostic group ratings for C/D were lower than A/B and E/F, with E/F receiving the highest ratings.

A 2×3 mixed model ANOVA with a between subjects factor of task type (Causal vs Diagnostic) and a within-subjects factor of cue-outcome pairing (A/B – O2/O4, C/D – O2/O4, E/F – O2/O4), revealed a significant main effect of cue-outcome pairing, $F(2, 124) = 10.39$, $MSE = 2198.64$, $p < .001$, $\eta_p^2 = .14$. There was no main effect of task type, $F(1, 62) = .04$, $MSE = 1485.43$, $p = .83$, $\eta_p^2 = .01$. The interaction between these two variables was significant, $F(2, 124) = 3.29$, $MSE = 2198.64$, $p < .05$, $\eta_p^2 = .05$. Simple main effects were performed to reveal the source of the interaction. There was an effect of task type for stimuli A/B with participants in the Causal group providing higher ratings than the Diagnostic group, $F(1, 62) = 5.35$, $MSE = 1439.03$, $p < .05$, $\eta_p^2 = .08$, but no effect for the C/D or E/F stimuli, largest $F(1, 62) = 3.16$, $MSE = 2136.36$, $p = .08$, $\eta_p^2 = .05$. To further explore the main effect of cue-outcome pairings, Bonferroni adjusted paired samples t-tests were performed (adjusted $p = .017$). These t-tests revealed that ratings for A/B-O2/O4 were higher than ratings for C/D-O2/O4, $t(63) = 3.93$, $p < .001$,

but not E/F-O2/O4, $t(63) = .65, p = .51$. Ratings for C/D-O2/O4 were also lower than E/F-O2/O4, $t(63) = -3.57, p = < .001$.

In summary, Experiment 8 revealed that after employing a one trial outcome blocking procedure an outcome facilitation effect was present, but not the outcome blocking effect observed in Experiment 7 (with the addition of the Kamin control). These results are inconsistent with the Mackintosh (1975) account of the experiments reported in this chapter. According to the Mackintosh model the effects which have been observed in Experiments 5 - 7 are dependent upon the prediction error from the first trial in Stage 2 influencing the associability of cues on the next trial. As such, any differences should not be present after one trial. However, the results reported here are inconsistent with this account. In contrast they appear to provide partial support to the SOP account.

According to the SOP account both outcome facilitation and blocking should be observed after one trial. This account is partially supported due to the presence of the outcome facilitation effect. However, the absence of the outcome blocking effect is inconsistent with the account derived from the model. To account for the outcome blocking effect observed in Experiment 7 it was suggested that the novelty of the Kamin control cues ensures that they fully enter into their A1 state relative to the other cues which will be primed by the context. At the outset of Stage 2, the novelty of these stimuli is particularly high, as such the effect should arguably be present at this Stage and perhaps at its strongest, yet this effect is not observed in the current experiment. Given that only one trial is presented, however, it might be that more trials are needed

to ensure that participants are confident of the associations involving these stimuli.

It is also worth noting the interaction effect (cue-outcome ratings $\times$ task type), with participants providing higher ratings in the causal task than the diagnostic task for the blocking stimuli A/B. As noted previously, this could be accounted for by assuming that participants are more likely to consider other potential causes when completing the diagnostic task and therefore provided lower ratings in the diagnostic task. As to why it is only observed with blocking stimuli, one possible explanation could be that this is less likely to be an issue with the C/D and E/F stimuli as uncertainty or novelty (respectively) will be more salient factors for these trials in both versions of the tasks, rather than employing causal or diagnostic reasoning, particularly given that participants are only presented with one trial. McLaren et al. (2018) suggest that under circumstances of uncertainty associative processes take precedence over cognitive reasoning (e.g., the Perruchet effect, 1985). This could possibly suggest associative processes are informing ratings for the C/D and E/F stimuli rather than cognitive reasoning at this stage, but that this is not the case for the stimuli A/B which continue to predict the same outcome, thus the influence of task type impacts ratings on these trials.

3.5 General Discussion

Four experiments explored whether previous training with an outcome (e.g., A – O1) blocked learning about a novel outcome (e.g., O2), when these two outcomes were presented in compound (e.g., A – O1 O2). In Experiment 5 participants displayed a bias toward learning about the novel element of an

outcome, when presented in compound with an outcome which a cue had previously predicted, relative to a control where both elements of the outcome compound were surprising. As such, previous training with an element of an outcome appeared to facilitate learning about the novel outcome, as opposed to block learning about the novel outcome. Experiment 5 demonstrated this in a causal judgment task where participants were first presented with foods (i.e., cues) which predicted certain reactions (i.e., outcomes). Experiment 6 replicated this effect when participants completed either a causal judgment task (which followed the same design as Experiment 5) or a diagnostic judgment task where participants were first presented with reactions (i.e., cues) which were caused by foods (i.e., outcomes).

Experiment 7 included an additional type of control, consistent with that included by Kamin (1968). With the addition of this control (termed the Kamin control) an outcome blocking effect was observed alongside the outcome facilitation effect observed in the previous experiments. Experiment 8 sought to tease apart two theoretical accounts of the results reported in Experiment 7 – derived from the Mackintosh (1975) model and Wagner's (1981) SOP model – by employing a one-trial outcome blocking procedure. Although both accounts can accommodate the results reported above, they differ as to when these effects should manifest. According to the Mackintosh model any effect should only be present after two trials or more (when the associability of the stimuli can be updated based on the prediction error from the first trial). According to the SOP model one trial should be sufficient to produce the effect. In Experiment 8 an outcome facilitation effect was

observed, following one trial of training in Stage 2, but not an outcome blocking effect, providing partial support to the SOP account.

Given the heterogenous nature of the existing literature, these results are both consistent and inconsistent with previous studies which have explored outcome blocking. Notably, the outcome facilitation effect reported in Experiments 5-8 is inconsistent with previous animal studies such as Rescorla (1980) and Miller and Matute (1998), which demonstrated an outcome blocking effect (and also Price and Yates, 1995, who demonstrated no effect). Yet they are consistent with Experiment 3 of Flach et al. (2006). These authors demonstrated a comparable effect in a response priming task, when the pretrained outcome element was an auditory stimulus and the added outcome was a visual stimulus. However, a number of procedural differences exist between the aforementioned studies and the experiments reported in the chapter, particularly the animal studies. Firstly, these studies employed indirect conditioning procedures with animals which could potentially influence results. Indeed, Rescorla (1980) noted that as a result of employing a second order conditioning procedure embedded within an autoshaping task with biologically significant outcomes, it could be that the pigeons in this experiment were simply distracted by the ‘blocking’ outcome element, as opposed to a form of associative competition occurring between the elements of the outcome; although this issue is circumvented by Miller and Matute (1998) by employing a sensory pre-conditioning paradigm.

Differences also exist between the aforementioned studies and the experiments reported here, in terms of species, the experimental task completed, and the type of control employed. Perhaps most importantly, the

experiments reported in this chapter highlight how different controls can result in different effects being produced, as the introduction of the Kamin control in Experiment 7 resulted in an outcome blocking effect consistent with some of the studies cited above. Two possible theoretical accounts can account for why these different types of effects are observed, depending on the control employed. According to the Mackintosh (1975) model the outcome facilitation effect is observed due to an inconsistency in the predictiveness of the blocking and control stimuli included in Experiments 5 and 6, whilst the outcome blocking effect occurs due to the absence of this inconsistency and the novelty of the Kamin control. According to SOP, Stage 1 learning ensures that the respective outcome elements are in their A2 state, thus allowing the added outcomes on the blocking trials to fully occupy their A1 state (as they continue to predict the outcome from Stage 1) but not the control outcomes (as both outcome elements are novel and thus compete to enter into their A1 state), therefore producing an outcome facilitation effect. To account for the outcome blocking effect observed with the addition of the Kamin control, it could be derived from the model that better learning occurs to these stimuli, as the stimuli already presented in Stage 1 are primed by the context. Experiment 8 provided partial evidence for the SOP model as an outcome facilitation effect was observed but not outcome blocking when employing a one-trial outcome blocking procedure (according to the model both effects should be present after one trial). However, the findings of Experiment 8 provide less support to the Mackintosh account, according to which no differences should be present after only one trial of Stage 2.

It is worth noting, however, that the outcome facilitation effect in Experiment 8 may indeed be accounted for by an alternative explanation to that suggested by the SOP model. For instance, a within-compound association explanation might also account for the results of this experiment. To elaborate, in Experiment 8 the C/D control stimuli reliably predicted an element of an outcome compound over a number of trials in Stage 1, which they are no longer paired with in Stage 2, during which only one trial is presented. In fact, the outcome these stimuli were previously paired with is now presented in compound with another novel outcome which they have not previously experienced. Given the discrepancy between the number of trials participants will have been exposed to with those outcomes in Stage 1 and Stage 2, it is possible a within-outcome compound association was formed in Stage 2, between the outcome which the C/D stimuli were previously paired with and the added outcome element, thus biasing participants toward rating the stimulus as predictive of an outcome it was not paired with in Stage 2 (hence accounting for participants' mildly negative rating for this control in Experiment 8).

To further clarify, for example, the stimuli which are paired together most in Stage 2 are the outcome compounds (i.e., O1 O2 and O3 O4). In fact, there are more outcome-compound pairings than cue-outcome pairings in stage 2. Given that only one trial is presented in Stage 2 for each of the cue-outcome pairings, this may be insufficient for effective learning to take place about the relationship between the cues and the added outcomes. In light of this, it is possible that Stage 1 learning is mediating performance at test. Thus, when stimuli C/D are presented at test, this provokes participants to think of the

outcome these stimuli were paired with in Stage 1, which in turn, associatively activates the outcome which it was presented in compound with during Stage 2. For instance, if stimulus C was paired with O3 in Stage 1, and O3 was presented in compound with O4 in Stage 2, it is possible that upon being presented with stimulus C at test, participants rate C - O4 higher than C - O2, which C was paired with. This would not be a problem for stimuli A/B, as these stimuli are paired with their Stage 1 outcomes, which are also presented in compound with the added outcome which A/B are paired with in Stage 2. As such, it is possible that a within-compound association account might explain the positive ratings stimuli A/B receive. The ratings for the E/F stimuli are, however, difficult to accommodate according to this account though. Given that these stimuli are not presented in Stage 1, there is no opportunity for these stimuli to form associations with a Stage 1 outcome which could associatively activate an outcome from Stage 2. It is possible, however, that the novelty of these stimuli ensures that some learning takes place between these cues and their outcomes, but that less learning takes place than the A/B stimuli due to the absence of the within-compound association activation that A/B can benefit from. It is also worth noting though that this explanation would not account for the presence of the outcome facilitation effect in Experiments 5 – 7.

Interestingly, despite the contradictory nature of the effects produced in these experiments when employing different controls, both types of effects challenge the assumption of associative learning models which assume that outcomes do not interact and compete with one another in a comparable manner to cues. Furthermore, although the outcome facilitation effect observed

with a one-trial outcome blocking could be accounted for by assuming that within-compound associations may account for this effect, this assumption itself suggests that outcomes can influence one another in a manner most models would assume is typically confined to cues. Moreover, the observation of both facilitation and blocking in outcomes could indeed highlight a further similarity between the behaviour of cues and outcomes when presented in compound. Although competition effects are often observed when cues are presented in compound (e.g., blocking, overshadowing), facilitation effects have also been observed under certain circumstances (e.g., augmentation, potentiation), as noted in the General Introduction. Indeed, Urcelay (2017) notes that cue-interaction effects occur on a spectrum ranging from competition, to no effect, to facilitation.

An alternative explanation of the findings can also be provided, that appeals to an inconsistency in the number of associates between the blocking and control stimuli. For example, across the course of the experiment the control stimuli C/D are associated with a greater number of outcomes than the blocking stimuli. It is therefore possible that this disparity in the number of associates could result in reduced learning about the outcomes associated with the C/D control stimuli, relative to the blocking stimuli. For instance, Anderson (1974) observed that when a stimulus accrues a greater number of associates, participants display slower response times when tasked with identifying an associate of this stimulus, thus displaying the “fan effect”. It is thus possible this effect could account for the outcome facilitation results provided here. However, this would not account for the outcome blocking

effect which was observed with the Kamin control given this control type has a comparable number of associates to the blocking stimuli.

It also worth considering the impact on these results of the type of judgement task completed. According to associative learning theories the type of judgement task completed should have little effect. The key factor which will influence results is the temporal order compound stimuli are presented in: cues refer to the stimuli which are presented prior to an outcome; outcomes are the stimuli which are presented after a cue. According to causality models developed to account for human reasoning, the type of task completed should have an effect. For instance, causal model theory (Waldmann & Holyoak, 1992) assumes that participants are aware of the direction of causality when completing causal judgement-based tasks and can decipher between causes and effects regardless of which order they are presented in. Moreover, they make different judgements based on whether a cause of effect is presented first. The model assumes that when participants are presented with a diagnostic judgement task where they are presented with an effect first and then presented with the causes, they are more likely to consider alternative causes which could have accounted for this effect.

The results of these experiments provide slight support to this model. For instance, an effect of task type was found in Experiment 8 with participants in the causal group providing higher ratings to the blocking stimuli than the diagnostic group. However, collectively the results demonstrate little impact of the type of task completed. That is, whether or not an outcome interaction effect was observed was not influenced by the type of judgement task completed. Nevertheless, it is important to note that the experiments

reported in this chapter involved relatively simple associative tasks which arguably prevented participants from engaging in rational inferences regarding potential alternative causes or effects produced by the stimuli presented.

In summary, the experiments in this chapter examined whether prior learning about an element of an outcome can block learning about a novel outcome when these are presented together. An assumption common to many associative models of learning is that the associative history of an outcome should have little impact on the learning of a novel outcome when these two outcomes are presented together. The current studies question this assumption, providing evidence of outcome facilitation and outcome blocking effects dependent on the control employed. These experiments provide further insight into the nature of outcome associability and outcome interaction effects and present some factors which should be considered when studying outcome blocking.

Chapter 4: Outcome Overshadowing and Outcome Relative Validity

As noted in the General Introduction (Chapter 1), when two or more cues are presented in compound a range of interaction effects can be observed based on the experimental procedure employed. Although a large body of literature has focused on the blocking effect, where an element of a compound is presented prior to the compound, this is not the only cue interaction effect to be reported. Competition effects have been reliably reported when employing relative validity (e.g., Van Hamme & Wasserman, 1994; Wagner, Logan, Haberlandt, & Price, 1968; Wasserman, 1990) and overshadowing procedures (Gluck & Bower, 1988; Pavlov, 1927; Price & Yates, 1993) where only a single stage of training is necessary.

In relative validity tasks compounds of cues are presented to participants. For some of the trials in this task, referred to as the True-discrimination (TD) trials, participants are presented with compound cues (AX, BX) which consist of a distinctive element (A, B) and a common element (X). Each distinctive cue will either reliably predict the presence (AX+) or absence (BX-) of an outcome, resulting in the common element (X) being paired with the outcome 50% of the time. For the remaining trials, referred to as Pseudo-discrimination (PD) trials, there is no distinctive cue which reliably predicts the presence or absence of the outcome, that is, two cue compounds are paired with the outcome 50% of the time (CY-/+ , DY-/+). What is typically observed when employing these tasks, is better learning about the strength of the association between the common cue and the outcome following the

Pseudo-discrimination trials than the True-discrimination trials. Explanations of this effect appeal to the relative validity of the accompanying stimulus (e.g., Wagner, Logan, Haberlandt, & Price, 1968). That is, on true-discrimination trials the common stimulus is presented in compound with a more effective predictor (A); as such the common stimulus is a poorer predictor relative to the distinctive element and therefore receives a decrement in its associability. However, this is not the case on the pseudo-discrimination trials as the cue stimuli are all equally predictive of the outcome.

In overshadowing tasks, a cue element and a cue compound each reliably predict an outcome and are presented in a single stage of training. Although both the element and the compound each reliably predict an outcome, the cue element typically enters into a stronger association with an outcome than either an element of the cue compound (i.e., unilateral overshadowing; see Mackintosh, 1976), or both elements of the cue compound (i.e. reciprocal overshadowing, see Sánchez-Moreno, Rodrigo, Chamizo, & Mackintosh, 1999). Accounts of this effect posit that the elements of the cue compound compete to enter into an association with the outcome, either due to differences in the saliences of the cue elements (i.e., Mackintosh, 1975; Pearce & Hall, 1980) or the limited amount of processing supported by the outcome (e.g. Rescorla & Wagner, 1972).

In a similar manner to blocking, both relative validity and overshadowing have been reliably demonstrated in cues (e.g., Gluck & Bower, 1988; Pavlov, 1927; Price & Yates, 1993; Van Hamme & Wasserman; Hamme & Wasserman, 1994; Wasserman, 1990) with many of the models and explanations used to account for blocking also being able to accommodate

these effects. However, much less research has explored the presence of these interaction effects in outcomes. Again, it is likely that this is partially due to the assumption of key associative learning models which typically posit that outcomes do not interact with one another in the same manner as cues, thus assuming an absence of these effects (see General Introduction for further discussion). Nevertheless, there is some evidence to suggest that these effects may indeed be present in outcomes, particularly overshadowing.

As noted in the Chapter 1, Burger, Mallemat and Miller (2000) observed outcome overshadowing in rats using a conditioned lick suppression task embedded within a backward conditioning procedure. In this experiment the CSs were presented as subsequent events following the presentation of a US (i.e., a 1.0 mA, 0.5-sec foot-shock). In an overshadowing group, rats were presented with the US, followed by a CS compound consisting of a white noise and a flashing light. In a control group the rats were presented with the US followed by a single CS (e.g., white noise). During a test stage where the rats were presented with the white noise, rats demonstrated less lick suppression in the overshadowing group relative to a control group. It is worth noting, however, that given that a backward conditioning procedure was employed, the CSs were presented as the subsequent events as opposed to the US, and thus it is at odds with most standard conditioning procedures (i.e., it is defined as an outcome overshadowing based on the temporal order of events).

Arcediano, Matute, Escobar and Miller (2005) also employed an outcome overshadowing procedure using a food allergist paradigm similar to that reported in Chapter 3. In this study these authors employed more traditional conditioning methods with humans, as opposed to the backward

conditioning procedure employed by Burger, Mallema and Miller (2000). In a single stage of training either one food was presented which reliably predicted a single reaction (which served as a control), or a food which reliably predicted a compound of reactions. At a test stage, participants provided ratings regarding the strength of the association between each of the foods and the reactions. In this study participants provided higher ratings to the single outcome than the elements of the outcome compound, thus demonstrating an outcome overshadowing effect.

To account for this effect these authors appeal to the comparator hypothesis theory (Denniston, Savastano & Miller, 2001; Miller & Matzel, 1988). According to this theory stimuli do not compete to enter into associations with one another in the manner assumed by most associative learning theories (i.e., during a learning stage). Instead, all stimuli which are presented together in a contiguous fashion during learning will enter into relationships with one another and are therefore learned about. It is at the testing stage, however, that competition between the stimuli occur. More specifically, competition between stimuli will occur when the participant is required to make a judgement about the strength of the association between a stimulus and an outcome. This theory also stipulates that the nature of test question is critical in determining whether competition will occur, with questions framed in a competitive manner necessary to produce a competition effect (e.g., which of the stimuli reliably predicted 'x'). If, for instance, the test question simply requested participants to rate whether the two stimuli were paired with the outcome, the model suggests that competition effects will be absent. Critically, in interpreting the results of their study, Arcediano et al.

(2005) state that the competition between stimuli which may be displayed at a test stage – according to the comparator hypothesis theory – is not confined to causes or effects, or the temporal order of these events, thus enabling the model to account for interaction effects observed in outcomes regardless of task type (i.e., causal or diagnostic). Furthermore, the authors also note that although they used an overshadowing procedure, the competition effect they displayed should generalise to outcomes when blocking and relative validity tasks are employed, given that all three effects are assumed to arise due to the same processes (i.e., competition during learning or at testing).

Evidence from relative validity studies, however, has been heterogenous. One such study to consider is that by Van Hamme, Kao and Wasserman (1993) which demonstrated cue relative validity in humans. These authors employed a food allergist task where participants were presented with combinations of foods (i.e., AX, BX) which were differentially correlated with the presence of an allergic reaction. In total there were 5 different levels of stimulus correlation (.00, .25, .50, .75, 1.00). When the stimulus correlation was .00, both stimuli compounds (i.e., AX, BX) were paired with the outcomes 50% of the time. In the condition where the difference in correlation between the stimuli was 1.00, one stimulus compound reliably predicted the outcome (i.e. AX) whilst the other stimulus (i.e. BX) reliably predicted the absence of the outcome. In a causal judgement version of the task participants in the .00 condition demonstrated comparable ratings about X to A and B. In the 1.00 condition participants ratings about X were lower than A, therefore demonstrating that the relative predictive ability of the accompanying stimulus influenced learning about stimulus X. However, in a diagnostic version of the

task (where foods served as outcomes and reactions served as cues) the ratings of X were unaffected by the relative predictive ability of stimulus A, therefore highlighting the influence of task type (thus supporting causal model theory but not the comparator hypothesis theory). Importantly, however, the temporal order of the events was not manipulated, thus the standard relative validity effect in cues was displayed.

Table 13. Adapted design of Matute, Arcediano & Miller (1996), Exp. 2 and 3

<i>Group</i>	<i>Training</i>	<i>Test</i>
Outcome relative validity	A + O1 O3 and A – O2 O3	A: O1 - O3
Control group	A +/- O1 O3 and A +/- O2 O3	A: O1 - O3

Note. A + = cause (medicine) consumed; A – = cause not consumed; A +/- = cause consumed in 50% of cases. O1, O2 and O3 = three effects (i.e., syndromes). O1 = Huxley syndrome; O2 = Lindsay syndrome; O3 = Hamkaoman syndrome. Indexing of potential causes and effects changed for clarity.

Only one such study, by Matute, Arcediano and Miller (1996), appears to have directly tested whether relative validity can appear when multiple outcomes are presented following a single cue (see Table 13). In the study of interest, the authors employed a medical decision-making task where participants were presented with a medicine and were tasked with identifying the extent to which this medicine caused three possible fictitious syndromes (i.e. Huxley syndrome [O1], Lindsay syndrome [O2] and Hamkaoman syndrome [O3]) presented in compounds of two. In a control condition the medicine (A) was presented 50% of the time prior to two outcome compounds (O1 O3, O2 O3). In an experimental condition consumption of the medicine would reliably predict one outcome compound (O1 O3) but not the other

outcome compound (O2 O3). In Experiment 2 of their study the authors observed that participants' ratings of the strength of the association between the medicine (A) and the common outcome (O3) were comparable in the control and experimental condition. That is, the absence of relative validity was observed amongst outcomes.

In Experiment 3 of their study, however, participants provided higher ratings to the common outcome (O3) in the control condition than the experimental condition, when a change to the test question was made which according to the authors induced competition between the outcomes. The potential impact of test question will be discussed further in the Chapter 5. As such, the results of the above study appear to suggest that outcome relative validity can be observed amongst outcomes under certain procedures. That said, however, it is important to note that in this study the causes and outcomes were presented simultaneously in a list like manner, thus it is difficult to ascertain whether participants did indeed consider the syndromes as outcomes. According to associative learning theories events which are considered outcomes must unambiguously follow the presentation of a cue. In the study above, it is possible that participants considered the task as diagnostic and therefore treated the syndromes as cues of the medicine which was taken.

In light of the above, the experiments in this chapter sought to assess whether outcome interaction effects can be obtained using both outcome relative validity and outcome overshadowing procedures, when employing the experimental task used in Chapter 3. If outcome overshadowing can be observed in a comparable manner to cues, as has been reported in the literature, it would be anticipated that better learning would be displayed toward an

outcome element, as opposed to the outcome element which is presented in compound with an additional outcome. If outcome relative validity can be observed amongst outcomes in a manner comparable to cues, it would be anticipated that the common outcome on the pseudo-discrimination trials would be learned about better than the common outcome on the true-discrimination trials. If, however, a facilitation effect is observed comparable to that observed when an outcome blocking procedure was employed, it would be anticipated that better learning would be displayed toward the common outcome on the true-discrimination trials. Given the effect of task type in Van Hamme et al. (1993), where relative validity was observed in cues in a causal task but not a diagnostic task, the experiments in this chapter, like Chapter 3, also sought to assess if task type had an impact.

4.1 Experiment 9

Experiment 9 again comprised a variant of the food allergist task employed by Le Pelley et al. (2005), in which participants were asked to imagine themselves as a food allergist tasked with identifying which foods caused certain allergic reactions. In this experiment, however, participants completed an outcome overshadowing task. That is, participants were presented with cues which reliably predicted a single outcome element, alongside cues which predicted a compound outcome. In accordance with the previous chapter the direction of causality was also tested with both causal judgement and diagnostic judgement tasks being employed. As such, participants were either informed about the foods which a fictitious patient

(‘Mr. X’) had consumed and then presented with information about the reaction the patient experienced (i.e., the causal task), or they were presented with the reaction a patient experienced and then informed about the foods the patient had eaten (i.e., the diagnostic task). In light of the experiments reported in Chapter 3 where interaction effects were observed in an outcome blocking design, there is reason to suggest that an interaction effect will be observed in an outcome overshadowing task, particularly given previous literature. For instance, Arcediano et al. (2005) noted the presence of outcome overshadowing in their study. There is also reason to suggest that overshadowing may occur in outcomes based on the theoretical account derived from the SOP model to account for the findings of Chapter 3. According to Wagner’s (1981) SOP model, outcome overshadowing would arise due to both outcome elements of the outcome compound entering the A1 state and competing for the limited capacity available in this state, thus resulting in poorer learning than a cue which predicted a single outcome.

Table 14. *Design of Experiment 9*

	<i>Training</i>	<i>Test</i>
Overshadowing	A – O1 O2	A: O1 – O4
	B – O3 O4	B: O1 – O4
Control	C – O2	C: O1 – O4
	D – O4	D: O1 – O4

Note. A – D refers to cues, whilst O1 – O4 signalled outcomes. A and B trials represent the Overshadowing trials, whilst C and D trials represent the control trials.

The design of Experiment 9 can be seen in Table 14. As both causal and diagnostic versions of the task were used the stimuli which served as cues (i.e., A-D), were either represented by foods in the causal version of the

task, or reactions in the diagnostic version of the task. The stimuli which served as outcomes (i.e., O1-O4) were reactions in the causal task and foods in the diagnostic task. In Stage 1, four cues (A – D) were presented alongside four outcomes (O1 – O4), as outlined in Table 14 and participants were tasked with learning the associations between these stimuli. Two of the cues reliably predicted outcome compounds, respectively A – O1 O2 and B – O3 O4, whilst two cues predicted outcome elements, C – O2 and D – O4. Trials A and B represented the outcome overshadowing trials as outcome compounds were presented on these trials, whilst trials C and D served as control trials as single outcomes were presented on these trials. At a subsequent test stage, participants were presented with a ratings screen which required them to provide ratings concerning how predictive the cues (A - D) were of each of the outcomes. If an outcome overshadowing effect is present, it would be expected that participants' ratings to A – O2 / B – O4 would be lower than their ratings to C – O2 / D – O4, as a result of O2 and O4 being overshadowed by the additional outcomes present on the A/B trials (e.g., O1/O3). For example, O1 would prevent O2 from entering into as strong as an association with A, relative to the C – O2 association. If the A – O2 / B – O4 ratings are comparable to the ratings for C – O2 / D – O4, this would suggest the absence of an interaction effect when employing an outcome overshadowing procedure.

4.1.1 Method

4.1.1.1 Participants

Thirty-two participants (21 females; 11 males) were recruited from the University of Nottingham's School of Psychology. Participants ranged from 18 – 28 years of age ($M = 20.84$; $SEM = .56$). All other details were the same as Experiment 8, with an equal number of participants in both Group Causal ($n = 16$) and Group Diagnostic ($n = 16$).

4.1.1.2 Apparatus and Stimuli

Each of the foods and reactions were assigned to the letters and outcomes in Table 14 using a Latin-Squared counterbalancing technique. The same stimuli were used as in Experiment 8. However, since some outcomes were presented in compound and some in element form (the order of which was counterbalanced), this resulted in the scale response options appearing different during the training trials. All other details for both the causal and diagnostic versions of the tasks were identical to that employed in Experiment 8. For the causal judgment version of the task, participants were presented with a food in the centre of the screen (see Figure 35) and the reactions presented beneath a white scale, thus there were four options beneath the scale which participants could select from (O2, O1 O2, O4, O3 O4). For the diagnostic version of the task participants were presented with the reaction in the centre of the screen and foods were presented in compound form beneath the white scale.

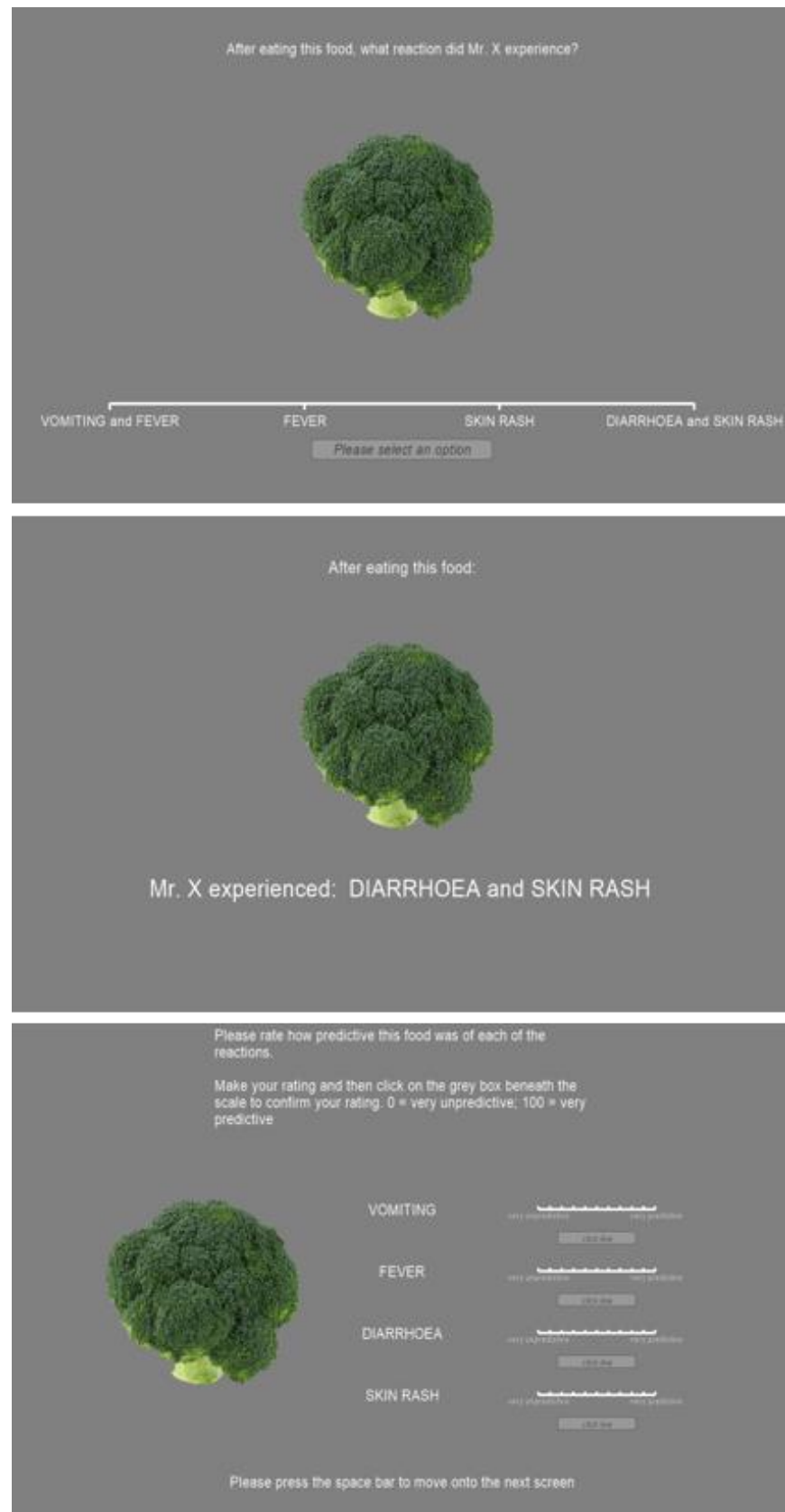


Figure 35. An example of each screen participants would be presented with during Stage 1 (top), the response they received (middle) and at test (bottom).

4.1.1.3 Procedure

Participants completed two blocks of training (with each block comprising 4 trials). All other details were identical to that of Experiment 8. Prior to starting the experiment participants were presented with the following cover text:

“In this experiment we would like you to imagine that you are an allergist (i.e. someone who investigates reactions to foods). You have just been presented with a new patient “Mr. X”, who suffers from different types of reactions as a result of eating certain foods. In an attempt to discover the relationship between the reactions Mr. X experiences and the different types of foods he has eaten, you observe the type of FOODS / REACTIONS⁷ he has EATEN / EXPERIENCED and try and workout which foods he had eaten.

On the following screens, you will be shown the FOODS / REACTIONS Mr. X has EATEN / EXPERIENCED, and you will be asked to predict which REACTIONS / FOODS he had EXPERIENCED / EATEN. Each REACTION / FOOD will be presented at the bottom of the screen. Make your prediction by selecting one of the REACTIONS / FOODS at the bottom of the screen. You will then be provided with feedback about what FOOD / REACTION Mr. X had EATEN / EXPERIENCED. You will have to guess at first, but with the aid of the feedback your predictions should soon start to become more accurate.”

⁷ The text either read FOOD or REACTION, hence the /. The first option always refers to the causal version of the task, whilst the second options always refers to the diagnostic version of the task.

Training Stage

Once the training stage commenced participants were then presented with the following text at the top of the screen when completing the causal version of the task: “Mr. X has eaten the following food, what did he experience?”. To move onto the next trial participants made a response by selecting a reaction/s (in word form) beneath the white scale. Once they had made a response, they were then presented with feedback informing them about the food Mr. X had eaten. In the diagnostic version participants were presented with the following text at the top of the screen: “Mr. X has experienced the following reaction, what did he eat?”. As the foods served as outcomes in this version of the experiment, the images of foods were presented beneath the white scale. Once participants had made their choice and received feedback on all trials, they then moved onto the Test Stage. The test stage was identical to Experiment 8.

4.1.2 Results and Discussion

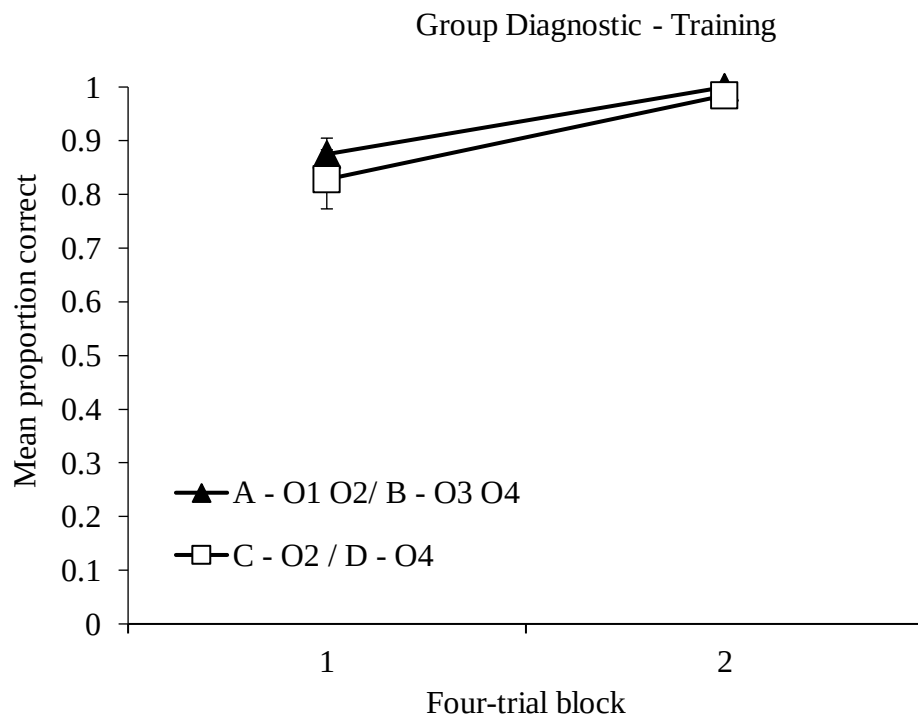
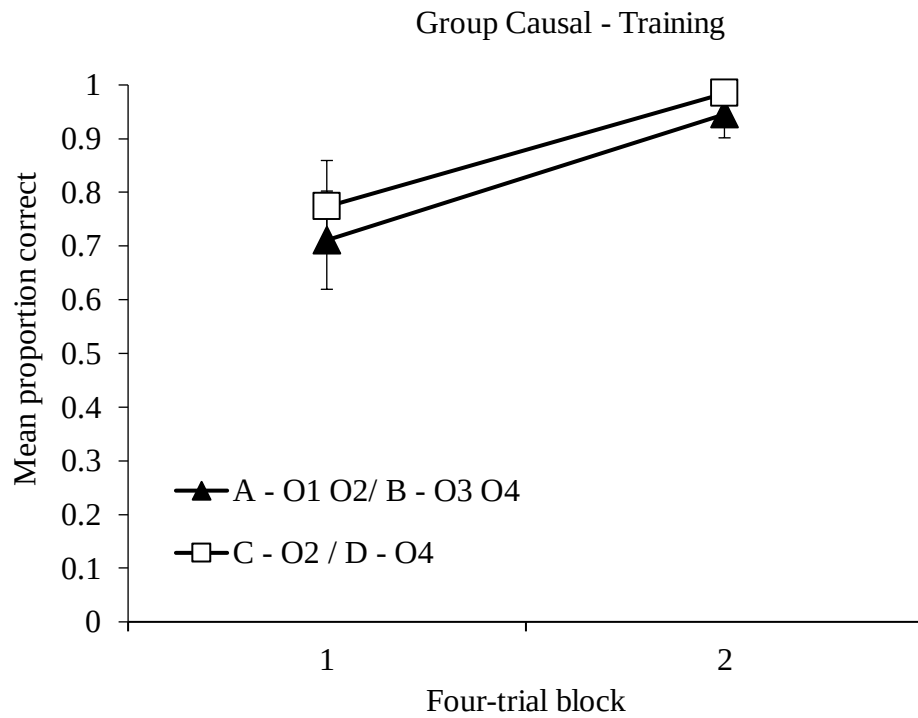


Figure 36. Mean proportion of correct responses for A-O1 O2/ B-O3 O4 and C-O2 /D-O4 for the causal version of the task. Error bars represent SEM.

Figure 37. Mean proportion of correct responses for A-O1 O3/B-O2 O3 and C – O2/D-O4 for the diagnostic version of the task. Error bars represent SEM.

4.1.2.1 Training Stage

Figures 36 and 37 show the mean proportion of correct responses for the cues (A-D) across the Training Stage. The data have been collapsed into 4 trial blocks in accordance with the training stage of the relative validity tasks. As can be seen participants quickly learnt the relationships between the cues and the outcome compounds (i.e., A – O1 O2 / B – O3 O4) and the cues and the outcome elements (i.e., C – O2 / D – O4), demonstrating a high proportion of correct responses after Block 1 and reaching near asymptote on the second block. Interestingly, participants appeared to learn the associations quicker in Group Diagnostic than Group Causal.

A three-way mixed model ANOVA was performed on these data with a between-subject factors of Task type (Causal vs Diagnostic), and within-subject factors of Stimuli (A/B vs C/D) and Trial-block (1 – 2). The ANOVA revealed a main effect of Trial-block, $F(1, 30) = 89.17$, $MSE = .01$, $p < .001$, $\eta_p^2 = .75$, a main effect of Task type, $F(1, 30) = 7.58$, $MSE = .02$, $p < .05$, $\eta_p^2 = .20$, and significant interactions between Trial-block and Group, $F(1, 30) = 4.55$, $MSE = .01$, $p < .05$, $\eta_p^2 = .13$, and Group and Stimuli, $F(1, 30) = 10.22$, $MSE = .00$, $p < .05$, $\eta_p^2 = .25$. All other interactions were non-significant, largest $F(1, 30) = 1.37$, $MSE = .00$, $p = .25$, $\eta_p^2 = .04$. To explore the source of the interactions simple main effects were conducted. These analyses revealed that participants in Group Diagnostic had a higher mean proportion of correct responses after one block of training than those in Group Causal, $F(1, 30) = 6.84$, $MSE = .01$, $p < .05$, $\eta_p^2 = .19$. Furthermore, participants in Group Causal learned more about the relationship between the

cues and the outcome elements, than the cues and the outcome compounds, $F(1, 30) = 9.64$, $MSE = .00$, $p = < .05$, $\eta_p^2 = .39$.

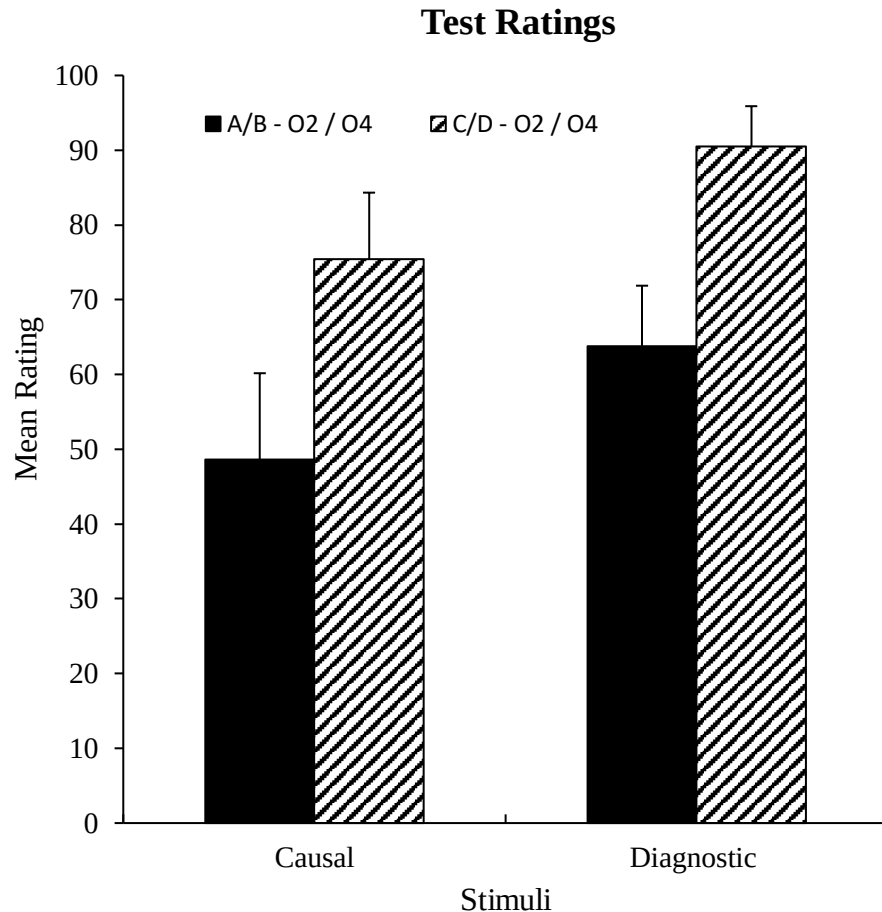


Figure 38. Mean of participants' difference ratings for A/B – O2 / O4 and C/D – O2 / O4 at test for both the causal and diagnostic version of the task. Difference ratings were calculated by subtracting ratings given to an outcome which was treated in an identical manner (e.g. O4) but was not paired with the cue (e.g. A), from the outcome (e.g. O2) which was paired with the cue (in accordance with Le Pelley and McLaren, 2003). Error bars represent SEM.

4.1.2.2 Test Data

Figure 38 shows participants' outcome specific ratings at test. These ratings assess whether participants' ratings were higher or lower for the specific outcomes of interest (i.e., O2 and O4). To provide an outcome specific measure of participants' ratings for cues A – D and O2 and O4 a difference

score was calculated. To calculate this measure, the rating for the outcome which a cue had never been paired with – but was otherwise treated in an identical manner – was subtracted from the rating for the outcome the cue had been paired with. For example, participants' ratings for A – O4 (which A had never been paired with, yet B had) were subtracted from their ratings for A – O2 (which A had been paired with). For the control stimulus C, participants' ratings for C – O4 were subtracted from their ratings for C – O2.⁷

As can be seen in Figure 38 participants' ratings for A/B – O2/O4 were lower than their ratings for C/D – O2/O4. Participants in Group Diagnostic also provided slightly higher scores for both sets of stimuli compared to Group Causal. A 2×2 mixed model ANOVA with a between-subjects factor of Task type (Causal vs Diagnostic) and a within-subjects factor of Stimuli (A/B – O2/O4 vs C/D – O2/O4) revealed a significant main effect of the factor Stimuli, $F(1, 30) = 18.54$, $MSE = 619.77$, $p < .001$, $\eta_p^2 = .38$. However, the main effect of Task type was non-significant, $F(1, 30) = 1.99$, $MSE = 1825.72$, $p = .17$, $\eta_p^2 = .06$. The interaction between these factors was also non-significant, $F(1, 30) = .00$, $MSE = 619.77$, $p = .99$, $\eta_p^2 = .00$. These results reveal that participants provided lower ratings to A/B – O2/O4 than C/D – O2/O4 regardless of task type, therefore demonstrating an outcome overshadowing effect.

⁷ For stimulus B, participants' ratings of B – O2 were subtracted from B – O4. For stimulus D, participants' ratings for D – O2 were subtracted from D – O4.

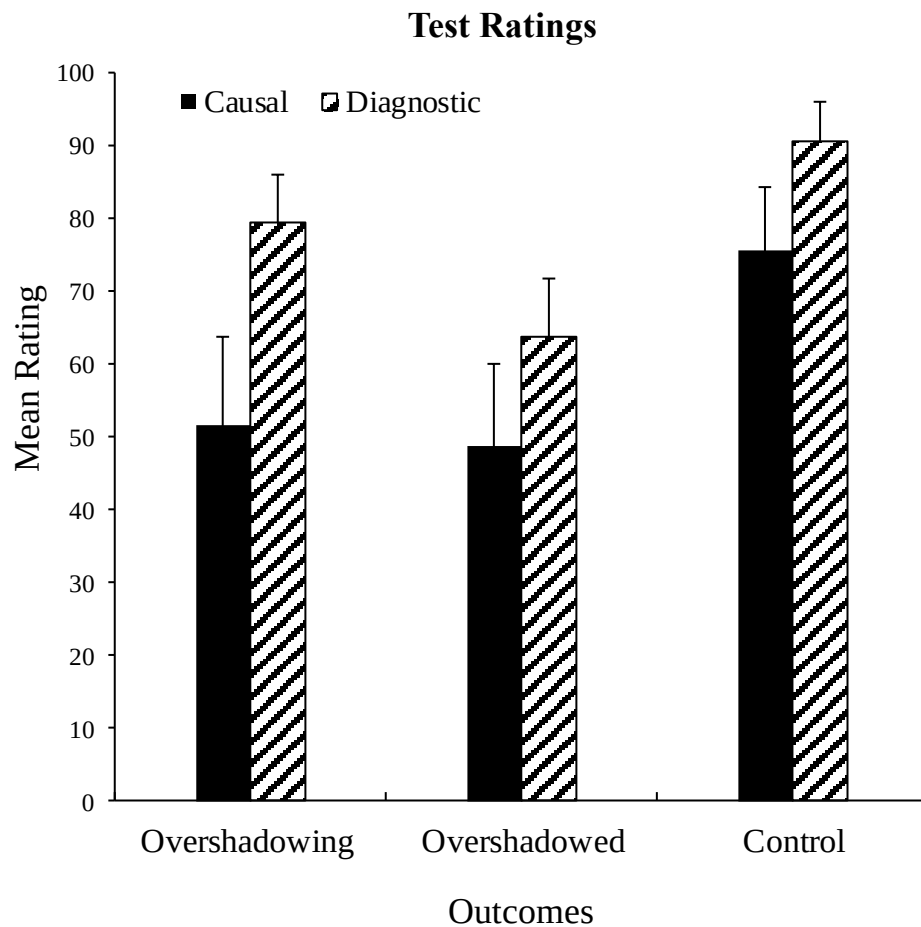


Figure 39. Mean of participants' difference ratings for the outcomes at test for both the causal and diagnostic version of the task. Difference ratings were calculated by subtracting ratings given to an outcome which was treated in an identical manner (e.g. O3) but was not paired with the cue (e.g. A), from the outcome (e.g. O1) which was paired with the cue (in accordance with Le Pelley and McLaren, 2003). Error bars represent SEM.

4.1.2.3 Type of Overshadowing

In previous studies which have reported outcome overshadowing (Arcediano et al., 2005; Burger, Mallemat & Miller, 2000) the nature of the outcome overshadowing effect was not discussed. For example, whether overshadowing occurred due to both elements of the compound being rated comparably low (i.e., reciprocal overshadowing) or due to one element of the outcome compound being rated lower than the outcome element (i.e., unilateral

overshadowing) was not discussed. Yet the nature of the overshadowing effect in cues has been a key test of predictions derived from theoretical models to account for this effect. For instance, according to the Rescorla and Wagner model (1972) overshadowing arises due to both outcome elements receiving a comparable share of the limited associability that is supported by the US, thus overshadowing should be reciprocal in nature. Yet, according to the Mackintosh model (1975), the overshadowing effect reported in cues arises due to the extent to which a component of a compound assumes the role of being the best predictor of the outcome, which is determined by the salience or associability of the relevant stimuli. According to Mackintosh, the stimulus with the lower salience will be learned about less than the stimulus with the higher stimulus. However, the stimulus with the higher salience will be learned about to the same extent as the cue presented on its own (1976). This effect is referred to as unilateral overshadowing. Given the absence of discussion concerning the nature of the overshadowing effect in outcomes, this will be explored in the following analysis.

To assess the nature of the outcome overshadowing effect further analysis was conducted on the ratings participants provided to cues A/B and the outcome elements, O1 and O3, of the outcome compounds O1 O2 and O3 O4. These outcome elements (i.e., A/B – O1/O3) will be referred to henceforth as the “Overshadowing” outcomes. This was done in order to assess whether the outcome overshadowing effect which was observed arose as a result of reciprocal overshadowing, where both elements of the outcome compound compete with each other when entering into an association with their

respective cue, or unilateral overshadowing, where only learning about only one element is impeded relative to the control.

To address this question participants' outcome specific measures were analysed. Outcome specific ratings were calculated in a manner consistent with the outcome specific measure of participants' ratings for O2 and O4 outlined above. For example, participants' ratings for the outcome which had never been paired with the cue of interest – but had served as an Overshadowing outcome – were subtracted from their ratings for the Overshadowing outcome the cue had been paired with. For instance, participants ratings for A – O3, with which A had never been paired (yet O3 had served as an Overshadowing outcome) were subtracted from their ratings for A – O1, which served as the Overshadowing outcome on the A trials. These data are presented alongside participants' outcome specific ratings for A/B – O2/O4, referred to as Overshadowed outcomes, and C/D – O2/O4 (referred to as Control outcomes), in Figure 39. As can be seen participants ratings for both the Overshadowing and Overshadowed outcomes were lower than the control outcomes regardless of task type.

A 2×3 mixed model ANOVA was performed on these data with a between-subjects factor of Task type (Causal vs Diagnostic) and a within-subjects factor of Stimuli (Overshadowing, Overshadowed, Control). This ANOVA revealed a significant main effect of Stimuli, $F(2, 60) = 9.78$, $MSE = 605.40$, $p < .001$, $\eta_p^2 = .25$, a non-significant main effect of Task Type, $F(1, 30) = 3.26$, $MSE = 2755.88$, $p = .08$, $\eta_p^2 = .10$, and a non-significant interaction between these factors, $F(2, 60) = .73$, $MSE = 605.40$, $p = .49$, $\eta_p^2 = .02$. To further explore the main effect of Stimuli, Bonferroni-corrected t-tests

were conducted (adjusted $p = .017$). These t-tests revealed significant differences between the Overshadowing outcome and the Control outcome, $t(31) = -3.09$, $p < .017$, and the Overshadowed outcome and the Control outcome, $t(31) = -4.38$, $p < .001$. However, there was no difference between the Overshadowing outcome and the Overshadowed outcome, $t(31) = 1.42$, $p = .16$. These results confirm that both outcomes of the outcome compound were rated lower than the control outcomes. That is, a reciprocal overshadowing effect was observed when analysing the outcome specific ratings.

Further analysis was also conducted on these data to assess if the reciprocal overshadowing effect was artificially produced as a result of determining the overshadowed (i.e., O2/O4) and overshadowing outcomes (i.e., O1/O3) a priori. It could be, for instance, that there were individual differences in which outcomes served as the overshadowed and overshadowing outcomes. To further elaborate, it might be that for one participant O1 is the overshadowing outcome and O2 is the overshadowed outcome, for another participant, however, O2 might be the overshadowing outcome and O1 the overshadowed outcome. If this was indeed the case then in some instances O1 will receive a higher rating and in others O2 will receive a higher rating, therefore when averaging these ratings for both outcomes across participants, both will receive a lower rating than a control outcome because of an averaging artefact. As such, the preceding analysis could possibly be erroneously detecting reciprocal overshadowing. To mitigate this, an analysis was also performed where the overshadowing outcome was defined as the outcome element of the compound which received the highest rating,

regardless of whether the outcome was also presented on its own on a control trial. Bonferroni corrected t-tests (adjusted $p = .017$) revealed that whilst the overshadowing outcome received a higher rating than the overshadowed outcome (which would be expected as the highest rated outcome was always selected as the overshadowing outcome), $t(31) = -4.14$, $p < .001$, both elements of the outcome compound were rated lower than the control outcome, smallest $t(31) = -2.58$, $p < .017$. These results therefore demonstrate a reciprocal overshadowing effect regardless of how the overshadowing stimulus is defined.

To summarise, Experiment 9 demonstrated that when employing an outcome overshadowing procedure, participants rated the outcomes which were presented on the overshadowing trials lower than the outcome elements which appeared on the control trials. These results provide evidence of an overshadowing effect which operates in a similar manner in both cues and outcomes. Further analysis concerning the nature of the outcome overshadowing effect also revealed that ratings for both elements of the outcome compound were lower than the outcome element presented on its own, thus demonstrating a reciprocal overshadowing effect. The observation of a reciprocal outcome overshadowing effect would appear to be best accommodated by Wagner's SOP model. According to this account, overshadowing would arise due to both elements of the outcome compound competing to enter into their A1 state, thus reciprocally preventing one another from being learned about to the same as extent as an outcome element presented on its own.

Additionally, the type of task participants completed appeared to influence their learning about these tasks in the training stage but not at test. For instance, in the diagnostic task participants displayed better learning than in the causal task. Moreover, in the causal judgement task participants learned about the outcome elements better across the course of training than the outcome compounds. One account of this effect is to simply suggest that the processing demands of learning about the additional outcome present in the outcome compound results in poorer learning. However, this would not account for why this effect was not present in the diagnostic task. Instead, the effect, in conjunction with the fact that participants displayed better overall learning in the training stage of the diagnostic task, would appear to support the idea that participants are able to better learn about multiple causes when the effects are presented first (in a diagnostic task), as is suggested by causal model theory (Waldmann & Holyoak, 1992). According to causal model theory when participants are first presented with the effect of a cause – as is the case in a diagnostic task – participants are more likely to consider that there are multiple possible causes as opposed to when completing a causal judgment task.

4.2 Experiment 10

Experiment 9 demonstrated an outcome overshadowing effect. These results are consistent with the findings of Arcediano et al. (2005). These results also provide further evidence of an interaction effect amongst the elements of an outcome. Given the presence of outcome interaction effects in both

blocking and overshadowing tasks, these results could be taken as evidence to suggest that comparable effects may be observed in an outcome relative validity task. Few previous studies, however, have explored whether a relative validity effect can be observed amongst outcomes and those that have been conducted, have provided heterogenous results. For instance, Matute, Arcediano and Miller (1996) demonstrated relative validity amongst the elements of an outcome but did so using a task where the cues and outcomes were presented simultaneously in a list like manner, thus it is possible that participants treated the task as diagnostic and therefore treated the outcomes as cues. In light of this, Experiment 10 sought to assess whether an outcome interaction effects can be observed when employing an outcome relative validity procedure with the same food allergist task as reported previously. The direction of causality was also tested with both causal judgement and diagnostic judgement tasks being employed. The design of Experiment 10 can be seen in Table 15. As both causal and diagnostic versions of the task were used the stimuli which served as cues (i.e., A-D), were either represented by foods in the causal version of the task, or reactions in the diagnostic version of the task. The stimuli which served as outcomes (i.e., O1-O6) were reactions in the causal task and foods in the diagnostic task.

Table 15. Design of Experiment 10.

	<i>Training</i>	<i>No. of trials</i>	<i>Test</i>
True	A – O1 O3	24	A: O1 - O6
Discrimination	B – O2 O3	24	B: O1 - O6
Pseudo	C – O4 O6	12	-
Discrimination	C – O5 O6	12	C: O1 - O6
	D – O4 O6	12	D: O1 - O6
	D – O5 O6	12	-

Note. A – D refers to cues, whilst O1 – O6 refer to outcomes. A and B represent the true-discrimination trials, whilst C and D trials represent the pseudo-discrimination trials.

In Stage 1, four cues (A – D) were presented which were paired with four different outcome compounds (O1 O3, O2 O3, O4 O6 and O5 O6 respectively). Two of these cues, A and B, reliably predicted outcome compounds, O1 O3 and O2 O3 respectively. These stimuli represented the true-discrimination trials as one element of each outcome compound was only ever predicted by A (i.e., O1) or B (i.e., O2), whilst the other element of the compound (i.e., O3) was predicted by both. The remaining two cues, C and D, were presented an equal number of times as A and B. However, cues C and D were each paired with outcome compounds O4 O6 and O5 O6, half the time they were presented. This therefore resulted in one element of each outcome compound (O6) always being present for both C and D, whilst the other element (i.e., O4/O5) would be present on half of all C and D trials. As such, these trials served as pseudo-discrimination trials, whilst one element of the outcome compound could be predicted (O6), it was impossible to predict the correct outcome compound on all occasions.

Participants were tasked with learning these associations between the cues and outcomes across the course of the experiment. Prior to a Test stage,

participants were informed that they would now be asked to make a rating concerning how predictive each cue (A – D) was of each of the respective outcomes (O1 – O6). They were then presented with each of the cues (individually) and asked to make their ratings for each of the outcomes (O1 – O6). If a comparable interaction effect to that observed in cues is demonstrated, it would be anticipated that the common outcome which occurs on the pseudo-discrimination trials (O6) would receive a higher rating than the common outcome on the true-discrimination trials (O3). If, however, a facilitation effect is observed comparable to that observed when an outcome blocking procedure was employed, it would be anticipated that better learning would be displayed toward the common outcome on the true-discrimination trials.

4.2.1. Method

4.2.1.1 Participants

Forty-eight participants (37 females; 11 males) were recruited from the University of Nottingham's School of Psychology. Participants ranged from 18 – 36 years of age ($M = 21.94$; $SEM = .60$). Participants received course credit for their participation or a cash payment of £3. Participants were randomly allocated (on the condition that there were an equal number of participants in each group) to either complete either the causal judgment task ($n = 24$) or the diagnostic judgement task ($n = 24$).

4.2.1.2 Apparatus and Stimuli

Each of the foods and reactions were assigned to the letters and outcomes in Table 15 using a Latin-Squared counterbalancing technique. The six food stimuli used were: broccoli, cabbage, mushroom, pepper, potato, tomato. The six reactions included were: diarrhoea, fever, headache, nausea, skin rash and vomiting. Both the causal and diagnostic versions of the tasks were identical to that employed in Experiment 9, with the only change being that the outcomes were always presented in compound form. At test each of the cues (either foods or reactions dependent on the version of the task) were presented individually on the left of the screen, on the right each of the outcomes were presented (see Figure 40). Next to each outcome a Likert scale would be displayed, given that there were six outcomes presented, six scales were presented on the right of the screen. The order of these screens was counterbalanced.

4.2.1.3 Procedure

As in Experiment 5-7 of Chapter 3 both groups experienced 6 blocks of 4 trials in a training stage before proceeding to a test stage. The procedure for the experiment was the same as Experiment 9, except that in this experiment there were two additional outcomes which helped form four outcome compounds. The Test stage was identical to Experiment 9, except that on each rating screen participants were presented with six outcomes to rate as opposed to four.



Figure 40. An example of a training trial (left), feedback screen (middle) and a test screen that participants would be presented with in the causal version of the task.

4.2.2 Results and Discussion

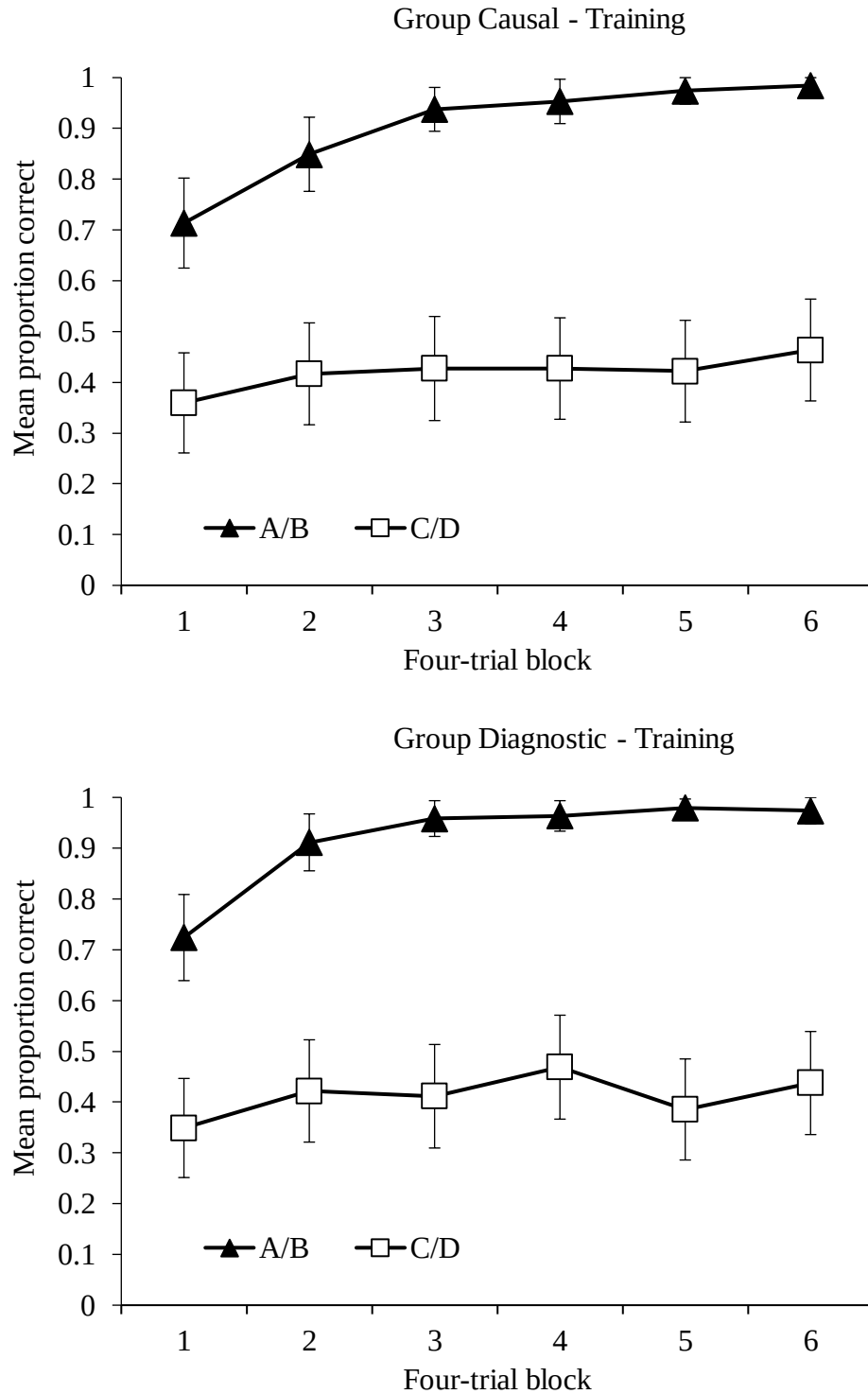


Figure 41. Mean proportion of correct responses for A-O1 O3/ B-O2 O3 and C-O4 O6/D-O5 O6 for the causal version of the task. Error bars represent SEM.

Figure 42. Mean proportion of correct responses for A-O1 O3/B-O2 O3 and C – O4 O6/D-O5 O6 for the diagnostic version of the task. Error bars represent SEM.

4.2.2.1 Training Stage

Figures 41 and 42 show the mean proportion of correct responses for the cue-outcome pairings (A-O1 O3/ B - O2 O3 and C - O4 O6/O5 O6 /D - O5 O6/O5 O6) across Stage 1 training. The data in Figures 41 and 42 have been collapsed into 6 blocks, each comprising 4 trials. As can be seen, regardless of the task completed (e.g., causal/diagnostic) participants displayed the same pattern of results, with participants rapidly learning the relationships between the cue-outcome associations for the true-discrimination trials (i.e., stimuli A/B) – appearing to reach asymptote by the end of training – and failing to achieve asymptote for the pseudo-discrimination trials (i.e., stimuli C/D) across Stage 1 training.

A three-way mixed model ANOVA was performed with a between-subject factors of task type (Causal vs Diagnostic), and within-subject factors of cue-outcome associations (A/B vs C/D) and trial-block (1 – 6). The ANOVA revealed a main effect of trial-block, $F(5, 194.58) = 18.01$, $MSE = .03$, $p < .001$, $\eta_p^2 = .28$, a main effect of cue-outcome associations, $F(1, 46) = 1245.39$, $MSE = .03$, $p < .001$, $\eta_p^2 = .96$, and a significant interaction between these two factors, $F(4.29, 197.20) = 6.75$, $p < .001$, $\eta_p^2 = .13$. There was no effect of task type, $F(1, 46) = .05$, $MSE = .07$, $p = .83$, $\eta_p^2 = .00$, and all other interactions were non-significant, largest $F(1, 46) = .70$, $MSE = .028$, $p = .41$, $\eta_p^2 = .02$. To further explore the source of the interaction between trial-block (1-6) and cues (A/B vs C/D) simple main effects were performed. There was an effect of stimuli on all blocks, smallest $F(1, 46) = 131.75$, $MSE = .02$, $p < .001$, $\eta_p^2 = .73$.

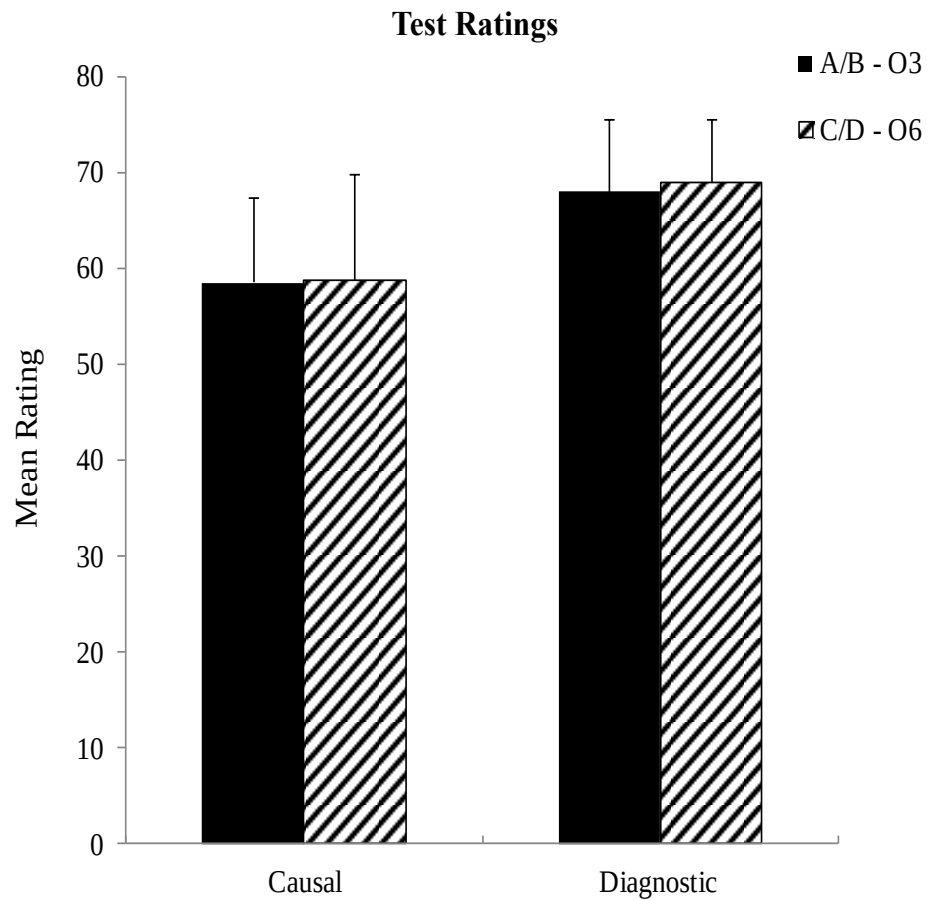


Figure 43. Mean of participant's difference ratings for A/B – O3 and C/D – O6 at test for both the causal and diagnostic version of the task. Difference ratings were calculated by subtracting ratings given to the common outcome which was not paired with the cue, from the common outcome which was paired with the cue (in accordance with Le Pelley and McLaren, 2003). Error bars represent SEM.

4.2.2.2 Test Data

Figure 43 illustrates participants' outcome specific ratings at test. One participant in Group Diagnostic failed to provide ratings for all the required outcomes. Outcome specific ratings were calculated in order to assess whether participants' ratings were higher or lower for the specific outcomes of interest (i.e., O3 and O6). To provide an outcome specific measure of participants' ratings for cues A – D and O3 and O6 a difference score was calculated. For example, for stimuli A/B, participants' ratings for these cues and the outcome

element that had always been predicted on the pseudo-discrimination trials (i.e., C/D - O6) – but never on the true discrimination trials – was subtracted from their ratings for these cues and the outcome element of interest which had always been presented on A/B trials. For stimuli C/D participants' ratings for these cues and the outcome which always occurred on the true discrimination trials (i.e., A/B – O3) – but never on the pseudo-discrimination trials – was subtracted from their ratings of the relationship between these cues and outcomes which did always appear on the pseudo-discrimination trials.

As can be seen in Figure 43, participants' ratings in Group Diagnostic are slightly higher for both sets of stimuli (i.e., A/B and C/D) than Group Causal. However, their ratings concerning the relationship between these stimuli and the outcomes they were most frequently paired with were similar, with participants rating both outcomes' at a comparable level regardless of task type. A 2×2 mixed model ANOVA with a between-subjects factor of task type (Causal vs Diagnostic) and a within-subjects factor of outcome (O3 vs O6), revealed no main effect of the factor outcome, with O3 and O6 receiving comparable ratings to each other, $F(1, 45) = .01$, $MSE = 885.36$, $p = .92$, $\eta_p^2 = .00$, and no main effect of task type, $F(1, 45) = .86$, $MSE = 2612.04$, $p = .36$, $\eta_p^2 = .02$. The interaction between these two variables was also non-significant, $F(1, 45) = .00$, $MSE = 885.36$, $p = .95$, $\eta_p^2 = .00$. As such, these results confirm participants' outcome specific ratings for the O3 and O6 were again comparable with task-type exerting little effect.

Experiment 10 demonstrated that when employing an outcome relative validity procedure, participants rated the common outcome which occurred on the true discrimination trials (i.e., A/B – O3) as equally predictive as the

common outcome which occurred on the pseudo-discrimination trials (i.e., C/D – O6). These results demonstrate the absence of an outcome interaction effect when employing an outcome relative validity paradigm. As such, they are inconsistent with Experiment 3 of Matute, Arcediano and Miller (1996) which demonstrated an outcome relative validity effect comparable to that observed in cues (they are, however, consistent with Experiment 2 reported by these authors). These results are also inconsistent with the findings from Chapter 3 and Experiment 9 which demonstrated interactions amongst the elements of an outcome. It is also worth noting the absence of an effect of task type. That is, participants provided comparable ratings regardless of whether they completed the causal or diagnostic version of the task.

4.3. Experiment 11

The results of Experiment 10 are somewhat surprising considering the findings reported in the previous experiments, all of which have observed outcome interaction effects of one kind (facilitation) or another (competition). Furthermore, given that cue interaction effects such as blocking, overshadowing and relative validity are assumed to be different manifestations of a common underlying mechanism in cues (e.g. Rescorla & Wagner, 1972), it would appear sensible to posit that the presence of outcome interaction effects may also be a manifestation of an overarching effect in outcomes. As such, the observation of outcome blocking and outcome overshadowing in the previous experiments, might also suggest that an interaction effect will be observed when an outcome relative validity task is employed. However, this was not the case in Experiment 10. It is possible, however, that an additional

factor may have influenced the results of Experiment 10. It could be, for instance, that the duration of the training which participants completed prior to the Test Stage influenced the results. Given the complexity of the task (relative to Experiment 9), the duration of training in Experiment 10 was extended to six blocks of four trials, which is three times the length of the equivalent stages where the manipulation of interest occurs in the experiments reported in Chapter 3 and Experiment 9 of this chapter. Importantly, the equivalent stages of training in the previous experiments were shorter to avoid any interaction effects being obscured by participants reaching asymptote. In light of the extended amount of training in Experiment 10, it is possible that participants' ratings for the outcomes are equally high as a result of them receiving more trials and potentially reaching asymptote.

As such, Experiment 11 sought to assess whether the extended amount of training in Experiment 10 impacted upon participants' learning by reducing the amount of training participants received. To ensure there was consistency with the experiments reported in Chapter 3, and Experiment 9, the duration of training provided was the same as the Stage 2 training of those experiments (excluding Chapter 3, Experiment 8). The experimental design was identical to that of Experiment 10 (see Table 16) with only the amount of training changing. If the duration of training in Experiment 10 was responsible for the comparable ratings provided to the outcomes at test, it might be suggested that the reduction in training will reveal any potential differences in learning about the outcomes. If the duration of training was not responsible for the

comparable ratings to the outcomes at test, the effect from Experiment 10 should be replicated.

Table 16. Design of Experiment 11

	<i>Training</i>	<i>No. of trials</i>	<i>Test</i>
True	A – O1 O3	8	A: O1 - O6
Discrimination	B – O2 O3	8	B: O1 - O6
Pseudo	C – O4 O6	4	-
Discrimination	C – O5 O6	4	C: O1 - O6
	D – O4 O6	4	D: O1 - O6
	D – O5 O6	4	-

Note. A – D refers to cues, whilst O1 – O6 refer to outcomes. A and B represent the true-discrimination trials, whilst C and D trials represent the pseudo-discrimination trials.

4.3.1 Method

4.3.1.1 Participants

Forty-eight participants (40 females; 8 males) were recruited from the University of Nottingham’s School of Psychology. Participants ranged from 18 – 22 years of age ($M = 19.00$; $SEM = 0.11$). All other details were the same as Experiment 9, with an equal number of participants in both Group Causal ($n = 24$) and Group Diagnostic ($n = 24$).

4.3.1.2 Apparatus and Stimuli

All stimuli used were the same as Experiment 10.

4.3.1.3 Procedure

The procedure was also the same as Experiment 10, with the only difference being that participants completed 2 blocks of training (with each block comprising 4 trials), as opposed to 6 blocks of training.

4.3.2 Results and Discussion

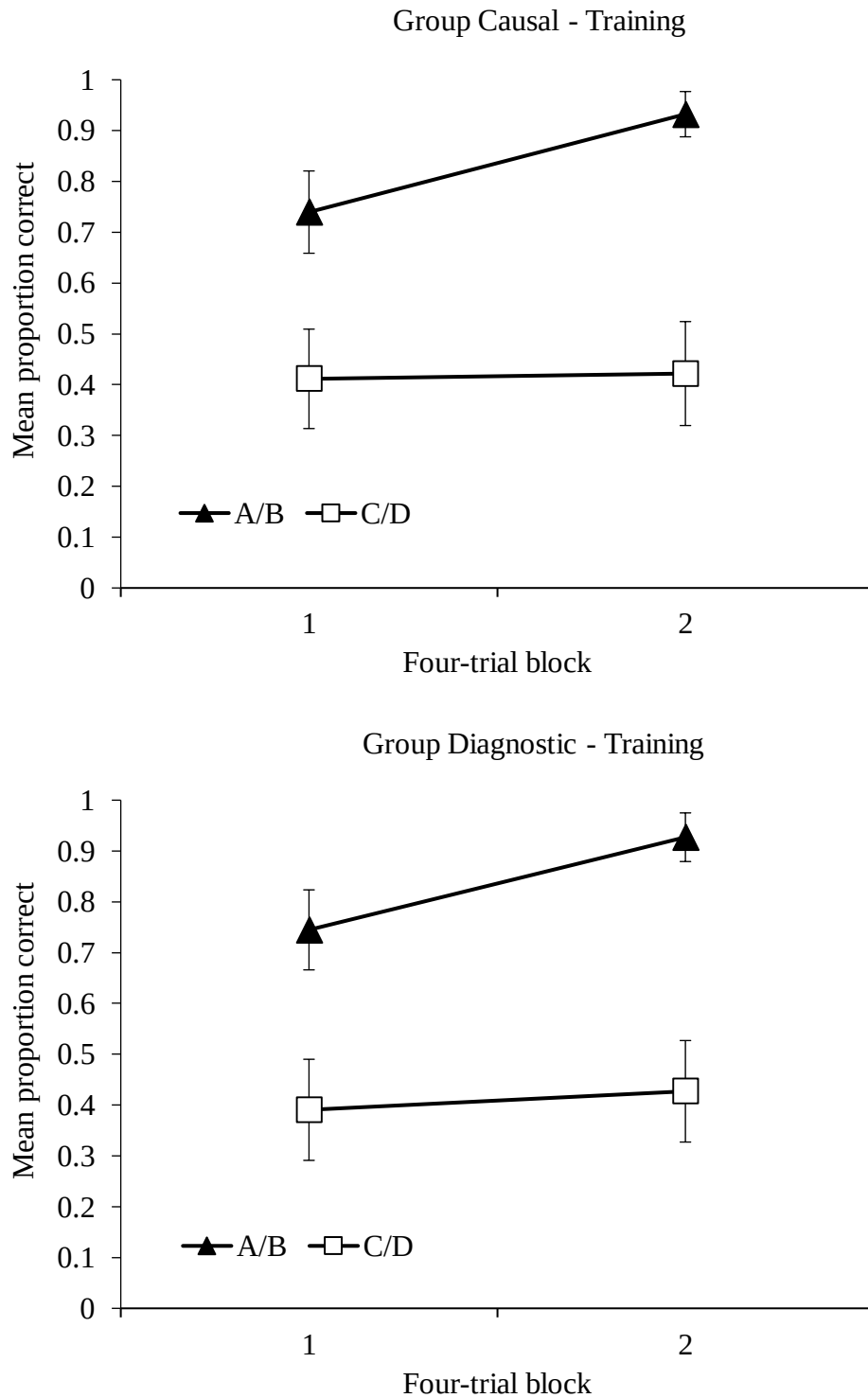


Figure 44. Mean proportion of correct responses for A-O1 O3/ B-O2 O3 and C-O4 O6/D-O5 O6 for the causal version of the task. Error bars represent SEM.

Figure 45. Mean proportion of correct responses for A-O1 O3/B-O2 O3 and C-O4 O6/D-O5 O6 for the diagnostic version of the task. Error bars represent SEM.

4.3.2.1 Training Stage

Figures 44 and 45 demonstrate the mean proportion of correct responses for the cues (A-D) across Stage 1 training. In keeping with Experiment 10 the data have been collapsed into 4 trial blocks. As in Experiment 10, participants quickly learnt the relationships between the cues and outcomes on the true-discrimination trials (i.e., stimuli A/B) but not the pseudo-discrimination trials (i.e., stimuli C/D), with participants mean proportion of correct responses staying consistently lower for the pseudo-discrimination trials. The task type participants completed also seemed to have little impact on their performance, also being consistent with Experiment 10.

A three-way mixed model ANOVA was conducted on these data with a between-subject factors of task type (Causal vs Diagnostic), and within-subject factors of stimuli (A/B vs C/D) and trial-block (1 – 2). The ANOVA revealed a main effect of trial-block, $F(1, 46) = 282.52$, $MSE = .03$, $p < .001$, $\eta_p^2 = .86$, a main effect of stimuli, $F(1, 46) = 24.34$, $MSE = .02$, $p < .001$, $\eta_p^2 = .35$, and a significant interaction between these two factors, $F(1, 46) = 20.31$, $p < .001$, $\eta_p^2 = .31$. There was no effect of task type, $F(1, 46) = .02$, $MSE = .04$, $p = .89$, $\eta_p^2 = .00$, and all other interactions were non-significant, largest $F(1, 46) = .25$, $MSE = .02$, $p = .62$, $\eta_p^2 = .01$. Simple main effects revealed there was an effect of stimuli on both blocks with participants producing a higher mean proportion of correct responses on the A/B trials, smallest $F(1, 46) = 125.58$, $MSE = .02$, $p < .001$, $\eta_p^2 = .73$.

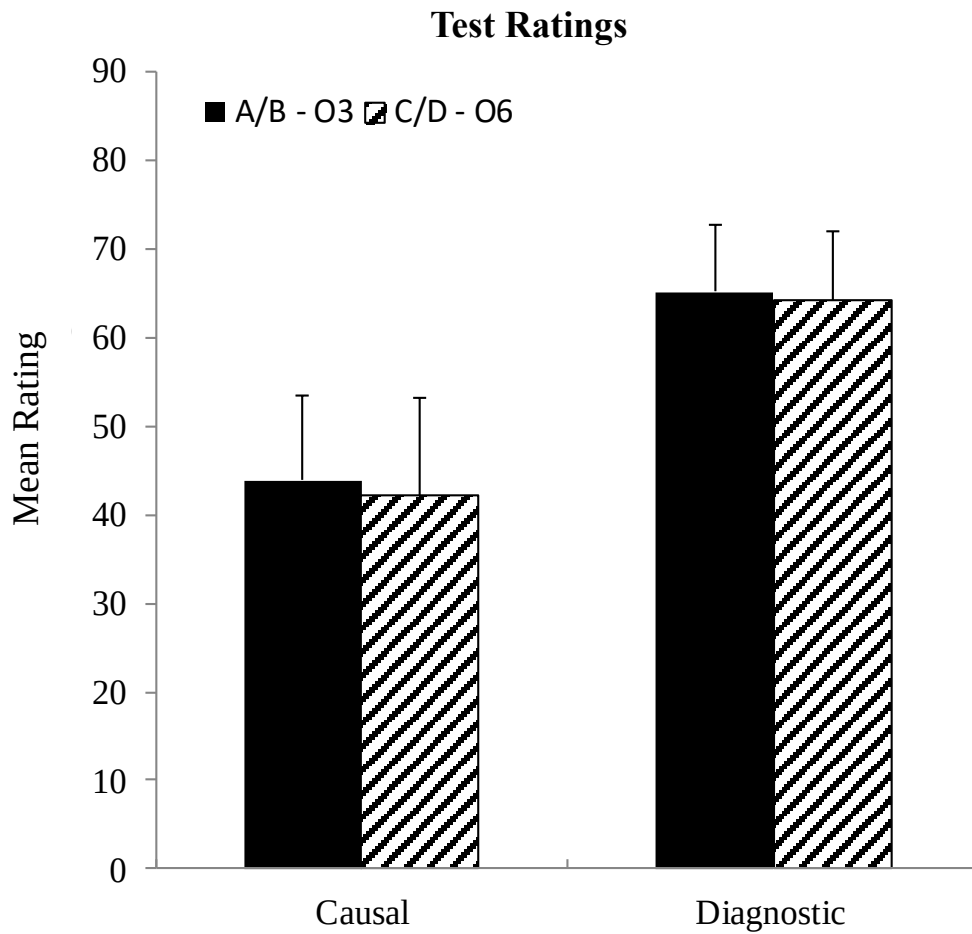


Figure 46. Mean of participants' difference ratings for A/B – O3 and C/D – O6 at test for both the causal and diagnostic version of the task. Difference ratings were calculated by subtracting ratings given to the common outcome which was not paired with the cue, from the common outcome which was paired with the cue (in accordance with Le Pelley and McLaren, 2003). Error bars represent SEM.

4.3.2.2 Test Data

Figure 46 demonstrates participants' outcome specific ratings at test. These ratings assess whether participants' ratings were higher or lower for the specific outcomes of interest (i.e., O3 and O6). To provide an outcome specific measure the difference score outlined previously in Experiment 10 was

calculated.⁸ As can be seen in Figure 46 participants ratings for the stimuli (i.e., A/B – O3 and C/D – O6) were again comparable within groups but appeared to differ between groups, with participants providing higher ratings for both sets of stimuli within Group Diagnostic, compared to Group Causal. A 2×2 mixed model ANOVA with a between-subjects factor of task type (Causal vs Diagnostic) and a within-subjects factor of outcome (O3 vs O6) revealed no main effect of the factor outcome $F(1, 46) = .04$, $MSE = 1013.04$, $p = .84$, $\eta_p^2 = .00$. The main effect of task type approached significance, $F(1, 46) = 3.81$, $MSE = 2933.67$, $p = .06$, $\eta_p^2 = .08$, however, the interaction between these two variables was also non-significant, $F(1, 46) = .00$, $MSE = 1013.04$, $p = .96$, $\eta_p^2 = .00$. These results confirm that whilst participants' ratings differed slightly dependent on task type, the outcome specific ratings for O3 and O6 were again comparable within groups.

In summary, Experiment 11 reproduced the effect from Experiment 10, with participants again providing comparable ratings to the outcomes regardless of whether they appeared on the true or pseudo-discrimination trials. This effect was observed even though the duration of training was reduced, thus demonstrating that participants extended training in Experiment 10 had not influenced the results. Given the null effect obtained in both experiments an additional Bayesian analysis was conducted on the combined test data of the two experiments. This was done to determine if the similarity in ratings between the outcomes supports the null hypothesis (Dienes, 2011).

⁸ For the A/B stimuli participants ratings for A/B and O6 would be subtracted from their ratings for A/B and O3. For the C/D stimuli participants ratings for the C/D and O3 would be subtracted from their ratings for the C/D and O6.

A Bayesian analysis was performed as outlined by Dienes (2014). This analysis produces a Bayes Factor, where values of less than .33 provide support for the null hypothesis; values > 3 provide evidence for the alternative hypothesis, whilst values ranging from .33 – 3 are inconclusive, providing support to neither hypothesis. In order to conduct this analysis an estimated effect size is required. To provide this value the mean difference ($M = 26.17$) and Standard Deviation (47.27) between conditions for Experiments 5 and 6 from Chapter 2 were calculated.⁹ The differences obtained in these experiments were used as an estimated effect size given the conceptual similarity between the types of interaction effect being explored (i.e., blocking and relative validity) and the procedural similarities in the task participants completed (e.g., both tasks feature outcome compounds in the critical stage of training). By calculating the mean difference ($M = .03$) and standard error of the mean difference ($SEM = 2.93$) for the collapsed Experiment 10 and 11 data, along with the estimated effect size of Experiments 5 and 6, this provided a Bayes factor of .05, providing support for the null hypothesis. Bayesian analysis further revealed that the likelihood of the null being true is .13 whilst the likelihood the alternative hypothesis being true is .00. As such, these results provide evidence of an absence of interaction effects amongst the elements of an outcome when employing an outcome relative validity task.

⁹ Experiment 7 and 8 were omitted as they employed additional trials and were procedurally different to the current experiments (i.e., they employed varying amounts of Stage 2 training).

4.3 General Discussion

Three experiments investigated the presence of interaction effects amongst the elements of an outcome compound when employing outcome overshadowing and relative validity procedures. In Experiment 9 an outcome overshadowing design was employed. In this experiment outcome compounds were reliably predicted in addition to outcome elements. The outcome elements received higher ratings than the elements of the outcome compound, thus demonstrating an outcome overshadowing effect. In Experiments 10 and 11 an outcome relative validity task was employed. In these tasks a control outcome was presented alongside an outcome element which was inconsistently predicted, whilst an experimental outcome was presented alongside an outcome element which was consistently predicted. In these experiments no difference was observed in participants' ratings to the outcomes, thus demonstrating the absence of an outcome interaction effect.

These results further highlight the unique nature of interaction effects in cues and outcomes. Typically, studies employing relative validity procedures with multiple cues (e.g., Wagner, Logan, Haberlandt, & Price, 1968) demonstrate that learning about the cue (i.e., X) present in both cue-compounds (i.e., AX+, BX-) on the true discrimination trials is learned about less than the cue present (i.e., Y) in both cue-compounds (i.e., CY+/-, DY+/-) on the pseudo-discrimination trials.¹⁰ Yet, no such interaction effect was observed in these studies. Conversely, however, overshadowing was observed in the outcome overshadowing procedure employed in Experiment 9, thus

¹⁰ Wagner et al. (1968) employed a between-subject design. The example provided here refers to a within-subjects design to ensure consistency with the Experiments 10 and 11 reported in this Chapter.

providing evidence of an interaction effect which operates in a similar manner in both cues and outcomes.

The results of Experiment 9 would appear to be best accommodated by the SOP model provided by Wagner and the account derived from this model for the effects reported in Chapter 3. According to Wagner overshadowing will be observed amongst the elements of an outcome compound as both elements compete to enter into their A1 state and reciprocally prevent one another from fully occupying this state. Notably, further analysis concerning the nature of the overshadowing effect revealed the presence of reciprocal overshadowing. That is, both elements of the outcome compound received comparably low ratings relative to the outcome element, as opposed to a unilateral overshadowing effect, where one outcome of the compound received a lower rating than the outcome element (i.e., the overshadowed outcome), yet the accompanying outcome receives comparable ratings to the outcome element (i.e., the overshadowing outcome).

The results of this experiment are less easily accounted for by the Mackintosh model though, which an explanation was also derived from in order to account for the findings of Chapter 3. It is possible to suggest that stimuli C/D become better predictors of the outcome element as these stimuli exclusively predict this outcome, as opposed to stimuli A/B which predict two outcomes, therefore permitting the model to account for the effect. However, stimuli A/B are equally reliable predictors of an outcome as stimuli C/D, as they consistently predict the same set of outcomes throughout the course of training, thus it is arguable that according to the model all stimuli should be learned about comparably and as such outcome overshadowing should be

absent. Furthermore, the nature of the overshadowing effect would also appear to be inconsistent with the Mackintosh model. For instance, assume that a cue (e.g., A) that predicted an outcome compound (e.g., O1 O2), could indeed be rendered an inferior predictor of the outcome element (e.g., O2) that is exclusively predicted by a control stimulus, and this therefore results in a weaker association being formed (between A – O2). According to the model the reduced associative strength between the cue (e.g., A) and the outcome element (e.g., O2), would result in the cue receiving an enhancement in its association with the other outcome present in the compound (e.g., O1). As such, A will become a better predictor of O1 than O2, and the rating for A – O1 will be higher than for A – O2, resulting in a unilateral overshadowing effect. However, analysis of the ratings for each element of the outcome compound revealed that both elements received ratings lower than a control outcome, thus being inconsistent with this account of the type of overshadowing effect which might be derived from the model.

The results of Experiment 9 also demonstrated that the type of task participants completed influenced the extent to which participants learned about the outcomes during the training stage. In the diagnostic task participants displayed better learning than in the causal task, and in the causal judgement task participants learned about the outcome elements better across the course of training than the outcome compounds. One way of accommodating this effect is to suggest that the processing demands of learning about the additional outcome present in the outcome compound results in poorer learning. This would not account for why this effect was not present in the diagnostic task though. Instead, the effect, in conjunction with the fact that

participants displayed better overall learning in the training stage of the diagnostic task, would appear to support the idea that participants are able to better learn about multiple causes when the causes are presented following an effect, as is suggested by causal model theory (Waldmann & Holyoak, 1992). According to this theory when participants are first presented with the effect of a cause – as is the case in a diagnostic task – participants are more likely to consider that there are multiple possible causes as opposed to when completing a causal judgment task.

Nevertheless, this difference in task type during the learning stage had no impact on the test ratings (which is where any effect would be likely to be present). Additionally, an alternative explanation could also be provided as to why both the outcome elements and compounds were better learned about in the diagnostic task. For example, in the diagnostic task the reactions served as the cues and the food images served as the outcomes, yet in the causal task the reactions (presented in text form) served as the outcomes. Given that the food images are likely to have a higher salience as outcomes than the reactions, it is possible that participants were better able to learn about the outcomes when they were images as opposed to when they were text.

Given the presence of overshadowing in Experiment 9 it is important to consider why an outcome interaction effect was not present in the relative validity experiments. As noted previously, Matute, Arcediano and Miller (1996) also conducted a set of experiments to explore whether interaction effects could be observed amongst outcomes when an outcome relative validity procedure was employed. In their study participants completed a medical decision-making task where a medicine was correlated with the presence of

three possible syndromes which formed outcome compounds. In Experiment 2 of their study the authors did not observe any interaction between the elements of the outcome compound. However, in Experiment 3 of their study these authors observed a relative validity effect in outcomes.

The results of Experiments 10 and 11 are also inconsistent with the experiments reported in Chapter 3 when an outcome blocking procedure was employed. In these studies, both competition and facilitation were observed. Given that interaction effects (e.g. blocking, overshadowing and relative validity) are assumed to have similar properties, it would appear sensible to suggest that the results of Chapter 3, and Experiment 9, would indicate that comparable effects will be observed when other similar procedures are employed. The Mackintosh account of the findings in Chapter 3 would also suggest an interaction effect might be observed in Experiments 10 and 11. According to the Mackintosh model (1975) better learning would be expected about the outcomes presented on the true-discrimination trials, as the cues paired with these outcomes served as consistent predictors. As such, the associability of these cues would be expected to increase, relative to cues on the pseudo-discrimination trials which were paired with two outcome compounds 50% of the time. That is, a facilitation effect would be expected. The SOP account of the outcome relative validity experiments reported in this chapter would also appear to anticipate a facilitation effect, although it is more ambiguous. For instance, according to the model, it is possible that on true-discrimination trials learning about the common outcome (e.g., A-O3) will be overshadowed to some extent by the presence of the distinctive outcome (e.g., A-O1), as they compete to enter into their A1 state. On the pseudo-

discrimination trials (C/D) the presence of the two alternating outcomes (O4/O5) will also result in a form of competition with the common outcome (O6), however, the alternating O4O6 and O5O6 outcomes will result in extinction on the trials where O4 or O5 are expected, but not presented. This will result in these outcome elements (O4/O5) being promoted into A1 on each trial in which they are presented and thus interfering with learning about O6 more than the accompanying outcomes (O1/O2) on the A/B trials.

Given the absence of any difference in ratings between the outcomes on the true and pseudo-discrimination trials it is possible that both competition and facilitation effects were present simultaneously between participants in Experiments 10 and 11, thus cancelling out any differences between the ratings provided to the outcomes. It is possible to test this suggestion by assessing the differences between each participant's ratings for the outcome which appeared on the true discrimination trials and the outcome which appeared on the pseudo-discrimination trials. If a participant displayed a facilitation effect a higher rating for the outcome on the true-discrimination trials would be displayed. If a participant displayed a competition effect a higher rating would be displayed for the outcome which occurred on pseudo-discrimination trials. If both these effects were operating simultaneously between participants a comparable number of participants would have provided ratings in line with both effects. This analysis of individuals' difference ratings at test across both experiments provides some support to the notion that both effects are operating simultaneously. Of the 95 participants who provided ratings across these experiments, 42 participants provided higher ratings to outcome presented on the true-discrimination trials (i.e., a facilitation effect); 37 provided higher

ratings to the outcomes provided on pseudo-discrimination trials (i.e., a competition effect), whilst 16 participants provided tied ratings (i.e. the same ratings to both outcomes). An independent sample t-test comparing ratings of participants who provided a higher rating to the outcome on the true-discrimination trials and those participants who provided a higher rating to the outcome on the pseudo-discrimination trials revealed a significant difference between their ratings, $t(77) = -8.99$, $p = < .001$. As such, this analysis provides some support to the idea that there are individual differences in the interaction effects displayed by participants with some participants displaying a facilitation effect and some displaying a competition effect.

A growing body of literature has provided evidence that individual differences influence associative learning in humans, thus accounting for some of the variation in learning studies (Byrom & Murphy, 2016; Byrom & Murphy, 2018; Crookes & Moran, 2003). For instance, Byrom and Murphy (2016) found that attentional scope influenced configural associative learning in participants in a manner that could not be accounted for by appealing to the speed or strength of learning. However, one problem with this account of these findings is that it is difficult to ascertain what these individual differences may be or how they have impacted upon the results of these experiments without an additional task or measure of individual differences. Moreover, it is unclear as to why individual differences would influence participants' ratings when an outcome relative validity procedure was employed but not when outcome blocking, or overshadowing procedures were employed.

Given the results of Experiments 10 and 11 it is tempting to suggest that they provide support to the notion that outcomes do not compete for

associability in the same manner as cues, an assumption inherent in many learning models (e.g., Esber & Haselgrove, 2011; Le Pelley, 2004; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972). Nevertheless, this suggestion is undermined by the results of Experiment 9. This experiment provides evidence of outcome elements interacting and doing so in a manner consistent with the effect observed in cues. This is not the first report of such an effect either, with comparable findings reported in both animals and humans employing different experimental procedures (Arcediano, Matute, Escobar & Miller, 2005; Burger, Mallemat & Miller, 2000).

An alternative account of these findings could be provided when considering the procedural requirements of the outcome relative validity task. In Experiments 10 and 11 the common outcomes appearing on the true-discrimination trials and the pseudo-discrimination trials are presented more frequently than other outcomes in the experiment. This is also the case with the outcomes of interest in the outcome overshadowing experiment. Unlike the outcome overshadowing experiment, however, the outcomes of interest in the relative validity task are presented alongside several different outcomes which are either inconsistently predicted (on the pseudo-discrimination trials) or presented less frequently (on the true-discrimination trials) than the two primary outcomes of interest (O3 and O6). As such, it is possible that participants are inadvertently directed to focus on the two outcomes of primary interest, as these outcomes are either presented the most or predicted consistently, and therefore learn about these stimuli to the same extent.

The results of Experiment 10 and 11 also demonstrate that the type of task participants completed had little effect. According to causal model theory

competition effects will be exhibited amongst causal versions of a causal judgement task but not diagnostic tasks. Indeed, Van Hamme et al. (1993) employed both task types using a relative validity procedure in cues. In their study relative validity was observed amongst cues in a causal judgement task but not a diagnostic judgement task, therefore providing support to causal model theory. In Experiments 10 and 11, however, task type had little effect on participants learning or test ratings.

In summary, three experiments explored interaction effects amongst the elements of an outcome compound. In Experiment 9 an outcome overshadowing effect was observed which is consistent with a study reported by Arcediano et al. (2005) and comparable to the effect typically employed in cue-overshadowing studies. In Experiments 10 and 11 an outcome relative validity procedure was employed, and no interaction effect was observed, with the outcomes receiving comparable ratings which is inconsistent with the results of Matute, Arcediano and Miller (1996).

Chapter 5: General Discussion

5.1. Summary of findings

The experiments reported in the previous chapters sought to assess the extent to which the predictability of an outcome can influence subsequent learning featuring this outcome. As discussed in Chapter 1, a large body of literature has demonstrated that the prior associative history of a cue can impact upon subsequent learning involving this cue (e.g., Bonardi, Graham, Hall & Mitchell, 2005; Le Pelley, Turnbull, Reimers & Knipe, 2010; Le Pelley, Vadillo & Luque, 2013; O'Brien & Raymond, 2012). However, much less research has explored how the prior associative history of an outcome can influence subsequent learning involving this outcome.

In Experiments 1 to 4 participants completed a serial letter-prediction task. When completing this task in Experiment 1 participants demonstrated better learning about an outcome which had been previously well-predicted when paired with a cue it had not been paired with previously, relative to an outcome which had previously been inconsistently predicted. Experiment 2 replicated these findings and did so using a procedure which removed a potential confound in Experiment 1, in the form of an inhibitory association impeding learning about the less-well predicted outcome. Experiment 3 demonstrated that the effect was absent when the Stage 1 training trials that established the predictability of the outcomes were absent in Stage 2, potentially suggesting that the effect is sensitive to context change or rapid beta decay (i.e., the associability of the outcomes declines rapidly). Experiment 4 demonstrated that the effect is dependent on the presence of Stage 1 training.

In the absence of Stage 1 training both outcomes were learned about at a comparable rate. These results suggest the prior predictability of an outcome is important in influencing subsequent learning about this outcome when separated the cue which it was first paired with.

Experiment 5 – 8 sought to explore whether an outcome interaction effect could be observed in an outcome blocking procedure. Experiment 5 demonstrated that, in contrast to a blocking effect, a previously well-predicted outcome facilitated learning about a novel outcome when the two were presented in compound and predicted by the same cue. In this experiment participants completed a food allergist task where they were required to identify which foods produced allergic reactions in a fictitious patient. In this experiment participants completed a causal version of the task where the foods served as cues whilst the reactions served as outcomes. In Experiment 6 participants completed both a causal version of the task and a diagnostic version of the task. In the diagnostic version of the task participants were first presented with the reaction which a fictitious patient experienced prior to being presented with the foods which may have caused it. Again, the same effect was displayed regardless of task type. That is, a previously well-predicted outcome facilitated learning about a novel outcome when the two were presented in compound and predicted by the same cue. Experiment 7 sought to assess whether the effect could be observed when using an alternative type of control trial. On a control trial in Experiments 5 and 6, a cue reliably predicted one outcome in Stage 1, before predicting two outcomes the cue had not been paired with previously.

In Experiment 7, an additional control trial was employed, referred to as the Kamin control as the control is comparable to that employed by Kamin (1968, 1969) in his seminal cue-blocking studies. That is, the control trial involved simply presenting participants with a new cue in Stage 2 which reliably predicted a an outcome compound (i.e., these cues were not presented in Stage 1). When employing this additional set of control trials, an interaction effect was again observed, however, this time an outcome blocking effect was observed as opposed to an outcome facilitation effect. That is, participants demonstrated better learning about the novel outcomes paired with the Kamin control stimuli than the blocked outcome. Two accounts of the interaction effects observed in Experiments 5 – 7 were derived from Wagner’s SOP model (1981) and the Mackintosh model (1975). Interestingly, although both models can account for the findings of these experiments they do so in different ways.

According to the Mackintosh model a facilitation effect is observed with the control trial employed in Experiments 5 and 6 as the stimuli present on the control trials are rendered poor predictors, as they predict different outcomes during the two stages of training, and thus their associability declines relative to the stimuli presented on the blocking trials. With regards to the Kamin control, these stimuli do not have this reduced associability and instead are arguably the most reliable predictors of the novel outcomes as they have not been presented before and thus have not predicted any other outcomes in Stage 1. According to the SOP model facilitation occurs as the Stage 1 element of the outcome compound for the blocking trials is primed into its A2 state, thus enabling the novel outcome to fully occupy its A1 state where better learning can take place. The same cannot be said of the Kamin control trials,

however, as the outcomes are paired with novel cues. Given the novelty of the cues though it is possible to suggest they more fully occupy their A1 state relative to the stimuli which had been presented in Stage 1 and are therefore potentially primed into their A2 state.

Notably, the two accounts also differ in terms of when they anticipate an effect would be present. According to the Mackintosh model it is on the second trial of Stage 2 that the associability of the stimuli is updated and thus differences in learning about the stimuli should be displayed. According to the SOP model, however, the effect should be present after only one trial of Stage 2. Experiment 8 sought to test these diverging predictions by employing a one-trial outcome blocking procedure. In this experiment an outcome facilitation effect was present, but not the outcome blocking effect observed in Experiment 7 (with the addition of the Kamin control). These results provide partial support to the SOP account but are less easily accounted for by the Mackintosh model.

Experiments 9 sought to explore whether an interaction effect could be observed in an outcome overshadowing task. According to the SOP model an overshadowing effect should be displayed where both elements of an outcome compound reciprocally prevent one another from being learned about to the same extent as an outcome element presented on its own. In this experiment participants did indeed demonstrate better learning about an outcome element relative to the elements of the outcome compound, thus demonstrating a reciprocal overshadowing effect. Participants again demonstrated this effect regardless of whether they completed a causal or diagnostic judgement task. These results provide further support to the SOP

model and the account derived from the model to accommodate the results of Experiments 5 – 8. Experiment 10 assessed whether an interaction effect could be observed in an outcome relative validity task. Given the results of the previous experiments it was anticipated that an effect would be observed, however, no interaction effect was observed. Experiment 11 sought to assess if the duration of training in Experiment 10 might have influenced the results and potential obscured an interaction effect. Again, no effect was observed in Experiment 11 and this occurred regardless of whether participants completed a causal or diagnostic judgment task.

5.2. Outcome Predictability

Returning to the results of Experiments 1 to 4, these experiments suggest that the prior predictability of an outcome can influence subsequent learning involving this outcome when separated from the cue it was first paired with. Previous studies have provided evidence to suggest the prior predictability of outcomes can influence subsequent learning, however, in these studies the outcome has not been separated from the cue which first established the predictability of the outcome (e.g., Baker & Mackintosh, 1977; Baker & Mercier, 1982; Kimble & Ost, 1961; Kremer, 1971; Marcos & Redondo, 2001; Randich & LoLordo, 1979). The findings of these experiments are consistent with the findings of more recent studies, however, which have also separated the outcome from the cue which it was first paired with and which have demonstrated an outcome predictability effect (e.g., Griffiths, Erlinger, Beesley & Le Pelley, 2018; Griffiths, Mitchell, Bethmont & Lovibond, 2015; Thorwart, Livesey, Wilhelm, Liu & Lachnit, 2017). In these

studies, better learning has been observed about an outcome which has been well-predicted in the past.

Taken together the results of these experiments suggest that a revision to key associative learning models may be necessary to accommodate these findings, as these models remain silent as to whether or not the associability of an outcome can alter in the same manner as the associability of cues. Given these results, one possible suggestion is that a mechanism comparable to the Mackintosh rule for changes in alpha may be suitable to account for changes in beta. That is, outcomes which are well-predicted are learned about better than outcomes which are poorly predicted. For this to be possible a change to the Mackintosh model would be necessary, whereby the parameter β is granted an adjustable associability:

$$\Delta V_A = \alpha_A \beta_Y (\lambda - V_A) \quad (7)$$

In this equation, β_Y is the associability of the outcome paired with CS A and is determined by how well-predicted it has been on previous trials, which is determined by the following formulae:

$$\Delta \beta_Y > 0 \text{ if } |\lambda - V_Y^{n-1}| < |\lambda - V_Z^{n-1}| \quad (8)$$

$$\Delta \beta_Y < 0 \text{ if } |\lambda - V_Y^{n-1}| \geq |\lambda - V_Z^{n-1}| \quad (9)$$

In the above formulae β_Y represents the associability of the outcome ‘Y’. According to these formulae outcome ‘Y’ will receive an enhancement in

its associability in comparison to outcome 'Z', if 'Y' has been better predicted on a previous trial than 'Z'. If, however, other outcomes present are better predicted outcomes (or equally predictive as 'Y') then the associability of 'X' will decrease. Thus, in terms of the experiments presented here, the above formulae would suggest that the predicted outcome, as a result of its prior predictability, received an enhancement in its associability, therefore facilitating its ability to enter into novel associations when reliably paired with a novel cue in Stage 2. However, the results of Experiment 3, where the Stage 1 trials which established the predictability of the outcomes were omitted from Stage 2, pose a problem to this suggestion. In the absence of these trials no difference occurred in participants learning about the outcomes. One possible explanation which could account for the finding of Experiment 3 is that any modifications to the associability of an outcome decays rapidly in the absence of the Stage 1 training trials which first established the predictability of the outcome, or at least much faster than the associability of cues (McLaren & Mackintosh, 2000).

Another possible explanation is that the absence of these trials constitutes a change to the task context which results in a resetting in the associability of an outcome; of course both of these accounts could also be partially responsible. However, it is worth noting that other experiments do demonstrate the effect in the absence of the Stage 1 training trials such as Thorwart et al. (2017). Interestingly these authors also noted that whilst both alpha and beta can be modified based on their associative history, they do not transfer between the two when the functional nature of the two changes between stages (i.e. a previously predictable outcome serves as a cue in

another stage of training). This therefore suggests that if a Mackintosh approach is adopted to account for changes in the associability of an outcome, stimulus predictability alone is not sufficient in determining learning and that the functional nature of the stimulus must remain the same.

Another possible suggestion to account for the outcome predictability effect appeals to notion that lambda (λ) may be modifiable in these models as opposed to beta. Griffiths and Thorwart (2017) note that whilst both beta and lambda are often assumed to be fixed parameters, it is possible that the lambda may decrease, as opposed to beta, when an outcome is rendered unpredictable. Recall, that lambda in these models refers to the maximum amount of conditioning support by the outcome. If an outcome which is inconsistently predicted experiences a devaluation in the amount of learning it can support, this would account for the outcome predictability effect. As noted by Griffiths and Thorwart, this account could suggest that participants' motivation to learn dwindles after repeated difficulty in detecting when an outcome will be presented in an analogous manner to the learned helplessness effect reported by Seligman (1972).

A parsimonious explanation of the outcome predictability effect could be provided by appealing to a context-blocking account, as has been noted previously (see Chapter 2). According to this account an aspect of the context enters into an association with the unpredicted outcome and thus prevents the stimuli which are paired with the outcome from being learned about to the same extent as the well-predicted outcome. This occurs as the context becomes a predictor of the unpredicted outcome. However, an analysis concerning the relationship between the stimuli which are most likely to constitute the context,

and the effect in Experiment 2, revealed no correlation between the effect and the context. For instance, the extent to which participants formed an association between the context and the outcome did not correlate with the effect produced. Griffiths et al. (2015) also dismiss the context blocking account. Furthermore, it is not clear why this account would result in the effect being functionally specific, if the context is priming a stimulus it should not matter whether the stimulus is a US or a CS. An additional explanation could also be provided by suggesting that an interference effect influenced learning in the second stage. It is possible, for instance, that the presence of the Stage 1 training trials featuring the unpredicted outcome in Stage 2 results in an increased cognitive load concerning the unpredicted outcome. For instance, in this stage the unpredicted outcome is paired with four stimuli (e.g., F – Z, G – Z, W – Z, S – Z) whereas the well-predicted outcome is paired with two stimuli (e.g., P – X, H – X). It is therefore possible that participants have greater difficulty in recalling which stimuli are paired with the unpredicted outcome due to the greater number of associates, resulting in the effect that is observed. However, the results of Experiment 4 appear to undermine this explanation. If interference is producing the effect it would be expected that the effect would occur when Stage 2 training is provided in the absence of Stage 1 training, however, this was not the case.

5.3. Outcomes Interactions

The results of Experiments 5 and 6 demonstrate that prior learning about an outcome can facilitate learning about an additional outcome when these two outcomes are presented in compound in an outcome blocking task.

These results are somewhat inconsistent with previous studies which have employed an outcome blocking procedure. For instance, both Rescorla (1980) and Miller and Matute (1998) provide evidence of outcome blocking, where a reliably predicted outcome element blocked learning about a novel outcome when they were paired together. However, these previous studies have employed indirect conditioning methods (i.e., second-order conditioning or sensory preconditioning) with animals (i.e., rats and pigeons) using different control procedures, thus it is somewhat difficult to generalise these results to the experiments reported here. In addition, studies using comparable procedures to those used in Experiments 5 and 6 (i.e., causal judgement tasks with humans) have failed to detect an outcome interaction effect, (Cobos, Lopez, Caño, Almaraz & Shanks, 2002; Price & Yates, 1995). However, similar findings to those reported in Experiment 5 and 6 have been reported by Flach, Osman, Dickinson and Heyes (2006). These authors assessed the role of associative learning mechanisms in response priming. They found that pre-training with an element of an outcome compound enabled better learning about the added visual stimulus, therefore producing an effect comparable to that observed in Experiments 5 and 6.

Experiment 7, however, demonstrated that the interaction effect observed is dependent upon the control trial employed. When an additional control trial was employed in Experiment 7, referred to as the Kamin control, where a novel set of cues were paired with an outcome compound in Stage 2 of training, participants demonstrated better learning about the outcomes, relative to those outcomes which were presented alongside stimuli which had been previously well-predicted. That is, outcome blocking was observed. Taken

together the results of Experiments 5 to 7 provide further evidence that the prior associability of an outcome can influence subsequent learning featuring this outcome. Perhaps most importantly, these results demonstrate that the predictability of an outcome can not only influence subsequent learning about this outcome, as demonstrated in Experiments 1 and 2, but it can also influence learning about additional outcomes which are presented alongside previously predicted outcomes.

Given the results of Experiments 5 – 6, these experiments might suggest that the predictive history of an outcome is the key factor which determines the type of learning which will occur, in a similar manner to which a previously well-predicted outcome was learned about better in Experiments 1 and 2. As such, these results might suggest that a previously well-predicted outcome positively biases learning about outcomes that are paired with this outcome. The results of Experiment 7, however, suggest that the novelty of a cue-outcome pairing is also crucial in determining the amount of learning which will occur. When novel cues were paired with the outcome compounds, participants displayed better learning about the outcomes. The results of this experiment could therefore be taken to suggest that a novel cue enters into relationships with outcomes quicker than a predictive cue. Alternatively, it could be suggested that the Kamin control trials in Stage 2 produced better learning as they served as the most reliable predictors of the novel outcomes. For instance, it could be argued that the cues which reliably predicted an outcome element across the two stages of training (i.e., the outcome blocking cues) were deemed poorer predictors of the novel outcome in Stage 2 as they predicted different outcomes across different stages. That is, in Stage 1 they

predicted one outcome, in Stage 2 they predicted the same outcome and an additional outcome. It is therefore possible that participants questioned why this cue did not predict both outcomes in Stage 1. For the Kamin control cues, however, these cues only ever predicted the outcome compound. It is therefore possible that these cues were deemed the best predictors of the outcome. This interpretation would appear to be consistent with a Mackintosh (1975) account of the data.

As noted previously, Experiment 8 sought to test a Mackintosh (1975) account of the results by employing a one-trial outcome blocking procedure. According to the Mackintosh model the effects observed in Experiment 7 are dependent upon the prediction error from the first trial in Stage 2 influencing the associability of cues on the next trial. That is, on Trial 2 the blocking stimuli are established as good predictors, the controls from Stage 1 are rendered poor predictors and new contingency between the novel Kamin cue and the outcome is consolidated. Consequently, any differences should not be observed in learning about the outcomes when only one trial of training is provided in Stage 2, as the associability of the stimuli has not yet been updated. However, the results of Experiment 8 revealed that participants displayed an outcome facilitation effect at this stage, although notably outcome blocking was not displayed, perhaps providing some support to the Mackintosh model. A more complete account can be offered by the SOP model provided by Wagner (1981) though. According to this model a facilitation effect occurs as the two outcomes on the control trial in Stage 2 compete to enter into their primary A1 state where optimal associative learning can take place.

Further evidence in support of the SOP account was provided in Experiment 9. In Experiment 9 participants completed an overshadowing task and displayed an overshadowing effect consistent with that observed in cues. These results are consistent with the findings of Arcediano, Matute, Escobar and Miller (2005) who also employed an outcome overshadowing task. These results are what would be expected based on the SOP model. According to the SOP model, competition occurs between the elements of an outcome compound as they compete to enter into the A1 state which has limited capacity. However, on a control trial where an outcome element is presented on its own, the outcome encounters no competition and thus enters into its A1 state to a greater degree. Analysis of the nature of the outcome overshadowing effect also revealed a pattern of results consistent with this account. That is, participants displayed a reciprocal outcome overshadowing effect where both elements of the outcome compound are learned about less well than an outcome element presented on its own.

An alternative account of these experiments can be provided, however, which would appear to address the facilitation, blocking and overshadowing effects displayed in Experiments 5 to 9. It is possible to suggest, for instance, that a form of associative interference is producing the effects observed. As noted previously with the ‘fan effect’, Anderson (1974) observed that the more associates a stimulus accrues, the longer it takes for participants to recall information about this stimulus. In light of this, it is possible to suggest that an outcome facilitation effect occurs in Experiments 5 and 6, due to the greater number of associates involving the control stimuli. On these trials the cues were paired with 3 outcomes (e.g., C – O3, C – O1, C –

O2). On the blocking trials however, the cues are paired with two outcomes (e.g., A – O1, A - O2). It is therefore possible that the greater number of outcomes on the control trials results in poorer learning about the novel outcomes. The argument could also be made for the overshadowing effect observed in Experiment 9. On an overshadowing trial, participants are presented with a cue which is paired with two outcomes (e.g., A/B – O1 O2 / O3 O4). On a control trial however, participants are presented with a trial which is paired with only one outcome (e.g., C/D – O2/O4). Again, in accordance with the fan effect, this discrepancy could result in poorer learning about the outcomes on the overshadowing trials due to the greater number of outcomes.

However, this account would appear to encounter difficulty when considering the results of Experiment 7 and Experiments 10 and 11. In Experiment 7, the inclusion of the Kamin control trials resulted in an outcome blocking effect, in contrast to the outcome facilitation effect that was observed with a different type of control. On a Kamin control trial a cue is paired with two outcomes, the same as on a blocking trial, yet better learning is displayed about the outcome compound paired with the Kamin control, than the outcomes on the blocking trials. Given that both stimuli are paired with an equal number of outcomes it would appear difficult to account for this difference in learning. It is possible to suggest that the Stage 1 training featuring blocking trials, where a cue reliably predicts an outcome, results in greater difficulty in recalling the additional outcome which it was paired with in Stage 2. This would not be the case for the Kamin controls as they were not

presented in Stage 1. However, the results of Experiment 10 and 11 would appear to be more difficult to account for.

In Experiments 10 and 11 participants completed an outcome relative validity task. On true-discrimination trials cues were reliably paired with two outcomes. On pseudo-discrimination trials the cues were paired with three outcomes, two of which were inconsistently predicted, and one of which was consistently predicted. As such, in a similar manner to the control trials employed in Experiments 5 and 6, the cues on the pseudo-discrimination trials are paired with a greater number of outcomes. Yet, in Experiments 10 and 11, no difference was observed in learning about the outcomes of interest. If the fan effect was underpinning outcome facilitation and outcome overshadowing, it would be expected that a facilitation effect would also appear in the outcome relative validity experiments. That is, poorer learning would be demonstrated about the outcomes paired with the cues on the pseudo-discrimination trials. However, this was not the case in Experiments 10 and 11.

A Mackintosh (1975) account of Experiments 10 and 11 would also appear to encounter difficulty. According to the Mackintosh model stimuli which serve as good predictors will be learned better than stimuli which are relatively poor predictors. As such, the model would anticipate better learning about the outcomes present on the true-discrimination trials. That is, a facilitation effect comparable to that observed in Experiments 5 and 6 would be displayed. A hybrid account discussed previously, however, could accommodate these experiments, if it is assumed that two competing mechanisms are operating simultaneously producing competition for some individuals and facilitation for others. For instance, according to the SOP

model it is possible that a degree of competition occurs on the true-discrimination trials as the outcomes are both reliably predicted and compete to enter into their A1 state. On the pseudo-discrimination trials the accompanying outcomes (O4/O5) will also result in a form of competition with the common outcome (O6); however, the alternating compounds of O4O6 and O5O6 outcomes will result in extinction on the trials where O4 or O5 are expected, but not presented. This will result in these outcome elements (O4/O5) entering into A1 on each trial in which they are presented and interfering with learning about O6. If both of these effects are operating simultaneously, between or within participants, this could result in the outcomes being learned about to the same extent.

The absence of an effect in Experiments 10 and 11 is somewhat inconsistent with a similar study by Matute, Arcediano and Miller (1996). In Experiment 2 of their study these authors did not observe an outcome relative validity effect when employing a medical decision-making task. Yet, in Experiment 3 of their study they did observe an effect. In order to account for these findings, the authors suggested that the test question presented to participants was critical in determining whether an interaction effect would be observed. In Experiment 2 of their study participants were presented with a set of questions which focused on the causal role of cues (e.g., “is [cue name] the cause of the [effect name]). The authors noted that when these questions were employed in a cue competition paradigm where multiple cues were present, competition between the cues was displayed. However, when this question was asked in an outcome competition scenario where there is only one cause, no interaction effect was observed. This is assumed to have occurred as the causes

compete to be considered the primary cause of an effect, which the question prompts participants to do (i.e., *is [cue name] the cause of the [effect name]*). Yet, in the outcome competition task where there are multiple outcomes, competition did not occur as there was only one cue and the question prompts competitions between cues, not outcomes. In Experiment 3 of their study the test question was rephrased to focus on how *indicative* an effect was of a cue (e.g., *is the effect indicative of the cue*) and a relative validity effect amongst outcome was observed. The authors note this effect occurred when employing this test question, as participants compare the potential indicators of the effect, which the question they employed previously did not encourage participants to do. In essence, the question is the asymmetry of the question employed in their Experiment 2 which produced a competition effect in cues.

In Experiments 10 and 11 reported in Chapter 4, the test question asked participants how predictive a cue was of each of the outcomes and required them to make a rating concerning this relationship. The question therefore differed from that employed by Matute, Arcediano and Miller which placed emphasis on the outcome and how indicative this was of the cue. It is therefore possible that the test question could have contributed to the absence of an outcome relative validity effect. However, the same test question employed in Experiments 10 and 11 was used in Experiments 5 to 9 which displayed evidence of interaction effects amongst the elements of an outcome compound. Given the results of these experiments it is difficult to account for why the test question would be critical to producing an outcome relative validity effect but not during outcome blocking and outcome overshadowing tasks. One way to reconcile the findings of Experiments 10 and 11 with those

of Matute, Arcediano and Miller is to suggest that the way in which the cues and outcomes were presented may have influenced in the results. In the latter study the cues and outcomes were presented in a list-like manner, thus the temporal order of the cues was not clearly defined. It is therefore possible that participants completed the task diagnostically and considered the multiple effects as cues. However, this would still not account for why effects are observed in other outcome interaction tasks.

5.4. Future work

In light of the experiments reported in this thesis future work would appear to benefit from considering the mechanisms which are responsible for these effects, in order to accommodate them within associative learning models. As noted previously, given the outcome predictability results in Chapter 2, it is tempting to suggest that the Mackintosh equation for changes to alpha may be applicable to beta. However, the results of Experiments 3 suggest that changes to beta may differ from changes to alpha. When the Stage 1 training trials which established the predictability of outcomes were omitted from Stage 2 training the effect was absent. Two possible suggestions were provided to account for this finding: (1) beta decays rapidly in the absence of the trials which established the predictability of the outcome; (2) the effect is context sensitive, with the omission of Stage 1 training trials in Stage 2 constituting a change to the task context.

Both of these accounts would appear to warrant further consideration. One way of testing this first account is to systematically remove the Stage 1 trials from blocks of Stage 2 and see if this correlates with the weakening or absence of the effect. If the absence of the effect correlates with the removal of

the Stage 1 training trials, this would appear to suggest that the associability of an outcome needs to be topped up by maintaining the trials which first established the predictability of outcomes. In order to test the context sensitivity of the effect, an experimental design similar to that of Hall and Honey (1989) could also be employed. Recall, that Hall and Honey investigated the transfer of learning between two contexts with rats. Two pairs of Skinner boxes were used which were stored in different environments which served as contexts. To further differentiate the contexts, the two pairs of Skinner boxes had distinctive visual and olfactory features. This study demonstrated that whilst habituation (attenuated response) of an unconditioned response persisted despite a change in context, latent inhibition was not observed between different contexts. An analogue of this experimental design could be embedded within a serial letter prediction task, comprising of two groups. In order to form different contexts, different blocks (each consisting of two stages) containing unique stimuli pairings could be presented. One group of participants (Group Same) would experience the target outcomes in Stage 2, accompanied by the stimuli pairings (i.e. the letters) in the block they were previously presented in during Stage 1. Another group (Group Different) would receive the target outcomes in Stage 2 in a different context (i.e. presented alongside different letters) to which they were previously presented in during Stage 1. If the effect is present in Group Same but not Group Different this would suggest the effect is context sensitive in a similar manner to latent inhibition.

The role of attention in the outcome predictability effect would also appear to warrant further investigation; thus far the majority of studies which

have explored this effect have focused on causal judgement type tasks and the ratings data produced from these studies. One way of assessing the relationship between attention and learning in the outcome predictability effect is to employ an attentional dot probe task (e.g., MacLeod, Mathews & Tata, 1986), where participants are required to respond to a small dot when it is presented in a location that was occupied by a predicted or unpredicted outcome. Le Pelley, Vadillo and Luque (2013) employed an attentional dot probe task to assess the relationship between attention and learning in a learned predictiveness task. These authors found that when a small dot was presented in a location where a previously predictable cue was presented, participants responded faster to the dot than when it was presented in a location where a previously unpredictable cue was presented. As such, these authors demonstrated a relationship between attention and learning which would be anticipated by learning theories (i.e., Le Pelley, 2004; Mackintosh, 1975). That is, attention is biased toward previously predictable cues. Notably, a recent eye tracking study has also provided evidence in support of a comparable effect with outcomes (Griffiths & Le Pelley, in press).

The relationship between attention and learning would also help shed light on the mechanisms contributing toward the outcome interaction effects displayed in this thesis. For instance, a question of interest for future work concerns the relationship between the outcome facilitation and blocking effects reported in Experiments 5 to 8 and participants' overt attention as measured by eye gaze. It is possible that better learning is displayed about a novel outcome which has been presented in compound with a previously well-predicted outcome, as overt attention is biased toward the previously predictable

outcome, and in turn the novel outcome, due to their spatial contiguity. Measures of overt attention would allow this account to be tested. An additional way in which this could be tested is to spatially separate the elements of the outcome compounds so that they appear in a spatially distributed manner. If an outcome compound was still well-predicted, yet the elements of the compound were presented in the corner regions of the task display, this would allow one to assess if the spatial contiguity of the outcomes influenced learning. If the same effect occurred in this circumstance this would appear to undermine the account offered above.

Measures of overt attention would also elucidate how participants are interacting with the outcomes presented on the control trials when an outcome facilitation effect was produced. Recall, that in a first stage of training the control cues were reliably paired with an outcome element, and in a second stage of training the cues were paired with two additional outcomes. A measure of overt attention would provide an insight into the outcomes which are attracting participants' attention during Stage 2 training. It could be that participants' attention is drawn to the Stage 1 outcome which the cue predicted despite this being an ineffective strategy to learning the new contingency featuring this cue (for C/D in particular) in Stage 2. Le Pelley, Vadillo and Luque (2013) demonstrated that participants' attention seemed to be biased toward a location where a previously predictable cue was presented in an automatic manner, even when they were informed this was not the most effective strategy to locating the dot as it could occur in locations where either a predictable or unpredictable cue was presented. Eye tracking experiments would also shed light on where overt attention is allocated in an outcome

overshadowing task. The ratings data provided evidence of a reciprocal overshadowing effect, where both elements of the outcome compound were rated at a comparably low level, relative to an outcome element. It is unclear, however, whether attention is distributed to the outcomes in a comparable manner.

The fact that the different control trials employed in Experiment 7 produced different effects would also appear to warrant further investigation. A variety of different control procedures have been employed in cue-blocking studies which has led to questions concerning what constitutes an effective control (Blaser, Couvillion & Bitterman, 2006; Guez & Miller, 2008). Notably, the different control trials employed in cue-blocking studies each appear to encounter a range of issues. When employing a Kamin control trial the cue presented on these trials has had no pre-exposure prior to this second stage of training, which results in a discrepancy in the pre-exposure of cues on the blocking trials and the cues on the control trials. If, however, a control stimulus is pre-exposed ($A - O1$) and is paired with a different outcome prior to a Stage 2 of training ($AX - O3$), the predictive ability of the stimulus then becomes an issue (i.e., on a blocking trials the cue becomes a good predictor, on a control trial the cue becomes a poor predictor). Alternatively, if a stimulus is presented in Stage 1 but is not paired with an outcome, the stimulus could become latently inhibited. Typically, however, these control trials have produced comparable effects in cue-blocking procedures (i.e., blocking). Yet, when employing two of these control trials in an outcome blocking study different effects were observed. Future work could explore how additional types of

control trials influence the effects observed and examine what the optimal control trials are.

A further question for future research concerns how the type of judgement participants make may influence their ratings. In the experiments reported in this thesis whether or not the participants completed a causal or diagnostic judgement task had little impact on the key test questions. Although there appeared to be an influence of task type in the training stage of Experiment 9, collectively the nature of the task type appeared to have little impact on the key data of interest. These results are inconsistent with the causal model theory (Waldmann & Holyoak, 1992) which assumes that the direction of causality is crucial in determining whether interaction effects will be present. According to causal model theory when completing diagnostic judgement tasks participants are less likely to display interaction effects, as they are more likely to consider alternative causes of the effects. It is possible, however, that the tasks completed by participants in this thesis encouraged participants to view these tasks in purely associative manner due to their relatively simple nature. Moreover, participants may also have been influenced by the test question which focused on how predictive stimuli were. In future research the test question could be rephrased to focus on the causal nature of the stimuli to see if this encouraged participants engage in making rational inferences about the stimuli.

The future work discussed thus far has focused on gaining further insight into the mechanisms underpinning the observed effects. However, the research reported in this thesis has practical implications which would also appear to warrant further investigation. As noted in Chapter 1, associative

learning is essential to survival and completing a range of day-to-day tasks in a variety of contexts, be it learning the association between an auditory signal and danger (e.g., a fire alarm), or a particular medical symptom and an illness. The practical implications of this research for medical contexts would appear to be a particularly fruitful area for subsequent work to consider. Medical practitioners are often required to make causal or diagnostic judgements concerning the relationship between symptoms and illnesses which are displayed in a patient. In this context, symptoms can be deemed as cues and the illnesses as outcomes – although this can be reversed, a doctor may be presented with an illness and have to identify the cause. A key question of interest is whether medical practitioners display biases in learning about outcomes based on their associative history, or the history of other outcomes which are presented. For instance, consider a scenario where a medical practitioner may be presented with a symptom which reliably predicted an illness which is indicative of a particular condition. Subsequently, this symptom reliably predicts a second type of illness which occurs in compound with the first illness. In this context, the question of interest pertains to whether or not the identification of this second illness is more likely given its pairing with a previously well-predicted outcome (i.e., outcome facilitation) or whether learning about this additional outcome is impaired (i.e., outcome blocking). Alternatively, if two illnesses are presented simultaneously would this prevent both illnesses from being identified to the same extent as they would if presented alone (e.g., Quigley & Reed, 2017). Whether or not these biases are presented in medical contexts and how these biases can be overcome would appear to be an important area of consideration.

5.5. Conclusion

The experiments in this thesis sought to assess whether the prior predictability of an outcome can influence subsequent learning involving this outcome. A large body of research has demonstrated that the associative past of a cue can influence subsequent learning featuring this cue. However, considerably less research has explored how the associative history of an outcome can influence subsequent learning featuring this outcome. Key associative learning models reflect this discrepancy; whilst these models outline how the prior associability of a cue can influence subsequent learning involving this cue, they remain relatively silent about this matter in relation to outcomes (e.g., Esber & Haselgrove, 2011; Le Pelley, 2004; Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972). The experiments in this thesis suggest that the associative history of outcomes can, at the very least, influence future learning in a manner somewhat comparable to cues. In Chapter 2 the prior predictability of an outcome influenced learning about this outcome in a manner consistent with learned predictiveness studies: participants learned better about an outcome that was well-predicted relative to an outcome that was less well-predicted. The results of Chapter 3 also suggest that the prior predictability of an outcome can influence learning about outcomes which are presented in compound with the previously presented outcome. In Chapter 3 both facilitation and blocking were observed dependent on the control trial employed. That is, when a novel outcome element was presented in compound with a well-predicted outcome this resulted in better learning or impaired learning about the outcome dependent on the control trials

employed. These results are comparable to cue-interaction effects (e.g., blocking and augmentation). The results of Chapter 4 provide further evidence of interaction effects in an outcome overshadowing study, although not when employing an outcome relative validity task. A theoretical account derived from the SOP model (1981), together with that provided by Mackintosh (1975) would appear to best accommodate these results. The challenge for future work is to explore the mechanisms underpinning these results and to assess the generalisability of these findings to a range of contexts.

References

- Abramson, C. I., & Bitterman, M. E. (1986). Latent inhibition in honeybees. *Animal Learning & Behavior*, 14(2), 184-189.
- Allswede, D., Curley, K., Cullen, N., & Batsell Jr, W. R. (2014). Within-compound associations mediate augmented flavor aversions. *Learning and Motivation*, 48, 33-46.
- Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive psychology*, 6(4), 451-474.
- Arcediano, F., Matute, H., Escobar, M., & Miller, R. R. (2005). Competition between antecedent and between subsequent stimuli in causal judgments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31(2), 228.
- Bakal, C. W., Johnson, R. D., & Rescorla, R. A. (1974). The effect of change in US quality on the blocking effect. *Pavlovian Journal of Biological Science*, 9, 97-103.
- Baker, & Mackintosh. (1977). Excitatory and inhibitory conditioning following uncorrelated presentations of CS and UCS. *Animal Learning & Behavior*, 5(3), 315-319.

- Baker, A. G., & Mackintosh, N. J. (1977). Excitatory and inhibitory conditioning following uncorrelated presentations of CS and UCS. *Animal Learning & Behavior*, 5(3), 315-319.
- Baker, A. G., & Mackintosh, N. J. (1979). Preexposure to the CS alone, US alone, or CS and US uncorrelated: Latent inhibition, blocking by context or learned irrelevance? *Learning and Motivation*, 10(3), 278-294.
- Baker, A. G., & Mercier, P. (1982). Manipulation of the apparatus and response context may reduce the US Preexposure interference effect. *Quarterly Journal of Experimental Psychology Section B-Comparative and Physiological Psychology*, 34, 221-234
- Batsell Jr, W. R., & Batson, J. D. (1999). Augmentation of taste conditioning by a preconditioned odor. *Journal of Experimental Psychology: Animal Behavior Processes*, 25(3), 374.
- Batsell, W. R., Wakefield, E., Ulrey, L. A., Reimink, K., Rowe, S. L., & Dexheimer, S. (2012). CS-US interval determines the transition from overshadowing to potentiation with flavor compounds. *Learning & behavior*, 40(2), 180-194.

- Batson, J. D., & Batsell, W. R. (2000). Augmentation, not blocking, in an A+/AX+ flavor-conditioning procedure. *Psychonomic Bulletin & Review*, 7(3), 466-471.
- Blaisdell, A. P., Bristol, A. S., Gunther, L. M., & Miller, R. R. (1998). Overshadowing and latent inhibition counteract each other: Support for the comparator hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes*, 24(3), 335.
- Blaser, R. E., Couvillon, P. A., & Bitterman, M. E. (2006). Blocking and pseudoblocking: New control experiments with honeybees. *The Quarterly journal of experimental psychology*, 59(1), 68-76.
- Bonardi, C. (1991). Blocking of occasion setting in feature-positive discriminations. *The Quarterly Journal of Experimental Psychology*, 43(4), 431-448.
- Bonardi, C., & Hall, G. (1996). Learned irrelevance: No more than the sum of CS and US preexposure effects? *Journal of Experimental Psychology: Animal Behavior Processes*, 22(2), 183–191.
- Bonardi, C., & Yann Ong, S. (2003). Learned irrelevance: a contemporary overview. *The Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology*, 56(1), 80–89.

- Bonardi, C., Brilot, B., & Jennings, D. J. (2016). Learning about the CS during latent inhibition: Preexposure enhances temporal control. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42(2), 187.
- Bonardi, C., Graham, S., Hall, G., & Mitchell, C. (2005). Acquired distinctiveness and equivalence in human discrimination learning: evidence for an attentional process. *Psychonomic Bulletin & Review*, 12(1), 88–92.
- Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigms of Pavlovian learning. *Psychological bulletin*, 114(1), 80.
- Bouton, M. E., Jones, D. L., McPhillips, S. A., & Swartzen-Truber, D. (1986). Potentiation and overshadowing in odor-aversion learning: Role of method of odor presentation, the distal-proximal cue distinction, and the conditionability of odor. *Learning & Motivation*, 17, 115-138.
- Bowditch, W. A., Verbruggen, F., & McLaren, I. P. (2016). Associatively mediated stopping: Training stimulus-specific inhibitory control. *Learning & Behavior*, 44(2), 162-174.
- Brandon, S. E., Vogel, E. H., & Wagner, A. R. (2003). Stimulus representation in SOP: I: Theoretical rationalization and some implications. *Behavioural Processes*, 62(1-3), 5-25.

- Burger, D. C., Mallemat, H., & Miller, R. R. (2000). Overshadowing of subsequent events and recovery thereafter. *The Quarterly Journal of Experimental Psychology Section B*, 53(2b), 149-171.
- Byrom, N. C., & Murphy, R. (2016). Individual Difference in Configural Associative Learning. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42(4), 325-335.
- Byrom, N. C., Msetfi, R. M., & Murphy, R. A. (2018). Human latent inhibition: Problems with the stimulus exposure effect. *Psychonomic bulletin & review*, 25(6), 2102-2118.
- Chamizo, V. D., & Mackintosh, N. J. (1989). Latent learning and latent inhibition in maze discriminations. *The Quarterly Journal of Experimental Psychology*, 41(1), 21-31.
- Channell, S., & Hall, G. (1983). Contextual effects in latent inhibition with an appetitive conditioning procedure. *Animal Learning & Behavior*, 11(1), 67-74.
- Cheng, K., & Spetch, M. L. (2001). Blocking in landmark-based search in honeybees. *Animal Learning & Behavior*, 29(1), 1-9.

- Clarke, J. C., Westbrook, R. F., & Irwin, J. (1979). Potentiation instead of overshadowing in the pigeon. *Behavioral and Neural Biology*, 25(1), 18-29.
- Cobos, P. L., López, F. J., Cano, A., Almaraz, J., & Shanks, D. R. (2002). Mechanisms of predictive and diagnostic causal induction. *Journal of Experimental Psychology: Animal Behavior Processes*, 28(4), 331.
- Crookes, A. E., & Moran, P. M. (2003). An investigation into age and gender differences in human Kamin blocking, using a computerized task. *Developmental Neuropsychology*, 24(1), 461-477.
- Denniston, J. C., Savastano, H. I., & Miller, R. R. (2001). The extended comparator hypothesis: Learning by contiguity, responding by relative strength. *Handbook of contemporary learning theories*, 3, 65-117.
- Dickinson, A., Nicholas, D. J., & Mackintosh, N. J. (1983). A re-examination of one-trial blocking in conditioned suppression. *The Quarterly Journal of Experimental Psychology*, 35(1), 67-79.
- Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgement of act-outcome contingency: The role of selective attribution. *The Quarterly Journal of Experimental Psychology*, 36(1), 29-50.

Dienes, Z. (2011). Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science*, 6(3), 274-290.

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, 5, 1–17.

Donders, F. C. (1969). On the speed of mental processes. *Acta psychologica*, 30, 412-431.

Donegan, N. H. (1981). Priming-produced facilitation or diminution of responding to a Pavlovian unconditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes*, 7(4), 295–312.

Dopson, J. C., Pearce, J. M., & Haselgrove, M. (2009). Failure of retrospective revaluation to influence blocking. *Journal of Experimental Psychology: Animal Behavior Processes*, 35(4), 473.

Durlach, P. J., & Rescorla, R. A. (1980). Potentiation rather than overshadowing in flavor-aversion learning: an analysis in terms of within-compound associations. *Journal of Experimental Psychology: Animal Behavior Processes*, 6(2), 175.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of experimental psychology: human perception and performance*, 27(1), 229.

- 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.
- Estévez, A. F., Fuentes, L. J., Mari-Bêffa, P., González, C., & Alvarez, D. (2001). The differential outcome effect as a useful tool to improve conditional discrimination learning in children. *Learning and Motivation*, 32(1), 48-64.
- Evans, L. H., Gray, N. S., & Snowden, R. J. (2007). A new continuous within-participants latent inhibition task: Examining associations with schizotypy dimensions, smoking status and gender. *Biological Psychology*, 74(3), 365–373.
- Flach, R., Osman, M., Dickinson, A., & Heyes, C. (2006). The interaction between response effects during the acquisition of response priming. *Acta Psychologica*, 122(1), 11-26.
- Forrest, D. R. L., Mather, M., & Harris, J. A. (2016). Unmasking latent inhibition in humans. *The Quarterly Journal of Experimental Psychology*, 71(2), 380-395.

- Galef, B. G., & Osborne, B. (1978). Novel taste facilitation of the association of visual cues with toxicosis in rats. *Journal of Comparative and Physiological Psychology*, 92(5), 907.
- George, D. N., & Pearce, J. M. (2012). A configural theory of attention and associative learning. *Learning & Behavior*, 40(3), 241-254.
- Gillan, D. J., & Domjan, M. (1977). Taste-aversion conditioning with expected versus unexpected drug treatment. *Journal of Experimental Psychology: Animal Behavior Processes*, 3(4), 297.
- Ginton, A., Urca, G., & Lubow, R. E. (1975). The effects of preexposure to a nonattended stimulus on subsequent learning: Latent inhibition in adults. *Bulletin of the Psychonomic Society*, 5(1), 5-8.
- Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning: An adaptive network model. *Journal of Experimental Psychology: General*, 117(3), 227
- Good, I., Allswede, D., Curley, K., & Batsell Jr, W. R. (2015). Increasing trial number produces augmentation, not blocking, in flavor-aversion conditioning. *Learning and Motivation*, 49, 38-50.

Granger, K. T., Moran, P. M., Buckley, M. G., & Haselgrove, M. (2016).

Enhanced latent inhibition in high schizotypy individuals. *Personality and Individual Differences*, 91, 31-39.

Griffiths, O., & Le, M. P. (in press). The outcome predictability bias is evident in overt attention. *Journal of experimental psychology. Animal learning and cognition*.

Griffiths, O., Erlinger, M., Beesley, T., & Le Pelley, M. E. (2018). Outcome predictability biases cued search. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 44(8), 1215-1223.

Griffiths, O., Livesey, E., & Thorwart, A. (2019). Learned biases in the processing of outcomes: A brief review of the outcome predictability effect. *Journal of Experimental Psychology: Animal Learning and Cognition*, 45, 1-16.

Griffiths, O., Mitchell, C. J., Bethmont, A., & Lovibond, P. F. (2015).

Outcome Predictability Biases Learning. *Journal of Experimental Psychology: Animal Learning and Cognition*, 41(1), 1–17.

Griffiths, O., & Thorwart, A. (2017). Effects of Outcome Predictability on Human Learning. *Frontiers in Psychology*, 8(12).

Guerrieri, F., Lachnit, H., Gerber, B., & Giurfa, M. (2005). Olfactory blocking and odorant similarity in the honeybee. *Learning & Memory*, 12(2), 86-95.

Guez, D., & Miller, R. R. (2008). Blocking and pseudoblocking: The reply of *Rattus norvegicus* to *Apis mellifera*. *The Quarterly Journal of Experimental Psychology*, 61(8), 1186-1198.

Hall, G., & Channell, S. (1985). Differential effects of contextual change on latent inhibition and on the habituation of an orienting response. *Journal of Experimental Psychology: Animal Behavior Processes*, 11(3), 470.

Hall, G., & Honey, R. C. (1989). Contextual Effects in Conditioning, Latent Inhibition, and Habituation - Associative and Retrieval Functions of Contextual Cues. *Journal of Experimental Psychology: Animal Behavior Processes*, 15(3), 232-241.

Hall, G., & Pearce, J. M. (1979). Latent inhibition of a CS during CS-US pairings. *Journal of Experimental Psychology: Animal Behavior Processes*, 5(1), 31.

Haselgrove, M., Esber, G. R., Pearce, J. M., & Jones, P. M. (2010). Two kinds of attention in Pavlovian conditioning: evidence for a hybrid model of

learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 36(4), 456–70.

Holland, P. C. (1984). Unblocking in Pavlovian appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 10(4), 476.

Honey, R. C., & Hall, G. (1988). Overshadowing and blocking procedures in latent inhibition. *The Quarterly Journal of Experimental Psychology*, 40(2), 163-180.

Hull, C. L. (1930). Simple trial and error learning: A study in psychological theory. *Psychological Review*, 37(3), 241.

Hullinger, R. A., Kruschke, J. K., & Todd, P. M. (2015). An evolutionary analysis of learned attention. *Cognitive science*, 39(6), 1172-1215.

Jones, P. M., & Haselgrove, M. (2013). Overshadowing and associability change: Examining the contribution of differential stimulus exposure. *Learning & Behavior*, 41(1), 107–117.

Kamin, L.J. (1968). “Attention-like” processes in classical conditioning. In M.R. Jones (Ed.). *Miami Symposium on the Prediction of Behavior: Aversive Stimulation*. Coral Gables, Florida (pp. 9-31). Florida: University of Miami Press.

- Kamin, L.J. (1969a). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and Aversive Behavior* (pp. 279-296). New York: Appleton-Century-Crofts.
- Kimble, G. a, & Ost, J. W. (1961). A conditioned inhibitory process in eyelid conditioning. *Journal of Experimental Psychology*, 61(2), 150–156.
- Kremer, E. F. (1971). Truly random and traditional control procedures in CER conditioning in the rat. *Journal of Comparative & Physiological Psychology*, 76(3), 441–448.
- Le Pelley, M. E. (2004). The role of associative history in models of associative learning: a selective review and a hybrid model. *The Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology*, 57(3), 193–
- Le Pelley, M. E., & McLaren, I. P. L. (2003). Learned associability and associative change in human causal learning. *The Quarterly Journal of Experimental Psychology*, 56B(1), 68–79.
- Le Pelley, M. E., Mitchell, C. J., & Johnson, A. M. (2013). Outcome value influences attentional biases in human associative learning: Dissociable effects of training and instruction. *Journal of Experimental Psychology: Animal Behavior Processes*, 39(1), 39.

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.

Le Pelley, M. E., Oakeshott, S. M., & McLaren, I. P. L. (2005). Blocking and

unblocking in human causal learning. *Journal of Experimental Psychology: Animal Behavior Processes*, 31(1), 56.

Le Pelley, M. E., Pearson, D., Griffiths, O., & Beesley, T. (2015). When goals

conflict with values: Counterproductive attentional and oculomotor capture by reward-related stimuli. *Journal of Experimental Psychology: General*, 144(1), 158.

Le Pelley, M. E., Turnbull, M. N., Reimers, S. J., & Knipe, R. L. (2010).

Learned predictiveness effects following single-cue training in humans. *Learning and Behavior*, 38(2), 126–144.

Le Pelley, M. E., Vadillo, M., & Luque, D. (2013). Learned predictiveness

influences rapid attentional capture: evidence from the dot probe task. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 39(6), 1888–900.

- Livesey, E. J., & McLaren, I. P. L. (2007). Elemental associability changes in human discrimination learning. *Journal of Experimental Psychology. Animal Behavior Processes*, 33(2), 148–159.
- LoLordo, V. M., & Droungas, A. (1989). Selective associations and adaptive specializations: Taste aversions and phobias. In: S. B. Klein & R. R. Mowrer (Eds.), *Contemporary learning theories: Instrumental conditioning theory and the impact of biological constraints on learning* (pp. 145-179). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Lubow, R. E. (1973). Latent inhibition. *Psychological bulletin*, 79(6), 398.
- Lubow, R. E., & Moore, A. U. (1959). Latent inhibition: the effect of nonreinforced pre-exposure to the conditional stimulus. *Journal of comparative and physiological psychology*, 52(4), 415.
- Lubow, R. E., Markman, R. E., & Allen, J. (1968). Latent inhibition and classical conditioning of the rabbit pinna response. *Journal of Comparative and Physiological Psychology*, 66(3), 688.
- Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82(4), 276–298.

Mackintosh, N. J. (1975b). Blocking of conditioned suppression: Role of the first compound trial. *Journal of Experimental Psychology: Animal Behavior Processes*, 1, 335-45.

Mackintosh, N.J. (1976). Overshadowing and stimulus intensity. *Animal Learning and Behaviour*, 4, 186-192.

Mackintosh, N. J., & Little, L. (1969). Intradimensional and extradimensional shift learning by pigeons. *Psychonomic Science*, 14(1), 5–6.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of abnormal psychology*, 95(1), 15.

Marcos, J. L., & Redondo, J. (2001). Relation between conditioned stimulus-elicited responses and unconditioned response diminution in long-interval human heart-rate classical conditioning. *The Spanish Journal of Psychology*, 4(1), 11–8.

Matute, H., Arcediano, F., & Miller, R. R. (1996). Test question modulates cue competition between causes and between effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(1), 182.

Matzel, L. D., Schachtman, T. R., & Miller, R. R. (1988). Learned irrelevance exceeds the sum of CS-preexposure and US-preexposure deficits.

Journal of Experimental Psychology: Animal Behavior Processes,
14(3), 311.

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

McLaren, I. P. L., McAndrew, A., Angerer, K., McLaren, R., Forrest, C.,
Bowditch, W., ... & Verbruggen, F. (2019). Mackintosh lecture:
Association and cognition: Two processes, one system. *Quarterly
Journal of Experimental Psychology*, 72(2), 98-117.

Melara, R. D., & Algom, D. (2003). Driven by information: A tectonic theory
of Stroop effects. *Psychological Review*, 110, 422–471.

Miller, R. R., & Matute, H. (1998). Competition between outcomes.
Psychological Science, 9(2), 146-149.

Miller, R. R., & Matzel, L. D. (1988). The comparator hypothesis: A response
rule for the expression of associations. In: *Psychology of learning and
motivation* (Vol. 22, pp. 51-92). Academic Press.

Miller, R. R., Barnet, R. C., & Grahame, N. J. (1995). Assessment of the
Rescorla-Wagner model. *Psychological bulletin*, 117(3), 363.

Miyashita, Y., Nakajima, S., & Imada, H. (2000). Differential outcome effect in the horse. *Journal of the Experimental Analysis of Behavior*, 74(2), 245-253.

Mok, L. W., & Overmier, J. B. (2007). The differential outcomes effect in normal human adults using a concurrent-task within-subjects design and sensory outcomes. *The Psychological Record*, 57(2), 187-200.

O'Brien, J. L., & Raymond, J. E. (2012). Learned predictiveness speeds visual processing. *Psychological Science*, 23(4), 359-363.

Overmier, J. B., Bull, J. A., & Trapold, M. A. (1971). Discriminative cue properties of different fears and their role in response selection in dogs. *Journal of Comparative and Physiological Psychology*, 76(3), 478.

Pavlov, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, Trans.). London: Oxford University Press.

Pearce, J. M., & Bouton, M. E. (2001). *Theories of associative learning in animals. Annual Review of Psychology*, 52, 111–139.

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), 532–52.

- Pearson, D., Donkin, C., Tran, S. C., Most, S. B., & Le Pelley, M. E. (2015). Cognitive control and counterproductive oculomotor capture by reward-related stimuli. *Visual Cognition*, 23(1-2), 41-66.
- Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1-2), 8–13.
- Peirce, J. W. (2008). Generating Stimuli for Neuroscience Using PsychoPy. *Frontiers in Neuroinformatics*, 2, 10.
- Perruchet, P. (1985). A pitfall for the expectancy theory of human eyelid conditioning. *The Pavlovian journal of biological science: official journal of the Pavlovian*, 20(4), 163-170.
- Pirolli, P. (2007). *Information foraging theory: Adaptive interaction with information*. Oxford University Press.
- Pirolli, P., & Card, S. (1999). Information foraging. *Psychological review*, 106(4), 643.
- Prados, J., Alvarez, B., Acebes, F., Loy, I., Sansa, J., & Moreno-Fernández, M. (2013). Blocking in rats, humans and snails using a within-subjects design. *Behavioural processes*, 100, 23-31.

Price, P. C., & Yates, J. F. (1995). Associative and rule-based accounts of cue interaction in contingency judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(6), 1639.

Price, P.C., & Yates, J. F. (1993). Judgmental overshadowing: Further evidence of cue interaction in contingency judgment. *Memory & Cognition*, 21, 561-572.

Quigley, M., & Reed, P. (2017). Over-selective Responding in a Diagnostic Judgment Task. *Applied Cognitive Psychology*, 31(5), 558-564.

Randich, A., & LoLordo, V. M. (1979). Associative and nonassociative theories of the UCS preexposure phenomenon: Implications for Pavlovian conditioning. *Psychological Bulletin*, 86(3), 523–548.

Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of comparative and physiological psychology*, 66(1), 1.

Rescorla, R. A. (1980). *Pavlovian second-order conditioning: Studies in associative learning*. Hillsdale, NJ: Erlbaum.

Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In: Black, A. H., & Prokasy, W. F. (Eds.), *Classical conditioning II*:

Current research and theory (pp. 64-99). New York: Appleton-Century-Crofts.

Robinson, J., & Bonardi, C. (2015). An associative analysis of object memory. *Behavioural brain research*, 285, 1-9.

Rusiniak, K. W., Hankins, W. G., Garcia, J., & Brett, L. P. (1979). Flavor-illness aversions: Potentiation of odor by taste in rats. *Behavioral and Neural biology*, 25(1), 1-17.

Sanderson, D. J., Jones, W. S., & Austen, J. M. (2016). The effect of the amount of blocking cue training on blocking of appetitive conditioning in mice. *Behavioural processes*, 122, 36-42.

Schnur, P., & Ksir, C. J. (1969). Latent inhibition in human eyelid conditioning. *Journal of Experimental Psychology*, 80(2p1), 388.

Seligman, M. E. (1972). Learned helplessness. *Annual review of medicine*, 23(1), 407-412.

Shishimi, A. (1985). Latent inhibition experiments with goldfish (*Carassius auratus*). *Journal of Comparative Psychology*, 99(3), 316.

Sissons, H. T., Urcelay, G. P., & Miller, R. R. (2009). Overshadowing and CS duration: Counteraction and a reexamination of the role of within-

compound associations in cue competition. *Learning & behavior*, 37(3), 254-268.

Slotnick, B. M., Westbrook, F., & Darling, F. M. (1997). What the rat's nose tells the rat's mouth: long delay aversion conditioning with aqueous odors and potentiation of taste by odors. *Animal Learning & Behavior*, 25(3), 357-369.

Spetch, M. L., Wilkie, D. M., & Pinel, J. P. (1981). Backward conditioning: A reevaluation of the empirical evidence. *Psychological Bulletin*, 89(1), 163.

Taylor, K. M., Joseph, V. T., Balsam, P. D., & Bitterman, M. E. (2008). Target-absent controls in blocking experiments with rats. *Learning & behavior*, 36(2), 145-148.

Thorwart, A., Livesey, E. J., Wilhelm, F., Liu, W., & Lachnit, H. (2017). Learned predictiveness and outcome predictability effects are not simply two sides of the same coin. *Journal of Experimental Psychology: Animal Learning and Cognition*, 43(4), 341.

Trapold, M. A. (1970). Are expectancies based upon different positive reinforcing events discriminably different?. *Learning and Motivation*, 1(2), 129-140.

- Trapold, M. A., & Overmier, J. B. The second learning process in instrumental learning. In: A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research*. New York : Appleton-Century-Crofts, 1972.
- Urcelay, G. P. (2017). Competition and facilitation in compound conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition*, 43(4), 303.
- Urcuioli, P. J. (1991). Retardation and facilitation of matching acquisition by differential outcomes. *Animal Learning & Behavior*, 19(1), 29-36.
- Urcuioli, P. J. (2005). Behavioral and associative effects of differential outcomes in discrimination learning. *Animal Learning & Behavior*, 33(1), 1-21.
- Vadillo, M. A., & Matute, H. (2010). Augmentation in contingency learning under time pressure. *British Journal of Psychology*, 101(3), 579-589.
- Vadillo, M. A., & Matute, H. (2011). Further evidence on the validity of web-based research on associative learning: Augmentation in a predictive learning task. *Computers in Human Behavior*, 27(2), 750-754.

- Van Hamme, L. J., & Wasserman, E. A. (1994). Cue competition in causality judgments: The role of nonpresentation of compound stimulus elements. *Learning and motivation*, 25(2), 127-151.
- Van Hamme, L. J., Kao, S. F., & Wasserman, E. A. (1993). Judging interevent relations: From cause to effect and from effect to cause. *Memory & Cognition*, 21(6), 802-808.
- Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in cognitive sciences*, 12(11), 418-424.
- Wagner, A. R., Logan, F. A., Haberlandt, K., & Price, T. (1968). Stimulus selection in animal discrimination learning. *Journal of experimental psychology*, 76(2p1), 171.
- Wagner, A.R. 1981. SOP: a model of automatic memory processing in animal behavior. In: Spear, N.E., Miller, R.R. (Eds.), *Information Processing in Animals: Memory Mechanisms*. Erlbaum, Hillsdale, NJ, pp. 5–47.
- Waldmann, M. R. (2000). Competition among causes but not effects in predictive and diagnostic learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(1), 53.

- Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: asymmetries in cue competition. *Journal of Experimental Psychology: General*, 121(2), 222.
- Wasserman, E. A. (1990). Attribution of causality to common and distinctive elements of compound stimuli. *Psychological Science*, 1(5), 298-302.
- Willner, J. A. (1978). Blocking of a taste aversion by prior pairings of exteroceptive stimuli with illness. *Learning & Motivation*, 9, 125-140.
- Young, A. M. J., Kumari, V., Mehrotra, R., Hemsley, D. R., Andrew, C., Sharma, T., ... Gray, J. A. (2005). Disruption of learned irrelevance in acute schizophrenia in a novel continuous within-subject paradigm suitable for fMRI. *Behavioural Brain Research*, 156(2), 277-288.

