Understanding the Origin of Pavlovian-Instrumental Interactions

Sabrina R. Cohen-Hatton

A Thesis Submitted to Cardiff University

For the Degree of Doctor of Philosophy

May 2013

DECLARATION

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed ……………………………………… (candidate) Date …………………………

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Signed ……………………………………… (candidate) Date …………………………

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.

Other sources are acknowledged by explicit references. The views expressed are my own.

Signed ……………………………………… (candidate) Date …………………………

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ……………………………………… (candidate) Date …………………………

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.**

Signed ……………………………………… (candidate) Date …………………………

Publications

Some of the research in Chapters 3 and 4 have been presented at conferences and described in two recent papers. Chapter 3 closely follows the published paper by Cohen-Hatton, Haddon, George and Honey (2013), while Chapter 4 is based upon a paper by Cohen-Hatton and Honey, which is in press at the time of writing.

<u>Conference presentations:</u>

- **European Brain and Behaviour Society** (Munich, 2013)
    - Renewal of extinguished instrumental responses: Independence from Pavlovian processes and dependence on outcome value.
- **Gregynog Associative Learning Symposium** (Wales, 2013)
    - Renewal of extinguished instrumental responses: Independence from Pavlovian processes and dependence on outcome value.
- **Federation of European Neuroscience** (Barcelona, 2012)
    - Pavlovian-to-Instrumental Transfer: Paradoxical effects of the Pavolvian Relationship explained.
- **Gregynog Associative Learning Symposium** (Wales, 2012)
    - Palvovian-to-Instrumental Transfer: Paradoxical effects of the Pavolvian Relationship.

<u>Peer-reviewed papers:</u>

- Cohen-Hatton, S.R., Haddon, J.E., George, D.N., Honey, R.C. (2013). Pavlovian-to-instrumental transfer: Paradoxical effects of the Pavlovian relationship explained. *Journal of Experimental Psychology: Animal Behavior Processes, 39*, 14-23.

- Cohen-Hatton, S.R., & Honey, R.C. (in press). Renewal of extinguished instrumental responses: Independence from Pavlovian processes and dependence on outcome value. *Learning & Behaviour.*

Summary

This thesis investigates interactions between Pavlovian and instrumental processes. The first chapter provides an evaluation of various theoretical analyses of how these two processes might interact in the context of two types of phenomena: Pavlovian-to-instrumental transfer (PIT) and the renewal of instrumental responses that have been extinguished. It is argued that the conditions under which both phenomena are observed do not sit readily with the theoretical analyses that have been offered for them. Chapter 2 reports three experiments that examined the conditions under which outcome-selective and general PIT occur in rats. Outcome-selective PIT was not increased by procedures that should increase the distinctiveness of the outcomes; but general PIT was more likely to be observed under conditions in which the distinctiveness of the outcomes should be low (Experiments 1-3). Chapter 3 contrasted the standard stimulus-outcome-response analysis of outcome-selective PIT with a novel theoretical analysis based on mediated stimulus-response associations that directly affect test performance (i.e., without the outcome becoming activated during the test). Experiment 4 demonstrated an outcome-selective PIT effect when the outcome (O) was embedded in the Pavlovian conditioned stimulus (S), and Experiments 5 and 6 showed that outcome-selective PIT was more likely to be observed after backward pairings (i.e., O-S) than after forward pairings (i.e., S-O). These results are consistent with the following analysis: Instrumental training establishes response-outcome and outcome-response associations, and during subsequent backward conditioning the outcome provokes its associated instrumental response during the stimulus and thereby allows a stimulus–response association to be acquired. This stimulus-response association then directly generates outcome-selective PIT at test. Experiment 7 provided direct evidence to support the assumptions upon which this analysis relies. These results, together with other paradoxical effects of the Pavlovian relationship, are incongruent with accounts of outcome-selective PIT that rely on a stimulus-outcome-response chain. Chapter 4 explored another instance where Pavlovian stimuli exert a powerful influence on instrumental performance: the case of instrumental renewal. Two fundamental issues were addressed: whether or not direct Pavlovian associations are responsible for the renewal effect, and whether or not renewed responses are controlled by goal-directed processes or stimulus-response associations. In Experiment 8, instrumental renewal was observed without concomitant involvement of any excitatory or inhibitory Pavlovian properties of the contexts involving the outcome; and in Experiment 9, renewed responding was sensitive to the current value of the outcome. Taken together, these results suggest that the extinction context exerts a direct (or hierarchical) inhibitory influence on the instrumental response-outcome association, the removal of which allows the impact of the response-outcome association of performance to be revealed. Chapter 5 explores the broader implications of these results for current theoretical analyses that rely on the idea that Pavlovian and instrumental processes interact through shared access to the features of the outcome.

Table of Contents

Chapter 1. General Introduction 1

- 1.1. Background 1
- 1.2. Learning Processes 2
  - 1.2.1. Pavlovian Conditioning 2
  - 1.2.2.Instrumental Conditioning 4
- 1.3. Interactions between learning processes: Pavlovian-to-instrumental Transfer 6
- 1.4. Theoretical analyses of PIT 7
  - 1.4.1. S-O-R: A two-process theory 7
  - 1.4.2. Associative-Cybernetic model 10
  - 1.4.3. Hierarchical Theory 13
  - 1.4.4. S-R Theory 16
- 1.5. Contextual influences on instrumental behaviour 19
- 1.6. Summary of rationale for new empirical work 21

Chapter 2. The elusive nature of outcome-selective Palovian-to-instrumental transfer 23

- 2.1. Summary 23
  - 2.2. Introduction 24
- 2.3. Experiment 1 27
  - 2.3.1. Method 28
  - 2.3.2. Results and discussion 31
- 2.4. Experiment 2 33

2.4.1. Method 34

2.4.2. Results and discussion 35

2.5. Experiment 3 41

2.5.1. Method 41

2.5.2. Results and discussion 41

2.6. General discussion 46

2.7. Conclusion 48

Chapter 3. Pavlovian-to-instrumental transfer: Paradoxical effects of the Pavlovian relationship examined 49

3.1 Summary 49

3.2 Introduction 50

3.3. Experiment 4 55

3.3.2 Method 57

3.3.3 Results and discussion 60

3.4. Experiment 5 62

3.4.5. Method 63

3.4.6. Results and discussion 64

3.5. Experiment 6 67

3.5.1. Method 68

3.5.2. Results 68

3.5.3. Discussion 72

3.6. Experiment 7 73

3.6.1. Method 74

3.6.2. Results and discussion 75

3.7. General discussion 77

3.8. Conclusion 81

Chapter 4. Contextual control of instrumental behaviour: The role of context-outcome and context-response associations 82

4.1. Summary 82

4.2. Introduction 83

4.3. Experiment 8 87

4.3.1. Method 89

4.3.2. Results 91

4.3.3. Discussion 95

4.4. Experiment 9 97

4.4.1. Method 98

4.4.2. Results and discussion 100

4.5. General discussion 107

4.6. Conclusion 109

Chapter 5. General Discussion 110

5.1. Overall summary 110

5.2. Summary of results 110

5.2.1. The elusive nature of outcome-selective Pavlovian-to-instrumental transfer.... 110

5.2.2. Pavlovian-to-instrumental transfer: paradoxical effects of the pavlovian relationship explored 111

5.2.3. Contextual control of instrumental behaviour: The role of context-outcome and context-response associations 112

5.3. Theoretical analyses of interaction 113

5.3.1. S-O-R: A two-process theory 113

5.3.2. Associative-Cybernetic model 115

5.3.3. Hierarchical theory 117

5.3.4. S-R theory 117

5.4. Implications for neuroscience: PIT and renewal 118

5.4.1. PIT 118

5.4.2. Renewal 119

5.5. Future directions 120

5.6. General implications 124

5.7. General conclusions 126

References 128

List of Figures

Figure 1. A typical Pavlovian conditioning experiment. Repeated pairings of a neutral stimulus such as a tone (a conditioned stimulus or CS) with food (an unconditioned stimulus or US) leads to some behavioural change that is indicative of the formation of an association between the CS and the US. As a result the animal comes to respond to the CS in a way that often resembles how it would respond to the US.

Figure 2. Associative structures for instrumental responding. The rat performs a response (R) that is reinforced by food (O) leading to a goal-directed R-O association; or the response manipulandum becomes a cue (S) associated with the response (R).

Figure 3. An S-O-R associative chain. At test, the CS (S) evokes a representation of the outcome (O) that, in turn, activates the response (R), via the O-R association.

Figure 4. An Associative-Cybernetic model of instrumental responding (black lines represent learned associations; grey lines represent fixed connections). An instrumental response (R) can be generated (i) through a link between the O in the associative system and the corresponding motor program via a reward memory (black dotted line); (ii) via a direct link between the response and the corresponding motor program via the R in the S-R (habit) system (grey dotted line). In the case of PIT, a Pavlovian CS can activate a response through a shared outcome memory (within the associative system) which is linked to the motor program through the incentive system. Alternatively, the Pavlovian CS may trigger the response via a bi-

directional O-R link, which activates the R in the habit system, and acts on the corresponding motor program independently of the incentive system.

Figure 5. The hierarchical model of instrumental conditioning according to which stimuli (Ss) act on the association between responses (Rs) and outcomes (Os).

Figure. 6. An S-R account of PIT. (A). A bi-directional R-O association is established in the first phase of training. (B). In the second phase of training, the presentation of a common outcome (O) evokes a representation of the response (R) via the backwards O-R association in the presence of the stimulus (S). This leads to an excitatory association between S and R. (C). At test, presentations of S evoke a response through the activation of this S-R association.

Figure 7: Experiment 1: Test. Mean rates of responding (in responses per minute, rpm) on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI) when no such stimulus was present.

Figure 8: Experiment 2: Pavlovian conditioning. Mean rates of magazine entries (in responses per minute, rpm; +SEM) during Pavlovian stimuli (responding collapsed across counterbalanced factors); and during the inter-trial interval (ITI) when no such stimulus was present. (A). Intermixed group. (B). Blocked group. (C). None group.

Figure 9. Experiment 2. Instrumental responding. Mean rates of lever press responses (in responses per minute, rpm; +SEM; collapsed over counterbalanced factors) over the course of instrumental conditioning for groups Intermixed, Blocked and None.

Figure 10. Experiment 2: Test. Mean rates of responding (in responses per minute, rpm) for groups Intermixed, Blocked and None on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI) when S1 and S2 were absent.

Figure 11: Experiment 3: Pavlovian conditioning. Mean rates of magazine entries (in responses per minute, rpm; +SEM) during Pavlovian stimuli (responding collapsed across counterbalanced factors); and during the inter-trial interval (ITI). (A). Intermixed group. (B). Blocked group. (C). None group.

Figure 12: Experiment 3. Instrumental conditioning. Mean rates of lever press responses (in responses per minute, rpm; +SEM; collapsed over counterbalanced factors) over the course of instrumental acquisition for groups Intermixed, Blocked and None.

Figure 13. Experiment 3: Test. Mean rates of responding (in responses per minute, rpm) for Groups Intermixed, Blocked and None on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI).

Figure 14. Experiment 4: Test. Mean rates of responding (in responses per minute, rpm) on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI) when no such stimulus was present.

Figure 15. Experiment 5: Test. Mean rates of responding (in responses per minute, rpm) on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was presented; and during the inter-trial interval (ITI) when no such stimulus was present. Original Pavlovian training had either involved forward pairings (S1-O1) or backward pairings (O2-S2).

Figure 16. Experiment 6: Test. Mean rates of responding (in responses per minute, rpm) during Block 1 (left panel) and Block 2 (right panel) on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI) when no such stimulus was present. Original Pavlovian training had either involved forward pairings (S1-O1) or backward pairings (O2-S2).

Figure 17. Experiment 7: Test. Mean rates of congruent and incongruent responding (in responses per minute, rpm) during S1 and S2 on backward (O1-S1 and O2-S2) and forward (S1-O2 and S2-O1) trials.

Figure 18. Experiment 8: Extinction. Mean rates (+SEM) of lever pressing (with R1 and R2 pooled) and magazine entries (during periods when the levers were retracted) in responses per minute, rpm, over the course of the extinction sessions.

Figure 19. Experiment 8: Test. Mean rates of instrumental responding (in responses per minute, rpm; +SEM) as a function of whether responding was tested in the same context as extinction (i.e., R1 in B and R2 in A) or in a different context (i.e., R1 in A or R2 in B). Rats had either been sated on the outcome associated with R1 and R2 during training (O1) or with a novel outcome (O2).

Figure 20. Experiment 8: Test. Mean rates of magazine entries (in responses per minute, rpm; +SEM) over the 5-min extinction test, as a function of whether the rats had

been sated on the outcome associated with R1 and R2 during training (O1) or with a novel outcome (O2).

Figure 21. Experiment 9: Extinction of lever pressing. Mean rates of instrumental responding (in responses per minute, rpm; +SEM) during extinction for group Extinction.

Figure 22. Experiment 9: Extinction of magazine entries. Mean rates magazine entry (in responses per minute, rpm; +SEM) for group Extinction and group No Extinction. Rats in both groups had been sated with O2 prior to the test, and while rats in group Extinction had received instrumental extinction in context B, those in group No Extinction had not. Magazine entries for group Extinction are shown for the periods when the levers were retracted.

Figure 23. Experiment 9: Test. Mean rates of two responses (R1 and R2; in responses per minute, rpm; +SEM) as a function of whether testing was conducted in the conditioning context (A) or in the extinction context (B). Group Extinction had received extinction with both responses (R1 and R2) and were sated with the outcome associated with R1 (i.e., O1) in context B; and group No Extinction were simply sated with O1 in context B.

Figure 24. Experiment 9: Test. Mean rates of magazine entries (in responses per minute, rpm; +SEM) over the 5-min extinction test, as a function of whether rats were tested in context A or B. Group Extinction (upper panel) had received extinction of both instrumental responses (R1 and R2) and were sated with the outcome associated with R1 (i.e., O1) in context B; and group No Extinction were simply sated with O1 in context B.

Figure 25. An Associative-Cybernetic model of instrumental responding (black lines represent learned associations; grey lines represent fixed connections). An extension to the current model would be required in order to account for excitatory (+) and inhibitory (-) S-R associations in the habit system.

List of Tables

Table 1. Design of a typical PIT Experiment

Table 2. Design of Experiments 1-3

Table 3. Pavlovian and instrumental conditioning during Experiment 3

Table 4. Design of Experiments 4-6

Table 5. Pavlovian and instrumental conditioning during Experiments 4-6

Table 6. Design of Experiment 7

Table 7. Design of Experiments 8 and 9

Table 8. Design of proposed Experiment (i; and idealised results)

Table 9. Design of proposed Experiment (ii; and idealised results)

Table 10. Design of proposed Experiment (iii)

Chapter 1

General Introduction

1.1. Background

This thesis investigates the interactions between Pavlovian and instrumental processes. These two fundamental learning processes permeate all aspects of human and animal behaviour. A key attribute of people and animals alike is the capacity to select appropriate responses in the face of changing conditions. This may be underpinned by the Pavlovian properties of cues that signal the presence or absence of certain desirable or undesirable outcomes, and thus guide behaviour (Doya, 2008). Understanding the interactions between these processes has applied relevance to forms of human psychopathology that involve conditioning - most notably drug dependence, anxiety disorders, and over-eating. Taking the example of drug addiction, cues associated with drug use can induce relapse after prolonged periods of abstinence, even when the drug is no longer desired (Robinson & Berridge, 2008). Furthermore, drugs themselves can act as cues, where one drug (e.g., alcohol) has been found to induce relapse to seeking other drugs (e.g., nicotine; Murray, Palmatier & Bevins, 2007). In associative learning terms, Pavlovian associations with drug cues are affecting an instrumental behaviour (i.e., drug seeking).

In modern society palatable foods with high calorific value are readily available, as are the cues (e.g., advertisements) that predict them. In many situations, Pavlovian cues seem to provoke instrumental behaviours, such as seeking food when hungry and water when thirsty (Perks & Clifton, 1997), that are adaptive. However, in some situations, the behaviours prompted by such cues can be mal-adaptive. For instance, cues associated with foods have been found to prompt eating, even in the absence of hunger. In one study, rooms or contexts associated with snack food were found to motivate children to eat more, even when the children had been pre-fed ice-cream (Birch, 1991). Indeed, overweight children may be particularly vulnerable to such cues that prompt over-eating (Jansen, Theunissen, Slechten, Nederkoorn, Boon, Mulkens, & Roefs, 2003). The idea that Pavlovian cues might prompt overeating, in the absence of hunger, might contribute to the rise in obesity associated with over-eating (Boggiano, Dorsey, Thomas, & Murdaugh, 2009).

A greater understanding of how Pavlovian cues affect instrumental behaviour may allow the development of better remedial treatments for maladaptive behaviours, or means of preventing their occurrence. This introduction considers, in detail, two examples of the interaction between Pavlovian and instrumental processes, and evaluates the theoretical analyses that have been developed to explain them. Before considering their interaction, it is necessary to briefly outline how they operate independently.

1.2. Learning Processes

1.2.1. Pavlovian Conditioning

In a typical Pavlovian conditioning experiment, neutral stimuli (such as lights or tones; called conditioned stimuli or CSs) are presented in some temporal relationship to a motivationally significant stimulus (e.g., an appetitive stimulus, such as food; or aversive stimulus, such as a shock; called unconditioned stimuli or USs). Repeated pairings of the CS and US, usually where the CS precedes the US, lead to the CS provoking some conditioned response or responses (CR) that often resembles a component or components of the unconditioned response (UR) to the US. For example, if a tone reliably precedes the delivery of food, then a rat will come to respond to the tone in a way that resembles how it responds to food (i.e., approaching the site of food delivery). Nowadays, this change in behaviour is most often attributed to the development of an association between the representations of the CS and US (e.g., Dickinson, 1980).

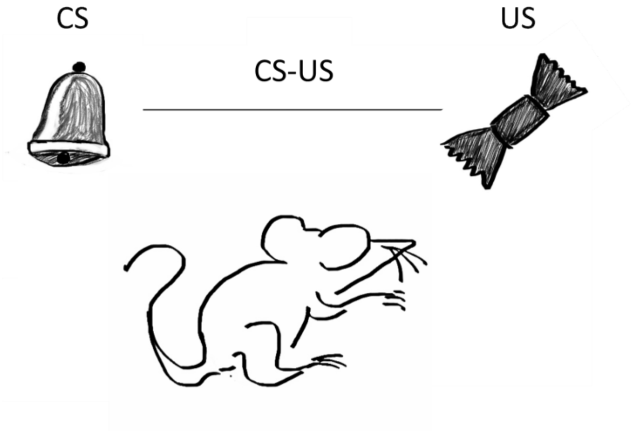


Figure 1. A typical Pavlovian conditioning experiment. Repeated pairings of a neutral stimulus such as a tone (a conditioned stimulus or CS) with food (an unconditioned stimulus or US) leads to some behavioural change that is indicative of the formation of an association between the CS and the US. As a result the animal comes to respond to the CS in a way that often resembles how it would respond to the US.

Pavlovian conditioning procedures involve different kinds of predictive relationships. These may be excitatory, in which the CS predicts the occurrence of a US, or inhibitory, in which the CS predicts the absence of a US. In addition, CS-US pairings can result in different kinds of CR: preparatory responses (such as approach behaviour or changes in heart rate) which involve the motivational status of the US; and consummatory responses (such as blinking for a CS predicting an air puff) which are specific to the identity of the predicted US. Konorski (1967) emphasised the theoretical implications of the distinction between preparatory and consummatory responding. He proposed a model (reviewed in Dickinson & Balleine, 2002) which identified two types of independent associations resulting from Pavlovian conditioning. One association was between a representation of the CS and the sensory properties of the US, resulting in US-specific consummatory responses. The second separate association is between the CS representation and the generic motivational properties of the US, which may be appetitive or aversive. This distinction is important when considering how Pavlovian processes might interact with instrumental processes (see Section 1.3).

1.2.2.Instrumental Conditioning

In Pavlovian conditioning procedures, the presentation of the reinforcer is independent of the behaviour of the animal, and is instead contingent on the occurrence of a stimulus. In contrast, for instrumental learning, reinforcement is dependent of the response of the animal. In an instrumental conditioning scenario, a specific response (R), such as pressing a lever can lead to a desirable outcome (O), such as food, in the presence of specific stimuli (S) for instance, an experimental chamber or context with a lever in it. Several associative structures have been offered to explain instrumental responding. The stimulus-response account suggests that the animal has learned to associate stimuli, such as the response manipulanda or other environmental cues (S), with that specific response (R; see Thorndike, 1911; Skinner, 1935). Here, the role of the outcome is to reinforce such S-R associations. An alternative model suggests the animal learns a response-outcome association (see Tolman, 1948; Dickinson, 1997). The critical difference between these models is the role of the outcome. In the S-R model, the key role for the reinforcer is during the acquisition of the association, though once established, the response is independent of outcome value. According to the R-O model, the outcome is directly represented in the association that directs responding, thus the response should be sensitive to changes in outcome value. More recent research has suggested that both associative structures are acquired, with instrumental behaviour being subdivided into goal-directed (R-O) and habitual (S-R) behaviour based on their sensitivity to the current value of the outcome (Dickinson & Balleine, 2002). This subdivision also suggests that responding based upon the two associations will be differentially affected by the presence or absence of the environmental cues (the S). Another clear distinction between these two structures is that while the R-O association allows instrumental performance to be affected by the presence of CS associated with the same O, performance that is mediated by S-R associations will only be affected by Pavlovian cues to the extent that there is response competition generated by the CS. I now move on to consider how the presence of a CS affects instrumental performance.

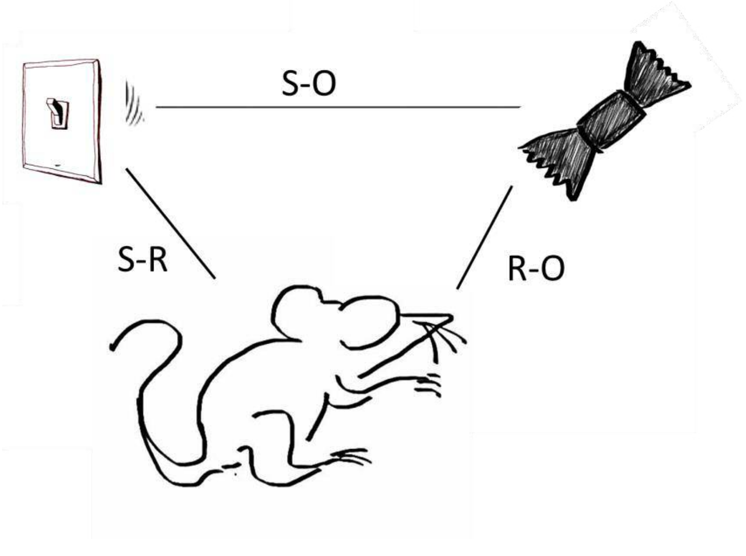


Figure 2. Associative structures for instrumental responding. The rat performs a response (R) that is reinforced by food (O) leading to a goal-directed R-O association; or the response manipulandum becomes a cue (S) associated with the response (R).

1.3. Interactions between learning processes: Pavlovian-to-instrumental Transfer

Pavlovian-to-instrumental transfer (PIT) refers to the fact that a Pavlovian CS, that predicts a given reinforcer, can elicit or increase instrumental responses associated with the same or similar reinforcers (e.g., Estes, 1948; Ostlund & Balleine, 2007). PIT procedures typically involve three stages (see Table 1): Pavlovian conditioning (e.g. pairings of a light and a food pellet), instrumental conditioning (e.g., pairings of a lever press and a food pellet), and an extinction test in which the influence of the CS is assessed on the instrumental response. PIT occurs in two forms: outcome-selective and general. Outcome-selective PIT refers to the ability of a stimulus to increase the likelihood of an instrumental response associated with the same outcome, whereas general PIT reflects the ability of a CS to elicit instrumental responses associated with reinforcers from the same affective class (for a recent review, Holmes, Marchand, & Coutureau, 2010). Demonstrations of PIT reveal that Pavlovian and instrumental learning processes are based upon shared systems (i.e., behavioural and neural), and have been used to investigate the behavioural and neural mechanisms that underlie response selection (Blundell, Hall, & Killcross, 2001; Corbit, Muir, & Balleine, 2001; Dickinson, Smith, & Mirenowicz, 2000; Hall, Parkinson, Connor, Dickinson, & Everitt, 2001; Holland & Gallagher, 2003; Johnson, Bannerman, Rawlins, Sprengel & Good, 2007), and as a model of components of drug addiction (e.g., Everitt & Robbins, 2005). In the following sections, I evaluate contemporary accounts of PIT and argue that these accounts fail to provide a coherent explanation for the elusive nature of PIT (cf. Holmes et al., 2010) or for some of the conditions under which it is observed. The elusive nature of PIT provides the rationale for the experimental work described in Chapter 2, where I consider the influence of outcome discriminability. The conditions under which PIT is observed prompted the development of an alternative, stimulus-response analysis, the validity of which is investigated in Chapter 3.

Table 1

Design of typical PIT experiment

Instrumental	Pavlovian	Test
R1-O1	S1+O1	S1 (Same) or S2 (Different): R1
R2-O2	S2+O2	S2 (Same) or S1 (Different): R2

Note: R1 and R2 refer to left and right lever presses; O1 and O2 denote food pellets and sucrose; and S1 and S2 are a tone and a light. The influence of S1 and S2 on instrumental responding during the test is assessed. Same and Different indicate whether the designated response (R1 or R2) had been paired with the Same or Different outcome as the stimulus (S1 or S2).

1.4. Theoretical analyses of PIT

There are several theoretical analyses for interactions between Pavlovian and instrumental processes. However, none of these analyses can account fully for the conditions under which PIT is observed. These analyses are now explored, in turn, and an alternative account is proposed that is capable of explaining some otherwise paradoxical findings within the literature.

1.4.1. S-O-R: A two-process theory

It has been argued that Pavlovian-instrumental interactions can be mediated by associations of the CS with sensory-specific as well as motivational components of the US (for review, see Dickinson & Balleine 2002; see also, Rescorla & Solomon, 1967). Balleine and Ostlund (2007) proposed a type of two-process account that was based on an S-O-R associative chain. This account was intended to provide a general analysis for many of the findings related to interactions between Pavolvian and instrumental processes, including outcome-selective and general PIT. They suggested that the R-O associations formed during instrumental training are bi-directional, allowing the R to activate the O and vice versa. Test presentations of the CS will activate a representation of the outcome, via the S-O association, which will generate a response associated with the same outcome, via the O-R association. This model accommodates outcome selectivity of PIT through the integration of S-O and O-R associations, where the O represents the sensory specific aspects of the reinforcer; while general PIT reflects those aspects of the O that are shared across a motivational class.

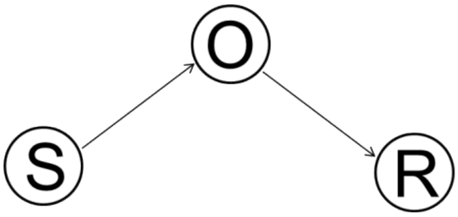


Figure 3. An S-O-R associative chain. At test, the CS (S) evokes a representation of the outcome (O) that, in turn, activates the response (R), via the O-R association.

There is, however, evidence that challenges this S-O-R analysis. First, it is known that instrumental performance, which is assumed to rely on R-O associations, is sensitive to the current value of the O (e.g., Adams, 1982; Adams & Dickinson, 1981). One might suppose then that (outcome-selective or general) PIT will also only be observed to the extent that the outcome is valuable during the test. However, PIT is unaffected by devaluation of the outcome immediately prior to the critical test (e.g. Holland, 2004). In this study, an outcome was paired with toxin-induced illness following the Pavlovian and instrumental training phases of a PIT procedure. A subsequent PIT test showed no attenuation of lever pressing, suggesting that devaluing the outcome did not influence the likelihood of observing the PIT effect. This observation is difficult to reconcile with accounts of PIT that rely on the excitement of a representation of a shared outcome, because devaluation should exert an outcome-selective depression of instrumental (R-O mediated) performance. Although it is possible that the outcomes were not completely devalued, consumption tests have shown that their value is substantially reduced with no concomitant effect on PIT.

Rescorla (1994) has argued that the CS in PIT experiments exert their effects through a signalling as opposed to a motivational role; and that transfer is mediated by the CS activating a representation of the outcome, the sensory aspects of which provide additional discriminative cues that evoke the instrumental response. As this analysis relies on the identity of O as opposed to its current value, there is no reason for devaluation to influence PIT: The sensory representation of O remains able to provide an additional discriminative cue for instrumental performance irrespective of the devaluation treatment. However, one could argue that this analysis is based upon special pleading: once the response is activated then its vigour should be determined (at least in part) by the current value of the outcome with which it is associated (see Balleine & Ostlund, 2007).

A second problem with S-O-R theory is that manipulations that should undermine the efficacy of S-O associations (extinction of the S; Delamater, 1996) should disrupt PIT. However, such extinction treatments have no effect on outcome-selective and general PIT. One could argue that these treatments did not affect the strength of the critical associations (for review, see Bouton, 2004), but they certainly should have affected their efficacy. Another problem with conventional analyses, is that outcome-selective PIT seems to be more reliable during a long CS, which immediately facilitates lever pressing, than a short CS, which has been found to have bi-phasic effects on lever pressing (Holland & Gallagher, 2003); and PIT is often observed when the outcomes are delivered within a long CS (e.g., Holland & Gallagher, 2003; Holland, 2004; Corbit et al., 2001; Holmes et al., 2010). These observations are curious given the fact that these conditions are unlikely to be conducive to the formation of especially strong S-O associations – associations that should, according to the S-O-R account, determine the strength of PIT. Perhaps, however, a more elaborate analysis, based upon the same general principles, might be better placed to deal with these anomalies. The associative-cybernetic model is one such analysis.

1.4.2. Associative-Cybernetic model

The Associative-Cybernetic model was advanced by Dickinson and Balleine (1993) as a general model of instrumental performance. The model is based on the general idea that there are parallel R-O and S-R systems. Within the R-O system, responses are linked to representations of outcomes, and performance can be generated through two routes: (i) a link between the O and the motor program via a reward memory, and (ii) a link between the R and the same motor program via the R within the S-R system (see Figure 4). Within the S-R system, stimuli (e.g., the sight of the lever, S) are linked to this representation of the response (i.e., R), and then to the motor program. To make this analysis more concrete, take the example of a rat that comes to perform a lever press and is reinforced by a food pellet for doing so. According to this model, the sight of the lever triggers a representation of the response in the S-R system, which in turns triggers a representation of the response in the R-O system that then activates the outcome. The rat then evaluates the current value of that outcome (O), and if it is positive (i.e., sufficient to activate the motor program of the associated response) the rat performs the lever press (R). Therefore, according to this model, both the selection of a response and the evaluation processes combine in order to initiate actions. When the performance of a response results in an outcome, the outcome representation is triggered in the associative memory and the reward memory, which updates the strength of the R-O association, and outcome value in the incentive system. The delivery of an outcome also reinforces the contiguously active S and R representations in S-R memory.

The influence of a Pavlovian CS on instrumental responding can be mediated by the capacity of such a CS to act on the outcome memory that is also linked to the instrumental response. Under these conditions, the pathway between the O and the motor program would be more active when the CS is present than when it is absent. However, this simple explanation of PIT would leave the effect sensitive to the value of the outcome, and, as I have mentioned, PIT is resistant to outcome devaluation treatments (Holland, 2004). A possibility that avoids this prediction is based upon the idea that the presentation of the CS triggers the response representation (via a reciprocal O-R association in the associative memory) which then activated the response representation held in S-R memory. Under these conditions, the influence of the Pavlovian CS on performance would be mediated by the S-R system and would be independent of the current value of the outcome (cf. Balleine & O’Doherty, 2009). This analysis, however, requires that the CS-outcome association has no effect on instrumental performance via the reward system, but that response-outcome associations do affect performance via this system. Moreover, the AC model fails to account for the fact that manipulations that reduce the efficacy of the

S-O association (e.g., Delamater, 1996) do not impact on the ability of a Pavlovian CS to increase instrumental responding. The model also does not provide a clear account for the observation that other variables that should reduce the strength of S-O associations during training, such as the use of an extended CS (Holland & Gallagher, 2003) or delivering the outcome during the CS (Holland & Gallagher, 2003; Holland, 2004; Cobit et al., 2001; Holmes et al., 2010), increase the likelihood of observing PIT. Rather, these manipulations might be expected to reduce the strength of the critical CS-O association and reduce the ability of the CS to activate the O and thereby affect instrumental performance.

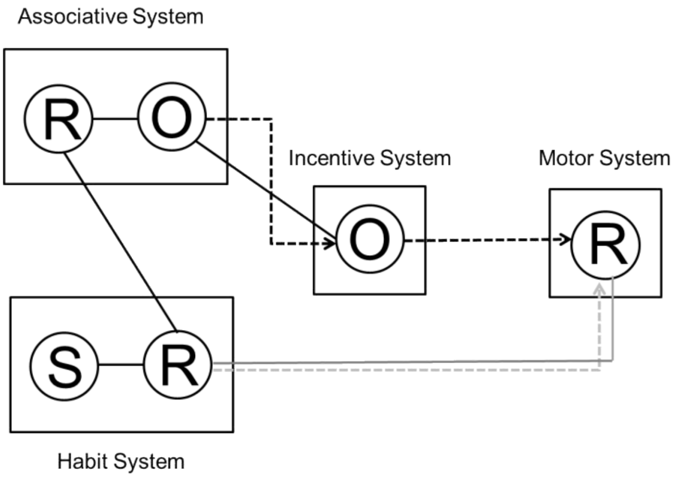


Figure 4. An Associative-Cybernetic model of instrumental responding (black lines represent learned associations; grey lines represent fixed connections). An instrumental response (R) can be generated (i) through a link between the O in the associative system and the corresponding motor program via a reward memory (black dotted line); (ii) via a direct link between the response and the corresponding motor program via the R in the S-R (habit) system (grey dotted line). In the case of PIT, a Pavlovian CS can activate a response through a shared outcome memory (within the associative system) which is linked to the motor program through the incentive system. Alternatively, the Pavlovian CS may trigger the response via a bi-directional O-R link, which activates the R in the habit system, and acts on the corresponding motor program independently of the incentive system.

1.4.3. Hierarchical Theory

Another analysis of the control of instrumental performance by accompanying stimuli relies on the idea that R-O associations might be hierarchically gated. Skinner (1938) suggested that stimuli might affect performance through an occasion-setting function. In associative terms, this view has been represented by the stimulus activating the association between the response and outcome: an S-(R-O) associative structure (see Figure 5). According to this analysis, instrumental performance is not a simple product of binary S-O, R-O or S-R associations, but reflects the hierarchical control by the S of the R-O association (Colwill & Rescorla, 1988; Rescorla, 1990a). There is some evidence to support this view. For example, Rescorla (1990a) designed an experiment in which, during pre-training, a light (L) signalled two particular response-outcome relations (R1-O1 and R2-O2). In the next phase of training, the same light was presented in compound with a tone (T), which also signalled R1-O1 and R2-O2. As L had previously signalled the same events and relations between them, the T would be blocked and therefore uninformative. A control stimulus (a noise; N) was also presented in compound with the light (L), which was informative about a new relation between these events; R1-O2 and R2-O1. The same individual events occurred in the presence of both N and T, and individual events were equally predicted by the light. However, for one auditory stimulus the relationships between events (the control stimulus - N) were not predicted by the light, and for the other, blocked stimulus they were predicted by the light. To assess the control of the two auditory stimuli, N and T, they were presented separately and outside of the compound with the light. Rescorla found that the control stimulus was more likely to provoke R1 and R2 at test than the blocked stimulus. This suggests the informativeness of stimuli with respect to instrumental *relationships*, between the response and the outcome, are important for the development of stimulus control; an observation that cannot be readily explained in terms of simple binary associations.

This analysis could be developed to explain PIT in the following way. First suppose that during Pavlovian conditioning reciprocal associations form between the components of the trials (i.e., S1 and O1; S2 and O2), and that such associations also form during instrumental conditioning (i.e., R1 and O1; R2 and O2). These associations will allow, for example, S1 to be evoked during the trace of R1 and O1, and S2 to be evoked during the trace of R2 and O2. Under these conditions, S1 might gain hierarchical control of the R1-O1 association and S2 to control the R2-O2 association. Now, S1 and S2 will be able to generate PIT through their ability to act directly on the relevant, specific associations. One virtue of this analysis of PIT is that hierarchical stimuli do not lose their properties when they are presented in isolation (i.e., they are not subject to extinction; see Rescorla, 1992; see also, Holland, 1989). Thus, one would not anticipate that presenting S alone should undermine PIT (Delamater, 1996). A second virtue is that it provides a ready analysis of why embedding the outcome within the Pavlovian stimulus is an effective way of generating PIT: this will allow the R to be activated when both the S and O are present. However, taken in isolation, this analysis provides no obvious grounds for anticipating that outcome devaluation will not affect PIT. During the test, the S will actuate the R-O association, and performance should be constrained by the current value of O.

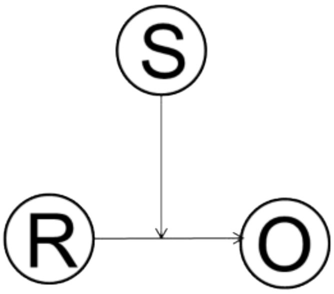


Figure 5. The hierarchical model of instrumental conditioning according to which stimuli (Ss) act on the association between responses (Rs) and outcomes (Os).

The analysis outlined in the previous paragraph makes use of the general idea that stimuli can acquire properties (hierarchical in the case under consideration above) when they are associatively provoked rather than being directly activated by the corresponding event in the world. This idea is not novel and there is evidence to support it from studies of simple conditioning. For example, Holland (1981) gave rats pairings of a stimulus with food, and then paired the stimulus with lithium chloride. This procedure resulted in a food aversion. One interpretation of this finding is that the stimulus evoked a representation of the outcome in the presence of the sickness induced by lithium chloride, and this retrieved memory of food entered into an association with nausea. The process of forming such mediated associations is not limited to instances of simple conditioning: mediated configural associations have been observed in rats using variants of a sensory preconditioning procedure (see Allman & Honey, 2006; Lin & Honey, 2010). However, once one allows that such a process of mediated learning occurs, then a simpler S-R analysis of PIT suggests itself. This analysis is better placed to explain many features of PIT, but has not, as far as I am aware, been suggested before.

1.4.4. S-R Theory

It has been argued previously that PIT is unlikely to be mediated by simple S-R associations (Blundell, et al., 2001; Colwill & Rescorla, 1988; Rescorla, 1990b). This argument is based on the assumption that the training conditions used in PIT procedures should not be capable of generating the S-R associations that would be necessary to support PIT. Consider the standard procedure in which instrumental training sessions (R1-O1 and R2-O2) occur separately from Pavlovian training sessions (S1-O1 and S2-O2). It is assumed that because S1 is not contiguous with R1 (they are trained in separate sessions) and S2 is not contiguous with R2, S1 and S2 should be no more likely to provoke R1 than R2. However, this analysis of the role of S-R associations does not consider the possibility of retrieval-mediated S-R learning (see Figure 6). In general, if rats receive instrumental training prior to Pavlovian training, there is the potential for O1 to activate R1 during S1, and for O2 to activate R2 during S2. These conditions might be sufficient to generate S1-R1 and S2-R2 associations that could directly produce outcome-selective PIT. Thus, take the example of a typical outcome-selective PIT procedure where two responses are trained with outcomes in separate sessions, this will result in bi-directional associations between R1 and O1 (i.e. R1-O1 and O1-R1) and between R2 and O2 (i.e., R2-O2 and O2-R2). Now, a phase of Pavlovian conditioning follows where S1-O1 and S2-O2 associations are formed. Given the standard procedure, where outcomes are embedded within CS presentations (cf. Holland & Gallagher, 2003; Holland, 2004; Corbit et al., 2001; Holmes et al., 2010), S1 and S2 will be present when O1 and O2 activate the motor programs of their corresponding responses (i.e. R1 and R2). Under these conditions, it is possible for S1 and S2 to become associated with R1 and R2, respectively, as a result of their temporal contiguity (see

Honey, Good & Manser, 1998). The resulting S1-R1 and S2-R2 associations would provide the basis for outcome-selective PIT at test without any mediation by the outcome. In the case where the order of the two stages of training are reversed (i.e. Pavlovian conditioning precedes the instrumental phase), if it is assumed that O1 and O2 come to evoke representations of S1 and S2 then these representations will be activated during later instrumental training where R1 and R2 are paired with O1 and O2. This should allow S1-R1 and S2-R2 to form. Whichever the order of training, the process of mediated stimulus-response learning (see Holland, 1983; Iordanova, Good, & Honey, 2011; Wheeler, Sherwood, & Holland, 2008; but see, Wagner, 1981) will depend on the stimuli or evoked representations, and responses (or corresponding motor programs) occurring in close temporal contiguity. Importantly, mediated S-R associations would be unaffected by the value of the outcome, meaning this analysis would be consistent with the finding that outcome devaluation immediately prior to test has no effect on responding (e.g., Holland, 2004). Additionally, presentations of stimuli prior to test, that should undermine the efficacy of the S-O association (see Delamater, 1996), might be expected to have rather less impact on S-R associations because there is no equivalent violation of the information contained in the S-R association when the outcome is not delivered. The possibility that this S-R mechanism could provide an explanation for the observations made during Pavlovian-to-instrumental transfer tests are discussed in depth in Chapter 3.

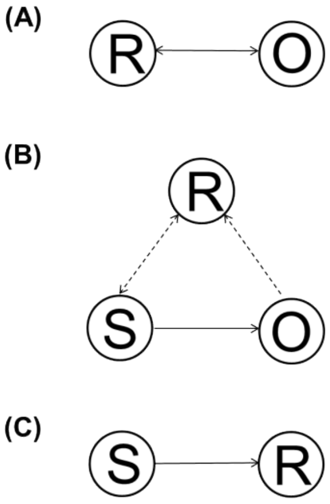


Figure. 6. An S-R account of PIT. (A). A bi-directional R-O association is established in the first phase of training. (B). In the second phase of training, the presentation of a common outcome (O) evokes a representation of the response (R) via the backwards O-R association in the presence of the stimulus (S). This leads to an excitatory association between S and R. (C). At test, presentations of S evoke a response through the activation of this S-R association.

The theoretical analyses of PIT have been evaluated with respect to a limited number of the characteristic features of PIT: the fact that it is immune to changes in the value of the reinforcer prior to test, and to manipulations that should influence the strength of simple associations between the Pavlovian CS and the outcomes. These features were chosen because they are theoretically interesting. However, it should be acknowledged PIT is rather more elusive than the discussion, up until this point, has implied (see Holmes et al., 2010). The boundary conditions under which PIT is observed have not been subject to a systematic investigation. In order to evaluate the accounts that have been developed in Chapter 1, it is first necessary to generate an outcome-selective PIT effect. The experiments presented in Chapter 2 were an attempt to provide such a demonstration. These experiments also investigated whether manipulations that should increase the distinctiveness of the outcome representations (prior to Pavlovian and instrumental training) increase the likelihood of observing PIT. In particular, these experiments examined whether pre-exposure to the outcomes prior to conditioning influenced PIT. Each of the accounts considered predict that such manipulations should have an impact on PIT. However, as will become apparent, the results of these experiments did not provide evidence that pre-exposure to the outcomes affected outcome-selective PIT. Nevertheless, the results of Chapter 2 did prompt the choice of the procedures that were employed in Chapter 3 to demonstrate outcome-selective PIT and to examine its origins.

1.5. Contextual influences on instrumental behaviour

Earlier, I outlined how cues associated with drug use can induce relapse, sometimes following prolonged periods of abstinence (Robinson & Berridge, 2008). PIT has been offered as a possible explanation for relapse behaviours where Pavlovian cues such as drug paraphernalia may serve to evoke drug seeking (c.f., Everitt & Robbins, 2005). The interaction between Pavlovian and instrumental conditioning that is evident as PIT is clearly in need of further experimental analysis. It might be argued, however, that the PIT procedure is relatively complex, and might allow a variety of different mechanisms to operate in parallel. The effect might be multiply determined. There are other, simpler, examples where Pavlovian stimuli modulate instrumental performance (Pearce & Hall, 1979; see also, Baker, Steinwald, & Bouton, 1991; Colwill & Rescorla, 1988; Rescorla & Solomon, 1967), which may prove to be more analytically tractable. A well-documented example of such an interaction, which has also been implicated in relapse behaviours, is the ABA renewal effect. Here, a conditioned response is first established to a stimulus in one context (A) and then extinguished in another context (B) prior to a final test in context A. This change in the context, after extinction, results in the conditioned response being made with renewed vigour (for a review, see Bouton, 2004). For example, the renewal of extinguished drug seeking behaviour (i.e., lever pressing) has been observed in rats when the environmental context was changed after extinction (Hamlin, Clemens, & McNally, 2008) or alcohol (Hamlin, Newby, & McNally, 2007).

Whilst the majority of examples of such renewal effects involve contextual modulation of Pavlovian CRs to CSs, there are examples of renewal using instrumental procedures and conventional reinforcers (e.g., Bouton, Todd, Vurbic, & Winterbauer, 2011; Nakajima, Tanaka, Urushihara, & Imada, 2000; Nakajima, Urushihara, & Masaki, 2002; Todd, Winterbauer, & Bouton, 2012). Although the origin of Pavlovian renewal effects has been examined in some detail (e.g., Bouton & King, 1983; Bouton & Swartzentruber, 1986; Rescorla, 2008), the origin of the control of instrumental performance by contexts is not yet clear.

One way in which contexts might come to control instrumental responding is through the Pavlovian properties that they acquire during training when the response is reinforced (creating an excitatory association with the outcome), and also during periods of nonreinforcement during extinction (creating an inhibitory association). The possibility that instrumental renewal effects are controlled by excitatory Pavlovian properties of the training context, although intuitive, is undermined by the finding that renewal is observed when testing occurs in a novel context (e.g. ABC renewal) even when extinction occurs in the training context itself (AAB renewal) (see Bouton et al., 2011). Another possibility is that renewal could be controlled by inhibitory associations formed during extinction (Bouton & King, 1983; Bouton & Swartzentruber, 1986). Assuming extinction causes an inhibitory association between the extinction context and outcome that masks responding, removal from this context would remove the influence of these inhibitory associations and hence restore responding. This analysis could account for both AAB and ABC renewal.

Chapter 4 explores theoretical explanations of instrumental renewal based on the inhibitory or excitatory Pavlovian associations between the contexts and the outcome. In this case, as in outcome-selective PIT, the contextual cues are influencing activity in the outcome, that in turn affects the instrumental response. However, a further possibility is that the context could act as an occasion-setting mechanism, on the links between the response and the outcome, or the absence of the outcome (see Bouton & Swartzentruber, 1986; Honey & Watt, 1999). Applied to the instrumental case this would involve the conditioning context (A) gating the excitatory response-outcome link or the extinction context (e.g., B) or gating the response-no outcome link, in AAB, ABA or ABC renewal (Bouton et al., 2011). Finally, instrumental renewal, like outcome-selective PIT, might result from excitatory or inhibitory context-response associations. Chapter 4 investigates these potential mechanisms for instrumental renewal and the control of instrumental performance by contexts more generally.

1.6. Summary of rationale for new empirical work

The overarching aim of this thesis was to increase our understanding of the mechanisms underlying interactions between Pavlovian and instrumental processes. To this end, Chapters 2 and 3 investigated the mechanisms that underlie PIT. Chapter 2 assessed the conditions under which outcome-selective and general PIT are observed; while Chapter 3 investigated a novel analysis of outcome-selective PIT. This analysis was based on mediated S-R learning, and the predictions derived from this analysis were contrasted with those derived from the conventional S-O-R analysis of PIT. Finally, Chapter 4 investigated the influence of contextual cues on instrumental performance in instrumental renewal procedures. This chapter assessed the origin of renewed instrumental responses, focussing on two theoretical issues: the extent to which renewed responding is based on Pavlovian associations with the outcome; and whether or not renewed responses are controlled by goal-directed processes.

Chapter 2

The elusive nature of outcome-selective Pavlovian-to-instrumental transfer

2.1. Summary

PIT has been widely cited and used to demonstrate important interactions between Pavlovian and instrumental learning processes. However, it could be argued that this phenomenon is somewhat elusive, given the reported difficulties in obtaining consistent results (e.g., Delamater & Holland, 2008). In Chapter 2, three experiments examined the conditions under which general and outcome-selective PIT occur in rats. In Experiment 1, general PIT, but no outcome-selective PIT, was observed when the Pavlovian and instrumental conditioning sessions with the two CSs and two responses were both arranged in blocks (e.g., R1→O1, R1→O1….R2→O2, R2→O2…. S1→O1, S1→O1….S2→S2, S2→S2). Experiments 2 and 3 explored the influence of procedures that should affect the discriminability of the two outcomes (O1 and O2) on Pavlovian-instrumental transfer, using differing schedules of pre-exposure (cf. Blair, Blundell, Galtress, Hall, & Killcross, 2003). In Experiments 2 and 3, rats received either intermixed pre-exposure (O1, O2, O1, O2….), blocked pre-exposure (O1, O1…O2, O2) or no pre-exposure, prior to either short (Experiment 2) or long (Experiment 3) stages of Pavlovian and instrumental conditioning. Pavlovian and instrumental conditioning were unaffected by pre-exposure, and evidence of a general PIT effect was restricted to Experiment 3. These results suggest that outcome-selective PIT is not encouraged by procedures that should increase the distinctiveness of the outcomes.

2.2. Introduction

Under certain conditions, Pavlovian and instrumental processes appear to interact. An example of such an interaction is Pavlovian-to-instrumental transfer (PIT) where a CS increases an instrumental response when the stimuli and the response have been paired with the same (or a similar) outcome (for a recent review, see Holmes et al., 2010). Although PIT tests have been widely used to demonstrate the interaction between these two fundamental learning processes, there is evidence to suggest that this phenomenon is elusive. Inconsistencies in results have been described which have been suggested to be attributable to relatively trivial differences in experimental chamber configurations (cf. Delamater & Holland, 2008). Determining the basis for these inconsistencies is made difficult by many differences between different PIT procedures, such as order and duration of conditioning, which have been found to influence the effect of a CS on instrumental responding, together with a range of other differences (see Holmes et al., 2010).

One way in which the procedural differences in studies of PIT might influence transfer effects is as a result of differences in how representations of the outcomes are encoded. The sensory properties of outcomes have been found to influence Pavlovian performance (e.g., Watt & Honey, 1997), and it has been further suggested the different sensory and affective properties of outcome representations provide the mechanism for general and outcome-selective PIT. For example, Dickinson and Balleine (2002) suggested that PIT can be mediated by the sensory-specific as well as motivational aspects of outcome representations. They argued that a Pavlovian CS gains access to motivational functions of the reinforcer, both directly (in the case of general PIT), and indirectly through associations with sensory features of that representation (in the case of outcome-selective PIT). If this analysis has merit, then one would expect circumstances where motivational aspects of the reinforcer are encoded more strongly than the sensory aspects to result in general PIT, and reduce outcome-selective PIT.

In keeping with the analysis outlined in the preceding paragraph, outcome-selective PIT has been observed using procedures that intermix Pavlovian and instrumental training on alternating days (Blundell et al., 2001), and also when separate Pavlovian and instrumental phases intermix S1-O1/S2-O2 or R1-O1/R2-O2 pairs on alternating days (e.g. Corbit et al., 2001; Holland, 2004; Delamater & Holland, 2008). Such procedures are likely to increase the discriminability of the outcome representations (Blair et al., 2003; see also, for example, Honey, Bateson & Horn, 1994; Symonds & Hall, 1995). For instance, Blair and colleagues assessed whether intermixed or blocked pre-exposure to two outcomes (O1 and O2) prior to pairing them with two responses (R1 and R2) would affect their discriminability. Following instrumental training, rats were sated on O1 only. This devaluation procedure reduced the likelihood of rats performing R1 relative to R2 in those given intermixed exposure to O1 and O2, but not those given blocked exposure. One explanation for this finding is that intermixed exposure to outcomes increased the discriminability of the outcomes and thereby allowed exposure to O1 to have a selective effect on R1 following devaluation. It is interesting to note, that there are no instances in which outcome-selective PIT has been observed when the procedure has involved blocked exposure to the outcomes within the instrumental and Pavlovian conditioning procedures.

In addition to the procedural conditions described above that appear to affect PIT, outcome-selective PIT has been found using widely differing amounts of training, ranging from as little as four Pavlovian trials (Delamater & Oakeshott,

2007) to 384 trials (Rescorla, 1994), and four instrumental sessions (Zorawski & Killcross, 2003) to 20 sessions (Delamater & Holland, 2008); with similar ranges being used in general PIT paradigms (e.g. Rescorla, 2000; Holland, 2004). A meta-analysis of PIT studies suggested variations in the amount of training affected transfer in a manner that depended on the order of training (Holmes et al., 2010). Specifically, the amount of Pavlovian training in the second phase increased the strength of outcome-selective PIT, and had no impact on general PIT. However, the amount of instrumental training influenced both outcome-selective and general PIT. Again, it is tempting to attribute these procedural effects to variations in the encoding of the outcomes. For example, it is possible that outcome features are differentially encoded when Pavlovian conditioning is conducted before or after instrumental training. When Pavlovian conditioning occurs in the absence of prior experience of the outcome, the CS may initially encode motivational features of the outcome, and encoding of sensory aspects only occurs with further training (cf. Konorski, 1967).

The overarching aim of the first series of experiments was to directly investigate whether manipulations that should affect outcome encoding influenced the likelihood of observing general and outcome-selective PIT. Experiment 1 assessed whether arranging the PIT procedure in a way that limited the extent to which the outcomes would become discriminable resulted in general rather than outcome-selective PIT. Experiments 2 and 3 examined the prediction that procedures that should encourage differences in outcome discriminability should modulate whether general or outcome-selective PIT is observed. To this end, rats received pre-exposure to outcomes in a blocked or intermixed fashion followed by a standard outcome-selective PIT procedure in which outcome pairings in each phase were intermixed either in short (Experiment 2) or long (Experiment 3) conditioning phases. A control group received no pre-exposure to the outcomes prior to the PIT procedure.

2.3. Experiment 1

Experiment 1 used a within-subjects PIT procedure in which the outcomes (O1 and O2) were embedded within the stimuli. This design is depicted in Table 2 and involved two stages of training, Pavlovian and instrumental conditioning, that were presented in blocks of consecutive sessions. The numbers of conditioning sessions were 16 and 12 respectively, which is similar to procedures that have successfully generated outcome-selective PIT (for review, see Holmes et al., 2010). Thus, rats received all S1-O1 sessions followed by all S2-O2 sessions for Pavlovian conditioning and all R1-O1 sessions followed by R2-O2 for instrumental conditioning. Following this training, rats received two extinction test sessions in which one lever was presented and responding was assessed in the presence of S1 and S2. An outcome-selective PIT effect would be evident if R1 was more frequent during S1 than S2, whereas if there was more R1 and R2 during S1 and S2 than during the inter-trial interval (ITI) it would represent an instance of general PIT.

Table 2

Design of Experiments 1-3 (and designation of instrumental responses)

Pre-exposure	Instrumental	Pavlovian	Test
Experiment 1			
	R1-O1	S1+O1	S1 (Same) or S2 (Different): R1
	R2-O2	S2+O2	S2 (Same) or S1 (Different): R2
Experiments 2 and 3			
Group Intermixed:			
(O1) (O2)… (O1) (O2)			
Group Blocked:	R1-O1	S1+O1	S1 (Same) or S2 (Different): R1
(O1) (O1)… (O2) (O2)			
Group None:	R2-O2	S2+O2	S2 (Same) or S1 (Different): R2
No pre-exposure			

Note: R1 and R2 refer to left and right lever presses; O1 and O2 denote food pellets and sucrose; and S1 and S2 are a tone and a light. S1+O1 and S2+O2 indicate that the outcomes were presented throughout S1 and S2. The influence of S1 and S2 on instrumental responding during the test was conducted with a single lever present. Same and Different indicate whether the designated response (R1 or R2) had been paired with the Same or Different outcome as the stimulus (S1 or S2).

2.3.1. Method

Subjects. 16 naïve male Lister Hooded rats were housed in pairs in standard rat cages, and maintained on a 12-h/12-h light/dark cycle (lights on at 7 am). Rats had unrestricted access to water in their home cages, and were maintained at 85% of their ad-lib weights (range: 300-325g) by giving them restricted access to food each day.

Apparatus. All experimental sessions occurred in eight identical conditioning boxes (H × W × D: 30cm × 24cm × 21cm; Med Associates, Georgia, VT) enclosed in sound-attenuating chambers. These boxes consisted of aluminium front and back walls, clear acrylic sides and top, with a floor consisting of 0.48cm diameter stainless steel rods spaced 1.6cm apart situated above a stainless steel tray. Food pellets (45mg; supplied by MLab; Richmond, IN) and sucrose solution (15%w/w) from a dipper (0.1ml volume) could be separately delivered to a recessed food well equipped with infrared detectors that was located in the centre of the left wall. A speaker mounted on the right wall of the box, opposite the food well, was used to present a tone (2000Hz, 80dB). A ventilation fan maintained background noise (at 68dB). The illumination of a light source (25mm in diameter), mounted 13.5 cm above the floor and 2.5 cm from the back wall, was used as the visual stimulus in an otherwise dark box. Two retractable levers, located 3cm to the left and right of the food well, could be inserted into the box. A computerised interface (Med-PC) was used to insert levers and deliver stimuli, and to record food well entries and lever presses.

Procedure. All rats were first trained to retrieve food pellets and sucrose from the food well in two separate 30-min sessions. Half of the rats in each group were trained to collect sucrose in the first session and pellets in the second, and the remainder received the reverse arrangement. The outcomes were delivered on a random-time (RT) 60-s schedule; each second there was a one in 60 chance that food would become available. Rats then received two stages of training, Pavlovian conditioning and instrumental conditioning. Each involved a series of daily sessions that were conducted at the same time of day for each rat. The order in which these forms of conditioning occurred was counterbalanced, with half of the rats receiving

16, 40-min Pavlovian conditioning sessions first and 12, 24-min instrumental conditioning sessions second, and the remainder receiving the reverse arrangement. The order in which the blocks occurred was counterbalanced.

The levers were retracted during Pavlovian conditioning, and were inserted during instrumental conditioning (a single lever per session). The two types of Pavlovian conditioning trial (i.e., S1-O1 and S2-O2) occurred in separate blocks of sessions. For example, rats received all S1-O1 sessions followed by all S2-O2 sessions. In both types of session, during 10, 2-min presentations of the designated stimulus (e.g., S1), the outcome (e.g., O1) was delivered on an RT 30- schedule. The inter-trial interval (ITI) was 2 minutes. The order in which the sessions occurred was counterbalanced. The rate of food well entries (in responses per minute, rpm) during stimulus presentation and the ITI was used to assess Pavlovian conditioning. The two types of instrumental conditioning trial (i.e., R1-O1 and R2-O2) also occurred in separate blocks of sessions. In these sessions, one lever was inserted and pressing this lever (e.g., R1; pressing the left lever) resulted in the delivery of the designated outcome (O1) on a random-interval (RI) 30-s schedule. Again, the blocked sequence in which the two types of sessions occurred was counterbalanced. The rate of lever pressing (in rpm) was used to assess instrumental conditioning. The identities of the stimuli (i.e., tone or light) that served as S1 and S2, the responses that served as R1 and R2 response (i.e., left or right press), and the outcomes (i.e., pellets or sucrose) that served as O1 and O2, were fully counterbalanced. Prior to the final PIT test trials, each rat received four “refresher” training sessions involving the type of conditioning from their first stage of training (either Pavlovian or instrumental training). This consisted of two Pavlovian sessions with each S1 and S2, or two instrumental sessions with each lever. These additional training trials were arranged in an identical manner to the corresponding conditioning trials from the first stage of training.

Over the following 2 days, rats received two 18-min test sessions in which no outcomes were presented. In one session, the lever corresponding to R1 was inserted, and in the other the lever corresponding to R2 was inserted. The order in which R1 and R2 sessions occurred was counterbalanced. During each test, there were three blocks of trials that each consisted of 2-min periods of three types of trial: S1 present, S2 present, and neither S1 nor S2 present (ITI). Within each block, the computer generated the pseudo-random order in which these three trials were presented. Comparing the rates of R1 and R2 during S1 and S2 assessed outcome-selective PIT. The effect would be evident if the rates of R1 during S1 plus R2 during S2 (Same responses) were higher than the rates of R2 during S1 plus R1 during S2 (Different responses). A general PIT effect, where the presentation of Pavlovian stimuli results in an increase in the rate of instrumental responding over the baseline level, would be evident if there was more instrumental responding during S1 and S2 than during the ITI.

2.3.2. Results and discussion

The mean rates of food well entry during S1 and S2 (pooled) during the final two sessions of Pavlovian conditioning, together with the rates of responding during the ITI in the corresponding sessions are shown in Table 3. Inspection of these scores reveals that the rate of responding during S1 and S2 is higher than during the ITI ($t(15) = 8.06, p < .001$). The mean rates of R1 and R2 (pooled) are also shown in Table 3 for the final two sessions of instrumental conditioning.

Table 3

Mean rates of food well entries (Pavlovian conditioning) and lever presses (instrumental conditioning) in responses per minute during the final sessions of Pavlovian and instrumental training.

	Pavlovian conditioning		Instrumental conditioning
	ITI	Stimuli	
Experiment 1:	9.62	16.92	10.5

The results from the critical test are shown in Figure 7, with the scores collapsed across the various counterbalanced factors. Inspection of this figure shows that there were more Same and Different responses than ITI responses, a general PIT effect. There was also a slight tendency for there to be more Same responses than Different responses. ANOVA confirmed a main effect of trial type (Same, Different or ITI; $F(2,30) = 1.21$, $p < .001$, MSE = 1.44). Subsequent tests revealed that while the rate of Same responses was higher than the rate of ITI responses ($t(15) = 4.92$, $p < .001$), and the rate of Different responses was higher than ITI responses ($t(15) = 3.78$, $p < .005$), the rate of Same responses was not higher than Different responses ($t(15) = .73$, $p > .48$). Thus, although instrumental and Pavlovian conditioning were both successful, Pavlovian CSs only augmented instrumental performance above the baseline levels. That is, Experiment 1 only provided evidence consistent with general PIT. Experiment 2 aimed to assess whether manipulations that should change the discriminability of the outcomes affect the nature of transfer during the PIT test.

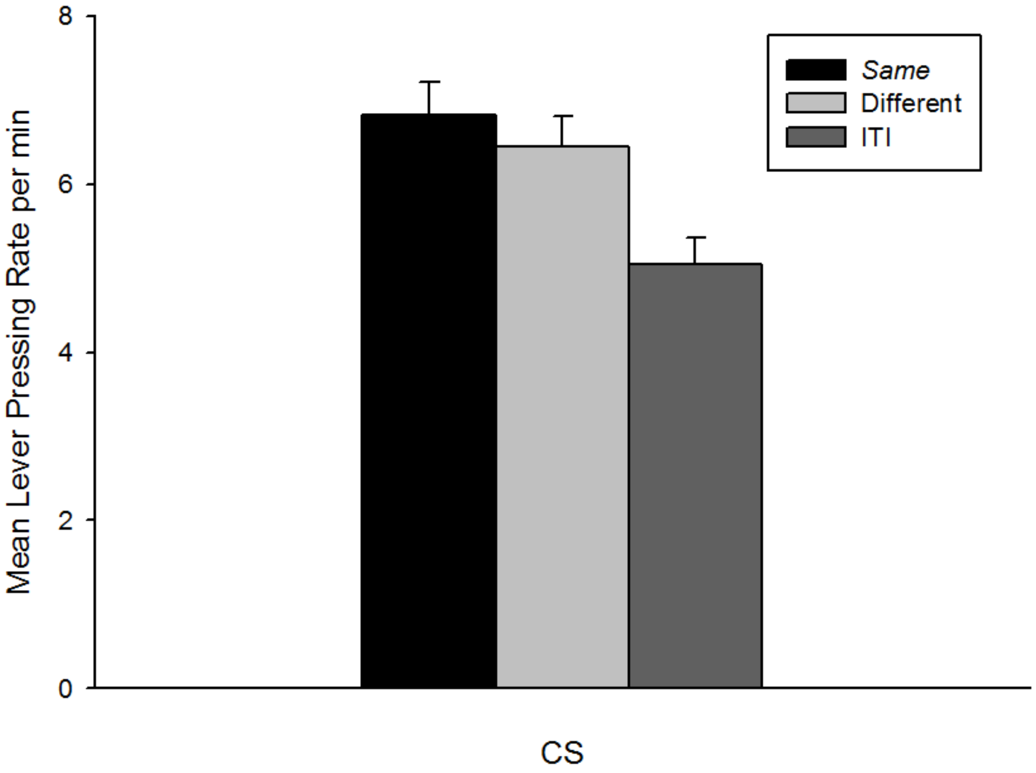


Figure 7: Experiment 1: Test. Mean rates of responding (in responses per minute, rpm; +SEM) on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI) when no such stimulus was present.

2.4. Experiment 2

The design of Experiment 2 is depicted in Table 2 and involved three stages of training. During the pre-exposure phase, rats either received sessions where both outcomes were presented in separate alternating sessions (group Intermixed), or each outcome was presented in a block of consecutive sessions (group Blocked), or was placed in the box with no outcomes being delivered (group None). These sessions were conducted in a chamber in which the response levers and CSs were absent. Following the pre-exposure phase, the same outcomes were paired with Pavlovian stimuli (S1: e.g., a tone; S2: e.g., a light; respectively) in separate sessions. During the instrumental conditioning phase, one response (R1: e.g., a left lever press) was paired with one of the outcomes (O1) and another response (R2: e.g., a right lever press) was paired with the other outcome (O2). Following this training, rats received successive test trials in which the opportunity to respond on either the left or right lever was assessed as a function of the presence of S1 and S2. An outcome-selective PIT effect would be evident if R1 was more frequent during S1 than S2 (and R2 was more frequent during S2 than S1).

2.4.1. Method

Subjects, Apparatus and Procedure. The subjects were 24 naïve male Lister Hooded rats, maintained in the same way as in Experiment 1. The apparatus was that used in Experiment 1. Rats were randomly allocated to the three groups, Intermixed, Blocked and None. Each group received eight pre-exposure sessions over four days. The Intermixed group received deliveries of each outcome in separate alternating 30-min sessions. Half of the rats in this group received sucrose in the first session followed by pellets in the second, and the remainder received the reverse arrangement. The outcomes were delivered on an RT 60-s schedule; each second there was a one in 60 chance that food would become available. The Blocked group received the same treatment with the exception that O1 and O2 were presented in consecutive sessions in a block, as opposed to alternating sessions. Again, half of the group received a block of sucrose followed by a block of pellets and the remainder received the reverse. Rats in group None were placed in the box for 30-min sessions and O1 and O2 were not presented.

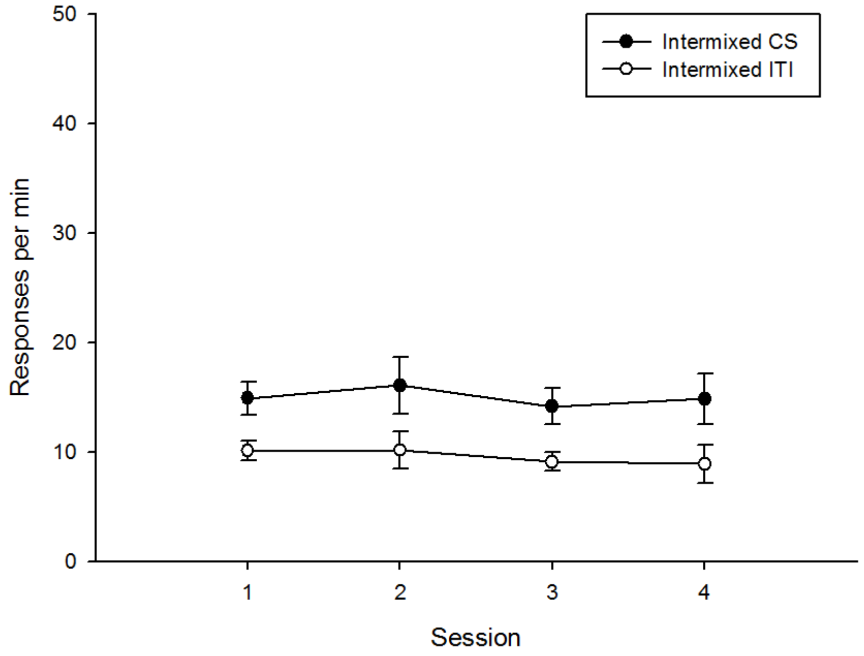
Rats then received Pavlovian conditioning and instrumental conditioning which involved a series of daily sessions that were conducted at the same time of day for each rat. The order in which these forms of conditioning occurred was counterbalanced, with half of the rats in each group receiving the four, 40-min Pavlovian conditioning sessions first and four, 24-min instrumental conditioning sessions second, and the remainder receiving the reverse arrangement. These sessions were conducted in the same way as Experiment 1, with the exception that each type of session within each conditioning phase was presented in alternating sessions as opposed to a consecutive blocks. For example, rats would receive alternating sessions of R1-O1 and R2-O2 in the instrumental phase and alternating S1-O1 and S2-O2 sessions during the Pavlovian phase.

Over the following two days, rats received two 18-min test sessions in which no outcomes were presented, which were conducted in the same manner as Experiment 1. An outcome-selective PIT effect would be evident if the rates of R1 during S1 plus R2 during S2 (Same responses) were higher than the rates of R2 during S1 plus R1 during S2 (Different responses). A general PIT effect, where the presentation of Pavlovian stimuli results in an increase in the rate of instrumental responding over the baseline level, would be evident if there was more instrumental responding during S1 and S2 than during the ITI.

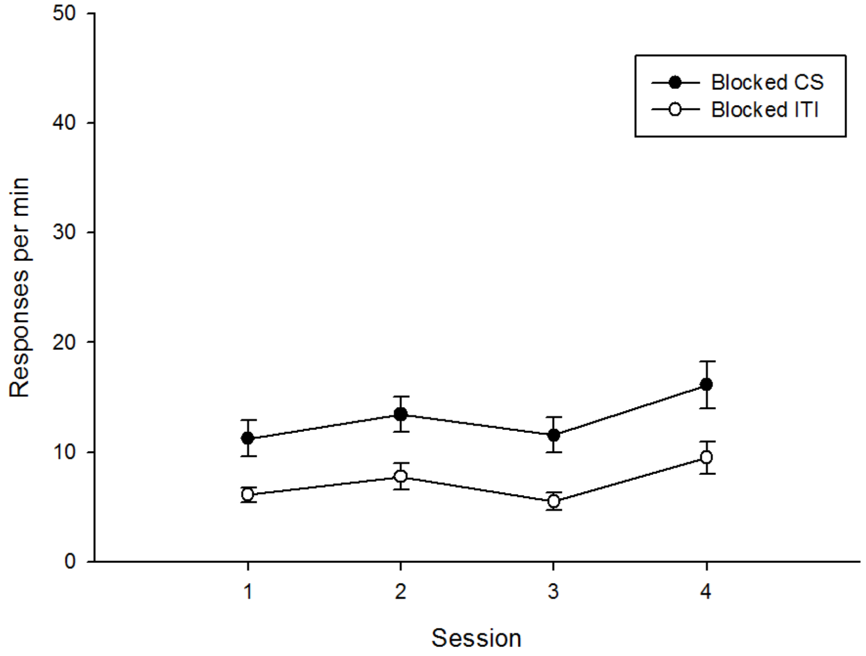
2.4.2. Results and discussion

The rate of responding over the four sessions of Pavlovian and instrumental conditioning are shown in Figures 8 and 9. Inspection of Figure 8 shows the mean rates of food well entries during S1 and S2 (pooled) was greater than during the ITI for all groups. This was confirmed by ANOVA with session (1-4) and trial (CS and ITI) as within-subjects factors and group (Intermixed, Blocked and None) as a between-subjects factor, which found a significant main effect of trial $F(1, 21) = 286.92, p < .001$, MSE = 5.26. There was no main effect of session $F(3, 63) = 1.74, p > .17$, MSE = 44.86; or group $F(2, 21) = 2.97, p > .07$, MSE= 104.85; and no interaction between these factors, $F(3, 63) = 1.62$, $p > .19$, MSE = 6.88. There was also no interaction between trial and group $F<1$, and there was no three way interaction between the three factors, $F<1$. All groups acquired Pavolvian conditioning which was uninfluenced by the pre-exposure treatments.

A)


Intermixed CS
Intermixed ITI
Responses per min
0
10
20
30
40
50
1
2
3
4
Session

B)


Blocked CS
Blocked ITI
Responses per min
0
10
20
30
40
50
1
2
3
4
Session

C)

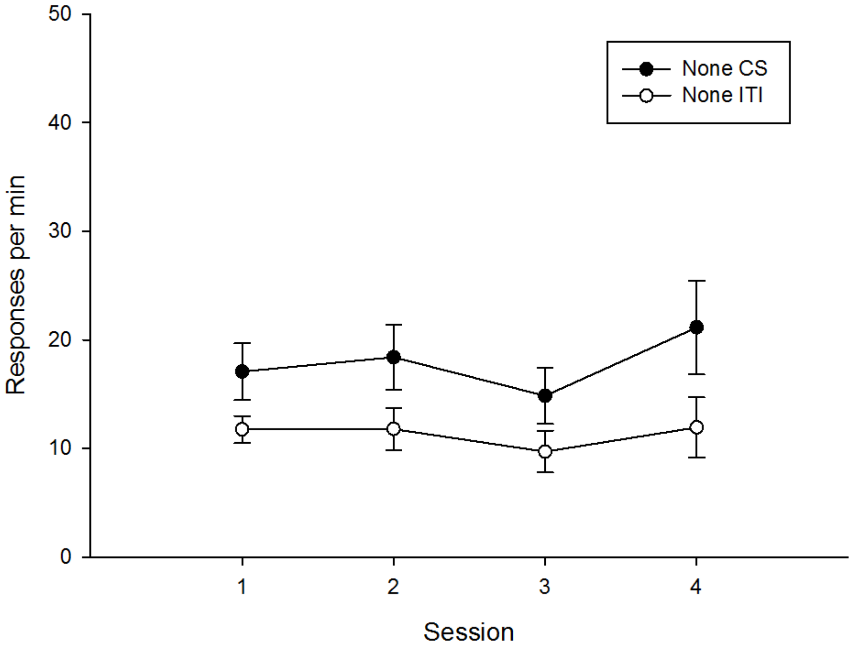


Figure 8: Experiment 2: Pavlovian conditioning. Mean rates of magazine entries (in responses per minute, rpm; ±SEM) during Pavlovian stimuli (responding collapsed across counterbalanced factors); and during the inter-trial interval (ITI) when no such stimulus was present. (A) Intermixed group. (B) Blocked group. (C) None group.

The mean rates of R1 and R2 (pooled) during the four sessions of instrumental conditioning shown in Figure 9. Inspection of this Figure shows responding increases steadily with no difference in responding as a function of pre-exposure treatment; there being some tendency for group None to respond at a lower rate than the groups that had received exposure to the outcomes. This description of the results was partially confirmed by ANOVA with a within-subjects factor of Session (1-4) and a between-subjects factor of group (Intermixed, Blocked and None), which found no significant main effect of group $F(1, 21) = 2.50$, $p > .11$, MSE = 22.83, a main effect of session, $F(3, 6) = 38.90$, $p < .001$, MSE = 4.73, and no interaction between these factors, *F*<1.

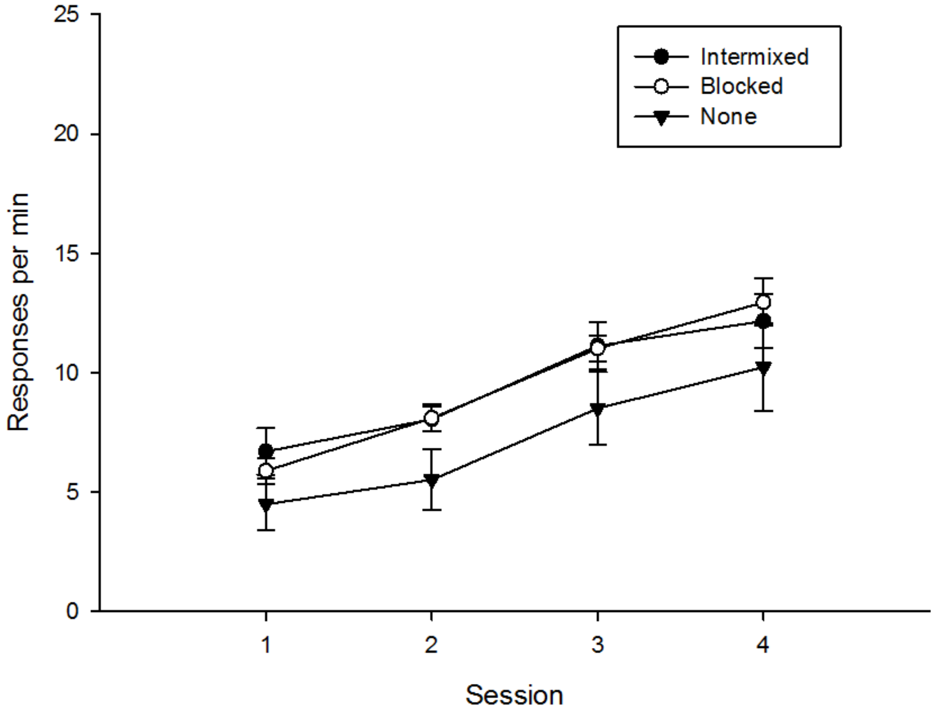


Figure 9. Experiment 2. Instrumental responding. Mean rates of lever press responses (in responses per minute, rpm; ±SEM; collapsed over counterbalanced factors) over the course of instrumental conditioning for groups Intermixed, Blocked and None.

The results from the critical test are shown in Figure 10, with the scores collapsed across the various counterbalanced factors. Inspection of this figure shows that there are more Same and Different responses than ITI responses in groups Intermixed and Blocked, but no similar tendency in group None. However, ANOVA showed that there was no main effect of trial type (Same, Different or ITI, *F*(2,42) = 1.21, *p* > .31, MSE = 4.80), no main effect of group, *F*<1, and no interaction between the two, *F*(2,42) = 1.09, *p* > .39, MSE = 4.80. Although it was apparent that all three groups acquired both Pavlovian and instrumental conditioned responding, there was no indication of either outcome-selective PIT or general PIT, and no effect of pre-exposure. The aim of Experiment 3 was to examine whether the failure to observe any form of PIT was a product of too little training (cf. Holmes et al., 2010). To do so, I extended the amount of conditioning sessions (from four to eight).

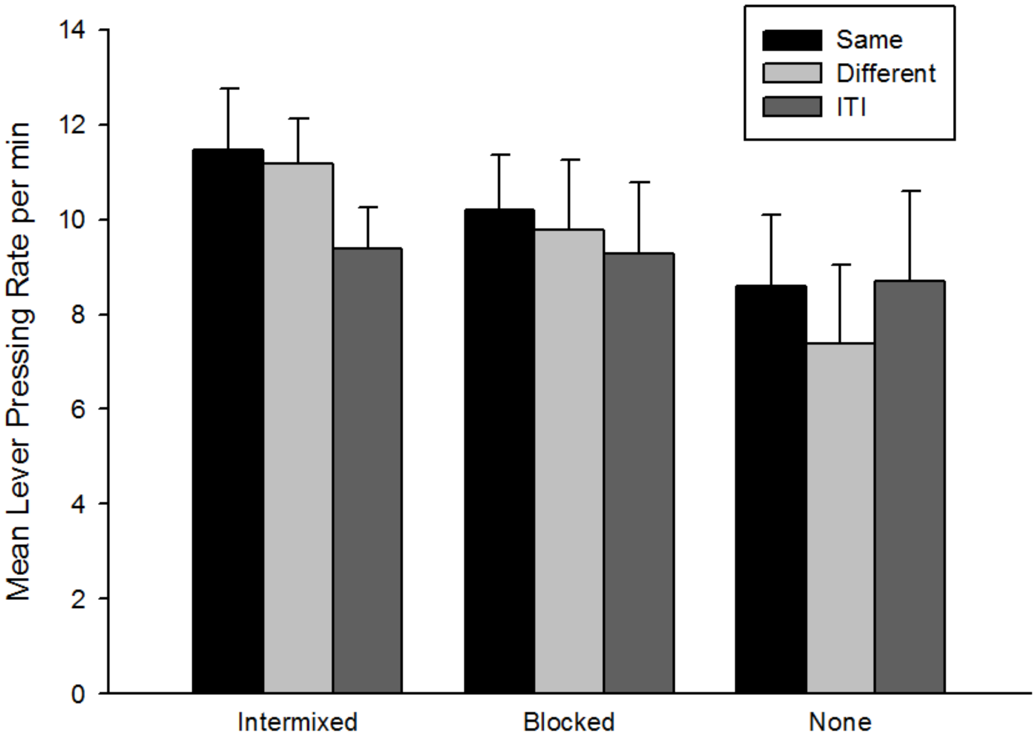


Figure 10. Experiment 2: Test. Mean rates of responding (in responses per minute, rpm; +SEM) for groups Intermixed, Blocked and None on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI) when S1 and S2 were absent.

2.5. Experiment 3

The design of Experiment 3 was similar to Experiment 2, and is summarized in Table 2. As in Experiment 2, rats in each group received a pre-exposure phase, an instrumental conditioning phase and a Pavlovian conditioning phase followed by two test sessions. The principle difference between Experiments 2 and 3 was the fact that in Experiment 3 all rats received eight sessions of both Pavlovian and instrumental conditioning phases as opposed to four sessions of both.

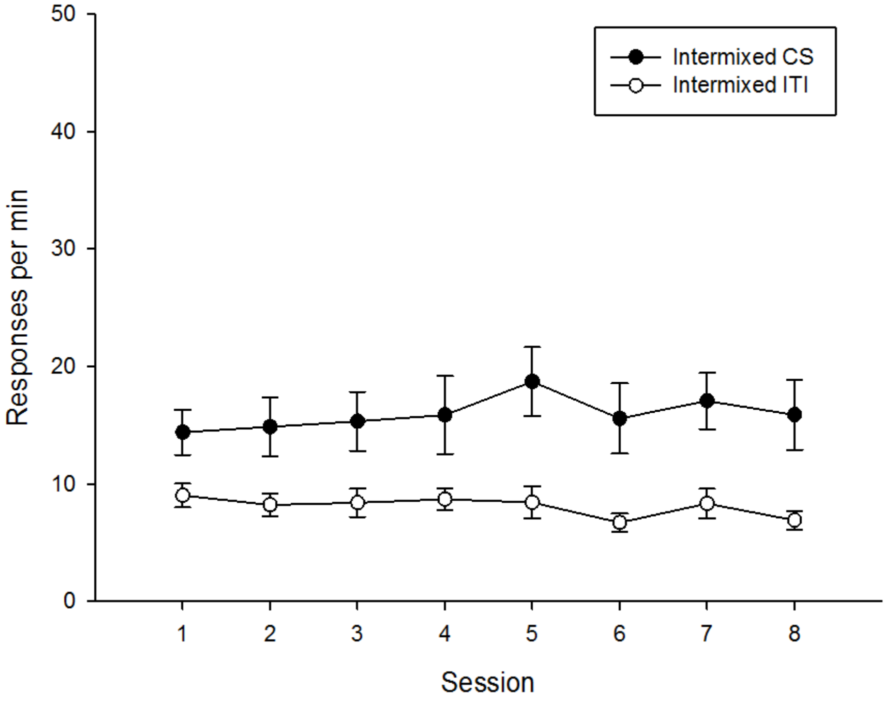
2.5.1. Method

Subjects, Apparatus and Procedure. The subjects were 24 naïve male Lister Hooded rats, maintained in the same way as in Experiments 1 and 2. The apparatus was the same as used in Experiments 1 and 2. After rats had received eight sessions of pre-exposure to the outcome they received eight sessions of instrumental conditioning and eight sessions of Pavlovian conditioning conducted in the same way as Experiment 2. These conditioning trials were arranged in the same was as Experiment 2. Finally, all rats received two extinction tests which were identical to those used in Experiment 2.

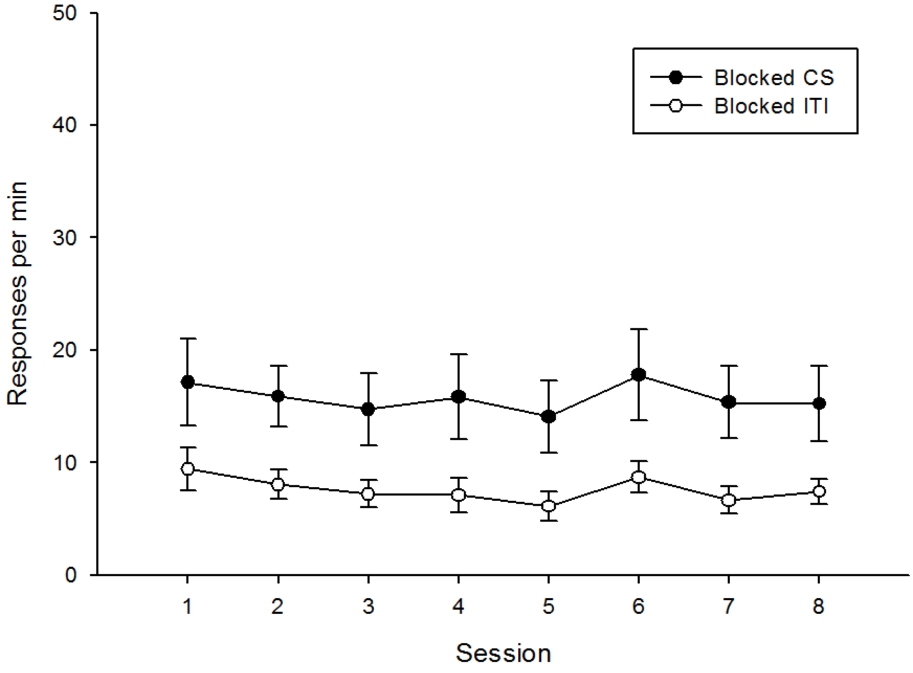
2.5.2. Results and discussion

The rate of acquisition over the eight sessions of Pavlovian and instrumental conditioning are shown in Figures 11 and 12. Inspection of Figure 11 shows the mean rates of food well entries during S1 and S2 (pooled) was greater than during the ITI for all groups. This was confirmed by ANOVA with session (1-8) and trial (CS and ITI) as within-subjects factors and group (Intermixed, Blocked and None) as a between-subjects factor, which found a significant main effect of trial $F(1, 21) = 155.15$, $p < .001$, MSE = 39.19. There was no main effect of session or group and no interaction between these factors, *Fs*<1.

A)


Intermixed CS
Intermixed ITI
Responses per min
50
40
30
20
10
0
1
2
3
4
5
6
7
8
Session

B)


Blocked CS
Blocked ITI
Responses per min
50
40
30
20
10
0
1
2
3
4
5
6
7
8
Session

C)

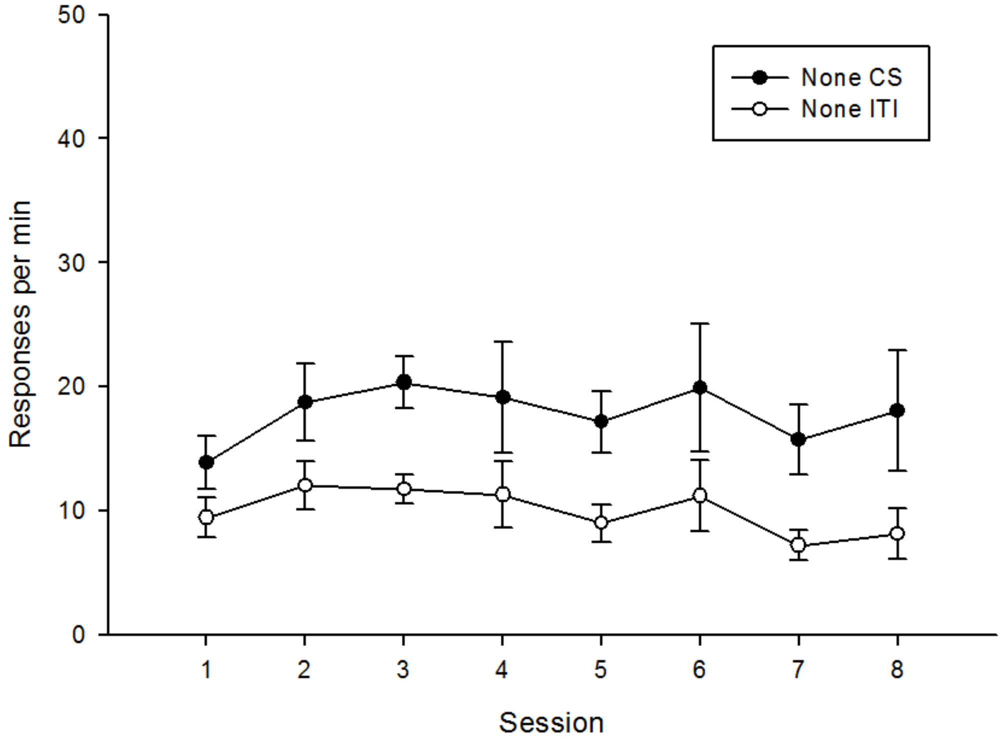


Figure 11: Experiment 3: Pavlovian conditioning. Mean rates of magazine entries (in responses per minute, rpm; ±SEM) during Pavlovian stimuli (responding collapsed across counterbalanced factors); and during the inter-trial interval (ITI). (A). Intermixed group. (B). Blocked group. (C). None group.

The mean rates of R1 and R2 (pooled) during the eight sessions of instrumental conditioning are shown in Figure 12. Inspection of this figure shows responding increased steadily with no difference in responding as a function of group. This description was confirmed by ANOVA, with a within-subjects factor of session (1-8) and a between-subjects factor of group (Intermixed, Blocked and None), which found no significant main effect of group, $F<1$, a main effect of session $F(7, 14) = 20.16$, $p < .001$, MSE = 14.52, and no interaction between these factors, $F<1$.

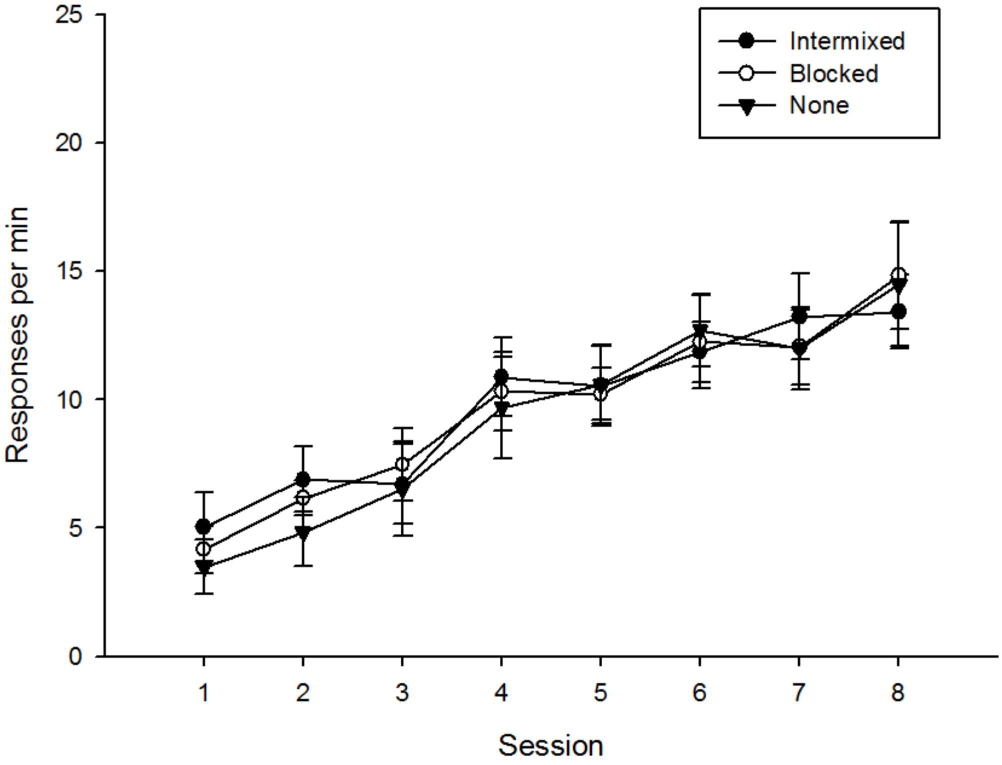


Figure 12: Experiment 3. Instrumental conditioning. Mean rates of lever press responses (in responses per minute, rpm; ±SEM; collapsed over counterbalanced factors) over the course of instrumental acquisition for groups Intermixed, Blocked and None.

The results from the critical test are shown in Figure 13, with the scores collapsed across the various counterbalanced factors. Inspection of this figure suggests there were more Same and Different responses than ITI responses in the case of the Blocked and None groups, whereas in group Intermixed there were more Same responses than both Different and ITI responses. There is also a tendency for the overall levels of responding to be greater in group Intermixed than in the other groups (particularly group None). However, statistical analysis only partially supported this description of the results. ANOVA confirmed that there was a main effect of trial type (Same, Different or ITI), $F(2,42) = 3.56$, $p < .05$, MSE = 53.88, no main effect of group, $F(2,21) = 1.25$, $p > .31$, MSE = 53.88, and no interaction between the two; $F<1$. Subsequent tests collapsed across group revealed a general PIT effect. That is, the levels of responding were greater than ITI during both Same trials; $t(23) = 2.11$, $p < .05$; and Different trials; $t(23) = 3.39$, $p < .005$; though there was no difference between the two, $t(23) = .14$, $p > .99$.

A comparison of the results of Experiments 2 and 3 suggest that increasing the amount of training, from four to eight sessions, resulted in a general PIT effect. Pre-exposure to the outcome, again, did not appear to influence general PIT or outcome-selective PIT. Although inspection of Figure 12 suggests that there were more Same responses than Different and ITI in the Intermixed group, whereas only evidence of general PIT in group Blocked, these differences were not statistically significant. This suggests that outcome-selective PIT is not encouraged by procedures that should increase the distinctiveness of the outcomes (cf. Blair et al., 2003).

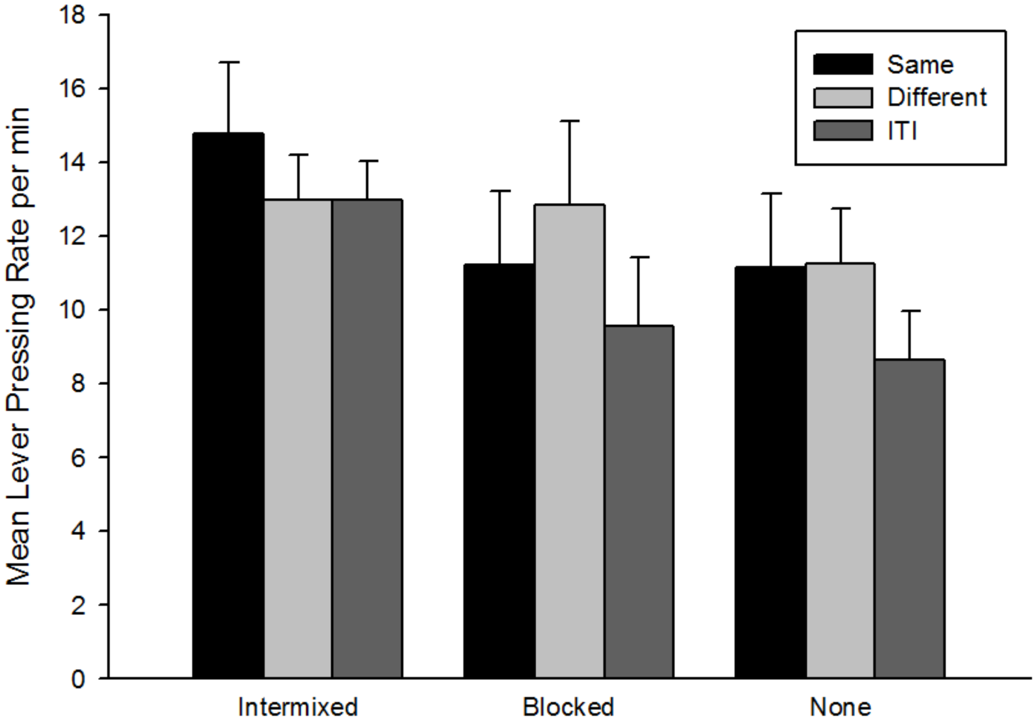


Figure 13. Experiment 3: Test. Mean rates of responding (in responses per minute, rpm; +SEM) for Groups Intermixed, Blocked and None on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI).

2.6. General discussion

The way in which Pavlovian and instrumental processes interact is fundamental to understanding response selection. Outcome-selective and general PIT demonstrate that instrumental performance can be modulated by Pavlovian stimuli. However, the conditions under which such effects are observed have not been systematically examined. Experiments 1-3 investigated whether conditions that should affect the discriminability of the outcomes influence whether outcome-selective or general PIT are observed. Experiment 1 confirmed that a general PIT effect, but not an outcome-selective PIT effect, could be observed when conditioning phases were blocked. This result could be taken to suggest that when a procedure is used that should promote the encoding of the common, affective properties of the outcomes, a general PIT effect is observed. Unfortunately, direct tests of this analysis in Experiments 2 and 3 failed to provide any evidence to support it: There was no indication of outcome-selective PIT in rats given differential exposure to the outcomes prior to Pavlovian and instrumental conditioning. It should be acknowledged, however, that there was no independent assessment of the discriminability of the two outcomes. Thus, although the procedures employed should increase the discriminability of the outcomes (cf. Blair et al., 2003) there was no internal evidence that the exposure treatments were effective. One cannot discount the possibility that pre-exposure to the outcomes only served to result in them coming to be coded as more as opposed to less similar, perhaps because they were presented in the same context (for a review, see Honey, Close & Lin, 2010). For example, it is plausible to suppose that presenting two outcomes in the same context resulted in associations forming between the outcomes and that context, thereby resulting in a reduction in the discriminability of the outcomes through a process of acquired equivalence. Clearly, under such conditions, the impact of outcome pre-exposure would not be to encourage outcome-selective PIT.

It was clear from Experiments 2 and 3 that four conditioning sessions in each phase was insufficient to generate outcome-selective PIT. Although others have reported such an effect using a limited numbers of training sessions (e.g., Zorawski & Killcross, 2003; Delamater & Oakeshott, 2007), it remains the case that the analysis provided by Holmes et al. (2010) suggests that increasing the amount of training (in particular, the second stage of training) increases the likelihood of observing outcome-selective PIT. It was with this observation in mind that I changed the procedure used in the next experiment (Experiment 4) and later more analytic experiments presented in Chapter 3.

2.7. Conclusion

The results of Experiments 1-3 serve to confirm that outcome-selective PIT is indeed an elusive phenomenon. They also provide little support for the idea that procedures that should increase the distinctiveness of the outcomes will increase the likelihood of observing outcome-selective PIT. However, as will become clear, this failure to replicate outcome-selective PIT did not reflect any general features of the procedure that was employed in Experiments 1-3: It is possible to observe this effect when rats are given more extensive, intermixed instrumental and Pavlovian conditioning, and this effect is modulated by how the Pavlovian conditioning trials are arranged. Chapter 3 provides clear evidence of outcome-selective PIT under conditions that begin to allow a more definitive statement about its associative origins. As will become evident, the failures to replicate outcome-selective PIT in Experiments 1-3, together with an evaluation of several features of the extant literature, suggested a novel theoretical analysis of outcome-selective PIT that is evaluated in Chapter 3.

Chapter 3

Pavlovian-to-instrumental transfer: Paradoxical effects of the Pavlovian relationship examined

3.1 Summary

Four further experiments with rats examined the origin of outcome-selective PIT. The design of these experiments reflected the failures of Experiments 1-3 to yield outcome-selective PIT, together with adopting aspects of the successful procedures that have been reported in the literature. Experiment 4 used a standard procedure, where outcomes were embedded within extended CSs, to demonstrate the basic effect: Pavlovian stimuli augmented instrumental lever presses that had been paired with the same outcomes. Experiments 5 and 6 showed that after instrumental conditioning, whereas a CS trained using a backward conditioning procedure produced outcome-selective PIT, forward conditioning with a CS did not. These results are consistent with the idea that backward conditioning results in the outcome provoking its associated instrumental response during the CS and thereby allows a stimulus–response association to be acquired that directly generates outcome-selective PIT at test. Experiment 7 provided direct support for the assumptions that underlie this stimulus-response analysis. These results, and other paradoxical effects of the Pavlovian relationship, are incongruent with accounts of outcome-selective PIT that rely on a stimulus-outcome-response chain. The implications of these findings for current theoretical accounts of PIT are explored.

3.2 Introduction

As already noted in Chapter 1, a fundamental aspect of human and animal behaviour is the capacity to select appropriate responses in the face of changing conditions. The study of the behaviour of rats, motivated by access to small quantities of food, has revealed two important influences on response selection. Consider first the simple situation in which rats are placed in an operant chamber and receive food (O1) for pressing one lever (R1) and sucrose (O2) for pressing another lever (R2). Under these conditions, rats will distribute their efforts evenly between the two levers. However, if the value of one of the outcomes (O2) is reduced by sating the rats with it, or by pairing it with an illness-inducing agent, then they become more likely to perform the response (R1), whose outcome remains valued, than the response (R2) whose outcome is no longer valued. Instrumental performance is not only determined by the current value of the outcome: If a conditioned stimulus (S1) is presented that has separately signalled one of the outcomes (e.g., O1), then rats are more likely to perform the response that is also associated with that outcome (i.e., R1) than the response associated with the other outcome (i.e., R2). These two influences have been taken to reflect the fact that instrumental training results in the formation of R1–O1 and R2–O2 associations, and that whereas outcome devaluation reduces the likelihood that a response associated with that now devalued outcome will be performed, a Pavlovian stimulus increases the selection of the response associated with the same outcome. The ability of a stimulus to increase the likelihood of a response associated with the same outcome is referred to as outcome-selective PIT (for a recent review, see Holmes et al., 2010). The origin of outcome-selective PIT is the focus of the experiments presented in this chapter.

One account of outcome-selective PIT is based on a chain of associations involving the Pavlovian stimulus and a response that is mediated by a shared, sensory-specific outcome representation (i.e., a stimulus-outcome-response or S–O–R chain). According to this account, instrumental conditioning establishes reciprocal associations between the response and a sensory-specific outcome representation (i.e., R1–O1 and O1–R1), and Pavlovian conditioning results in the formation of an association between the stimulus and the same representation (i.e., a S1–O1 association). The resulting S1–O1–R1 chain enables S1 to provoke R1 during the PIT test (for a review, see Balleine & Ostlund, 2007; see also Corbit & Balleine, 2005). This type of analysis is clearly appealing: Taken at face value it imbues an organism with a means of integrating the disparate forms of specific associative knowledge that are generated during two fundamental types of learning: Pavlovian and instrumental. Moreover, this analysis underpins the use of PIT to investigate the behavioral and neural mechanisms that underlie response selection (Blundell et al., 2001; Corbit et al., 2001; Dickinson et al., 2000; Hall et al., 2001; Holland & Gallagher, 2003; Johnson et al., 2007), and its use as a model of components of drug addiction (e.g., Everitt & Robbins, 2005).

As noted in Chapter 1, there is some evidence that seems to be incongruent with the theoretical analyses outlined in the immediately preceding paragraph. To recap: First, outcome-selective PIT is not influenced by devaluation of the outcome immediately before the critical test (e.g., Holland, 2004). Given the fact that such devaluation can exert an outcome-selective influence on instrumental performance, its failure to affect outcome-selective PIT is disquieting. In fact, it is just such a resistance to the effects of outcome devaluation that is taken to be diagnostic of habitual (i.e., stimulus-response) behavior (e.g., Adams & Dickinson, 1981; Adams, 1982; Dickinson, 1985). One could argue that the S–O–R chain, that is activated during the PIT test, influences instrumental performance independently of any R–O association (see Balleine & Ostlund, 2007); for example, because the O–R association involves only the sensory properties of the outcome. However, one could equally well argue that once the response is activated then its vigor should be (at least partly) determined by the current value of the outcome with which it is associated (see Balleine & Ostlund, 2007). Second, extinction manipulations that should reduce the efficacy of the S-O association (i.e., presenting the S alone immediately before the PIT test; e.g., Delamater, 1996) do not influence outcome-selective PIT. This finding might be taken to suggest that the S–O component of the S–O–R chain is left unaffected by the extinction procedure, and the S–O–R association therefore retains its capacity to influence responding during the PIT test. However, this analysis is tendentious: it relies on the assumption that outcome-selective PIT is dependent upon the integrity of the S–O association during the test; and I shall present an alternative view that does not require one to appeal to the operation of an S–O association during the test. Moreover, while it is now widely accepted that extinction treatments do not erase excitatory associations (Bouton, 2004), this is not to say that their efficacy is unaffected by extinction. Finally, outcome-selective PIT does not appear to be reduced by variables that should reduce the strength of S–O associations: the effect is more reliable with a long stimulus than a short stimulus (Holland & Gallagher, 2003), and PIT procedures typically use a long CS where outcomes are delivered within the stimulus rather than at the end (Holland & Gallagher, 2003; Holland, 2004; Balleine & Ostlund, 2007; Holmes et al., 2010). Both of these manipulations might be expected to result in relatively poor excitatory S–O associations, irrespective of whether one is considering the sensory or motivational properties of the outcome. This is not to deny the fact that outcome-selective PIT has been observed when the outcome is delivered at the offset of the CS (Trapold & Overmier, 1972; Kruse, Overmier, Konz, & Rokke, 1983; Delamater & Holland, 2008), but rather to point to some features of the boundary conditions of the effect that seem somewhat curious. It was with these observations in mind that I began considering an alternative way in which outcome-selective PIT could be generated: one where the effect does not rely on the integration of S–O and O–R associations but instead reflects a form of S–R learning, which occurs during the training stages of the PIT procedure.

I began by adopting the assumption that instrumental training results in bidirectional associations between responses and their outcomes: between R1 and O1 (i.e., R1–O1 and O1–R1), and between R2 and O2 (i.e., R2–O2 and O2–R2). Now, during a later period of Pavlovian conditioning (S1–O1 and S2–O2), S1 and S2 will be present when O1 and O2 activate the motor programs of their corresponding responses (i.e., R1 and R2). Under these conditions, S1 could become associated with R1 and S2 with R2 merely as a result of their temporal contiguity (see Honey et al., 1998). The resulting S1–R1 and S2–R2 associations will be capable of generating an outcome-selective PIT effect at test without any mediation by the outcome. A similar analysis can be developed for the case in which the order of the two stages of training is reversed: If it is assumed that O1 and O2 come to evoke representations of S1 and S2 as a consequence of Pavlovian conditioning (which will be likely when the outcomes are embedded in the stimuli; cf. Holland & Gallagher, 2003; Holland, 2004; Corbit et al., 2001; Holmes et al., 2010), then these representations of S1 and S2 will be activated during later instrumental training when R1 is paired with O1 and R2 is paired with O2. This should allow S1–R1 and

S2–R2 to form. Whatever the order in which Pavlovian and instrumental training takes place, the process of mediated stimulus-response learning (see Holland, 1981, 1983; Iordanova et al., 2011; Wheeler et al.,, 2008; but see, Wagner, 1981) will depend upon the stimuli (or their evoked memories) and responses (or their corresponding motor programs) occurring in close temporal contiguity. Importantly, the influence of the mediated S1–R1 and S2–R2 associations on instrumental performance will be unaffected by outcome devaluation before the PIT test (e.g., Holland, 2004).

Moreover, S-R associations might be expected to be less sensitive to extinction than S-O associations, as there is not an equivalent violation of the information contained in the S-R association when the outcome is not delivered. Therefore, presentations of the stimuli before the PIT test might not influence the likelihood of observing outcome-selective PIT in spite of the fact that they should undermine the efficacy of the S–O association during the test (see Delamater, 1996).

The aim of the four experiments reported in this chapter was to investigate the potential role of mediated S–R learning in outcome-selective PIT. Experiment 4 sought to replicate outcome-selective PIT using a procedure in which the outcomes were embedded within the stimuli, and in which rats received rather more intermixed Pavlovian and instrumental training than in Experiments 2 and 3 before the PIT test (cf. Holmes et al., 2010). In Experiments 5 and 6, rats first received instrumental training (R1–O1 and R2–O2) and were then given forward pairings of one stimulus with one outcome (S1–O1) and backward pairings of a second stimulus with the other outcome (O2–S2). The capacity of S1 and S2 to generate outcome-selective PIT was then assessed. This is based on the assumption that the process of excitatory mediated stimulus-response learning (cf. Delamater, LoLordo, & Sosa, 2003) should be most likely to occur if the outcome was presented and the motor program for the response evoked during the stimulus. On this basis, the backward (O2–S2) relationship should allow O2 to evoke R2 during S2, and foster the development of an excitatory S2–R2 association. This association will enable S2 to provoke R2 during the PIT test. However, the forward S1–O1 pairings should be less likely to result in S1–R1 learning, because S1 will not be present when R1 is activated by O1; and S1 will be less likely to activate R1 during the PIT test. In contrast, according to the S–O–R analysis, PIT should rely on the ability of the S to evoke the memory of O during the test. Given the fact that forward S1–O1 pairings should result in stronger S–O associations than will backward O2–S2 pairings, then the S–O–R analysis might predict that S1 would be more likely to generate PIT than will S2. Of course, as I shall discuss later, without independent assays of the strength of the critical outcome-specific S–O associations, this prediction is moot (cf. Delamater, 1995). However, Experiment 7 directly assessed the assumptions upon which the mediated stimulus-response analysis is based, assumptions that are not made by the S–O–R analysis.

3.3. Experiment 4

Experiment 4 used a within-subjects PIT procedure in which the outcomes (O1 and O2) were embedded within the stimuli (S1 and S2; cf. Holland & Gallagher, 2003; Holland, 2004; Corbit et al., 2001; Holmes et al., 2010). This design is depicted in Table 4 and involved two stages of training. During Pavlovian conditioning one stimulus (S1; e.g., a tone) was paired with one outcome (O1; e.g., a food pellet), and another stimulus (S2; e.g., a light) was separately paired with another outcome (O2; e.g., sucrose); during instrumental conditioning one response (R1; e.g., a left lever press) was paired with one of the outcomes (O1), and another response (R2; e.g., a right lever press) was paired with the other outcome (O2). As in previous studies of PIT (and Experiments 1-3), the two types of training trials, for both Pavlovian and instrumental training, were presented in separate sessions, to reduce the possibility that the two outcomes would be associated with both stimuli or with both responses.[1] After these stages of training, rats received successive test trials in which the opportunity to respond on either the left or right lever was assessed as a function of the presence of S1 and S2. An outcome-selective PIT effect would be evident if R1 was more frequent during S1 than S2 (and R2 was more frequent during S2 than S1). For simplicity, I refer to R1 during S1 and R2 during S2 as Same responses, and R2 during S1 and R1 during S2 as Different responses.

---

[1] It should be acknowledged that this arrangement also increases, relative to a procedure in which the trial types are intermixed, the likelihood that the outcomes could become discriminative stimuli for specific responses. This would provide another potential mechanism for PIT, with the Pavlovian stimulus activating the discriminative stimulus (i.e., the outcome) during the test and thereby triggering the response (see Ostlund &Balleine, 2007).

Table 4

Design of Experiments 4-6 (and designation of instrumental responses)

Instrumental	Pavlovian	Test
Experiment 4		
R1-O1	S1+O1	S1 (Same) or S2 (Different): R1
R2-O2	S2+O2	S2 (Same) or S1 (Different): R2
Experiment 5		
R1-O1	S1-O1	S1 (Same) or S2 (Different): R1
		or
R2-O2	O2-S2	S2 (Same) or S1 (Different): R2
Experiment 6		
R1-O1	S1-O1	S1: R1 (Same) or R2 (Different)
R2-O2	O2-S2	S2: R2 (Same) or R1 (Different)

Note: R1 and R2 refer to left and right lever presses; O1 and O2 denote food pellets and sucrose; and S1 and S2 are a tone and a light. S1+O1 and S2+O2 indicate that the outcomes were presented throughout S1 and S2, whereas S1-O1 and O2-S2 indicates that the outcome was presented after S1 but before S2. In Experiments 4 and 5, the influence of S1 and S2 on instrumental responding during the test was conducted with a single lever present, whereas in Experiment 6 both levers were available at the same time. Same and Different indicate whether the designated response (R1 or R2) had been paired with the Same or Different outcome as the stimulus (S1 or S2).

3.3.2 Method

Subjects, Apparatus and Procedure. The subjects were 32 naïve male Lister Hooded rats, maintained in the same way as in Experiments 1-3. The apparatus was that used in Experiments 1-3. All rats were first trained to retrieve food pellets and sucrose from the food well in two separate 30-min sessions. Half of the rats in each group were trained to collect sucrose in the first session and pellets in the second, and the remainder received the reverse arrangement. The outcomes were delivered on an RT 60-s schedule; each second there was a 1 in 60 chance that food would become available. Rats then received two stages of training, Pavlovian conditioning and instrumental conditioning, that each involved a series of daily sessions that were conducted at the same time of day for each rat. The order in which these forms of conditioning occurred was counterbalanced, with half of the rats receiving the 16, 40-min Pavlovian conditioning sessions first and 12, 24-min instrumental conditioning sessions second, and the remainder receiving the reverse arrangement.

The levers were retracted during Pavlovian conditioning and were inserted during instrumental conditioning (a single lever per session). The two types of Pavlovian conditioning trial (i.e., S1–O1 and S2–O2) occurred in separate, alternating sessions. In both types of session, during 10, 2-min presentations of the designated stimulus (e.g., S1), the outcome (e.g., O1) was delivered on an RT 30-s schedule. The ITI was two minutes. The order in which the sessions occurred was counterbalanced. The rate of food well entries (in responses per minute, rpm) during stimulus presentation and the ITI was used to assess Pavlovian conditioning. The two types of instrumental conditioning trial (i.e., R1–O1 and R2–O2) also occurred in separate, alternating sessions. In these sessions, one lever was inserted and pressing this lever (R1; e.g., pressing the left lever) resulted in the delivery of the designated outcome (O1) on an RI 30-s schedule. Again, the alternating sequence in which the two types of sessions occurred was counterbalanced. The rate of lever pressing (in rpm) was used to assess instrumental conditioning. The identities of the stimuli (i.e., tone or light) that served as S1 and S2, the responses that served as R1 and R2 response (i.e., left or right press), and the outcomes (i.e., pellets or sucrose) that served as O1 and O2, were fully counterbalanced.

Before the final PIT test trials, each rat received four “refresher” training sessions involving the type of conditioning from their first stage of training (either Pavlovian or instrumental conditioning). This consisted of two Pavlovian sessions with S1 and S2, or two instrumental sessions with each lever. These additional training trials were arranged in an identical manner to the corresponding conditioning trials from the first stage of training. Over the following 2 days, rats received two 18-min test sessions in which no outcomes were presented. In one session, the lever corresponding to R1 was inserted, and in the other the lever corresponding to R2 was inserted. The order in which R1 and R2 sessions occurred was counterbalanced. During each test, there were three blocks of trials that each consisted of 2-min periods of three types of trial: S1 present, S2 present, and neither S1 nor S2 present (ITI). Within each block, the computer generated the pseudorandom order in which these three trials were presented. Comparing the rates of R1 and R2 during S1 and S2 assessed outcome-selective PIT. The effect would be evident if the rates of R1 during S1 plus R2 during S2 (Same responses) were higher than the rates of R2 during S1 plus R1 during S2 (Different responses). A general PIT effect, where the presentation of Pavlovian stimuli results in an increase in the rate of instrumental responding over the baseline level, would be evident if there was more instrumental responding during S1 and S2 than during the ITI.

3.3.3 Results and discussion

The mean rates of food well entry during S1 and S2 (pooled) during the final two sessions of Pavlovian conditioning, together with the rates of responding during the ITI in the corresponding sessions are shown in Table 5. Inspection of these scores reveals that the rate of responding during S1 and S2 is higher than during the ITI, $t(31) = 8.49$, $p < .001$. The mean rates of R1 and R2 (pooled) are also shown in Table 5 for the final two sessions of instrumental conditioning.

Table 5

Mean rates of food well entries (Pavlovian conditioning) and lever presses (instrumental conditioning) in responses per minute during the final sessions of Pavlovian and instrumental training.

	Pavlovian conditioning		Instrumental conditioning
	ITI	Stimuli	
Experiment 4:	4.6	11.7	12.0
		Forward (S1)/Backward (S2)	
Experiment 5:	4.4	23.8 / 10.9	10.6
Experiment 6:	5.1	30.6 / 14.1	10.8

Note: Forward and backward refer to the order in which the Pavlovian stimuli (S1 and S2) were paired with their respective outcomes (O1 and O2; i.e., S1-O1 and O2-S1).

The results from the critical test are shown in Figure 14, with the scores collapsed across the various counterbalanced factors. Inspection of this figure shows that there were more Same responses than both Different responses and ITI responses. ANOVA confirmed that there was a main effect of trial type (Same, Different or ITI;

F(2, 62) = 7.56, *p* < .001, MSE = 3.01). Subsequent tests revealed that while the rate of Same responses was higher than both the rate of Different responses, *t*(31) = 2.08, *p* < .05 and ITI responses, *t*(31) = 4.81, *p* < .001, the rate of Different responses was not higher than the rate of ITI responses, *t*(31) = 1.42, *p* > .15.

The results of Experiment 4 confirm that a standard outcome-selective PIT effect can be observed when the outcomes are embedded within the stimuli during Pavlovian conditioning and more extensive intermixed training is given (cf. Experiments 2 and 3). Experiments 5 and 6 investigated which aspect of this procedure is critical to observing the effect: the fact that the S precedes O or the fact that the O precedes the S. As I have already noted, the S–O–R analysis of outcome-selective PIT predicts that the strength of the S–O association should be critical; to the extent that a forward S–O arrangement results in a stronger S–O association than does a backward O-S arrangement, then outcome-selective PIT should be more likely after forward than backward conditioning. In contrast, according to the S-R analysis, the backward O-S arrangement should ensure that the O activates the motor program for the R during a period when the S is present and thereby engender the S-R association that is responsible for outcome-selective PIT. This form of mediated S-R learning should be less evident after forward conditioning - when the outcome will evoke the motor program for the response when the stimulus is no longer present. Therefore, outcome-selective PIT should be more likely after backward than forward conditioning.

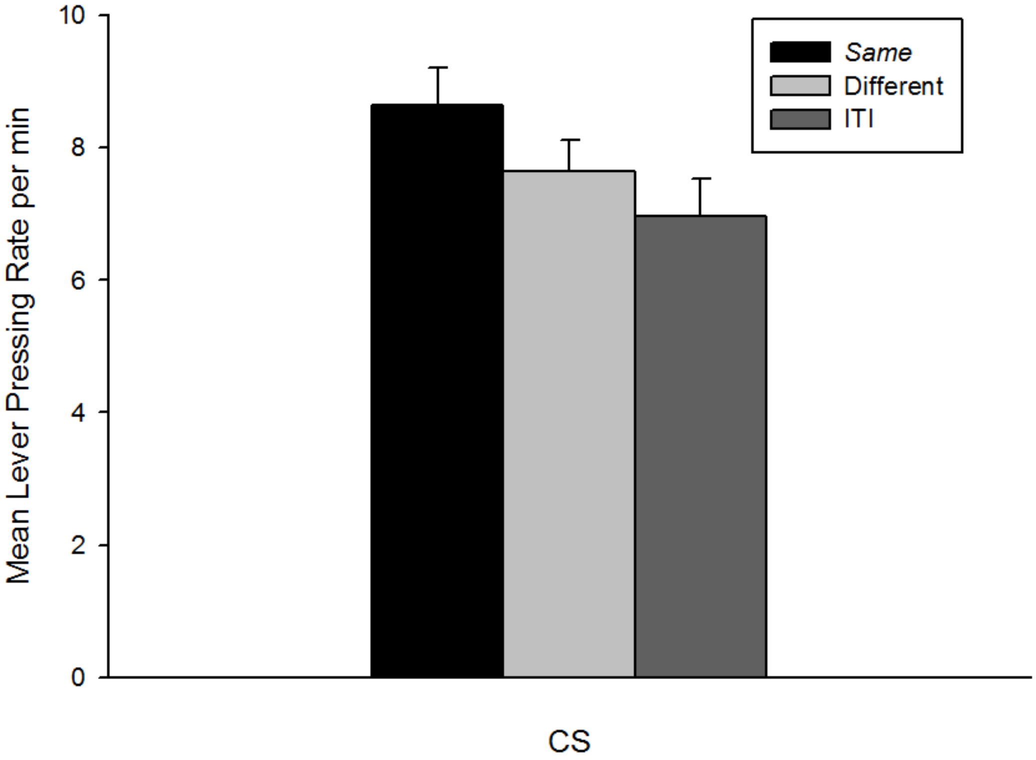


Figure 14. Experiment 4: Test. Mean rates of responding (in responses per minute, rpm; +SEM) on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI) when no such stimulus was present.

3.4. Experiment 5

The design of Experiment 5 was similar to Experiment 4 and is summarised in Table 4. In Experiment 5, however, all rats first received instrumental conditioning in which R1 was paired with O1 and R2 was paired with O2. Rats then received forward pairings of S1 with O1 (i.e., S1–O1) and backward pairings of O2 with S2 (i.e., O2–S2). As in any backward conditioning procedure, it is difficult to be certain that the order in which the events are scheduled to occur is the order in which they are experienced by the animal. For example, the rats might still be consuming O2 when S2 is presented, and the O2–S2 relationship might be better characterised as simultaneous rather than backward. However, the levels of conditioned responding to S1 and S2 should provide one independent assay of the strength of the Pavlovian association; albeit one that does not provide separate indices of the strengths of associations involving the sensory and motivational components of the outcomes. After Pavlovian conditioning with S1 and S2, half of the rats then received a test in which one lever was inserted into the box and S1 (i.e., Forward) and S2 (Backward) were presented, and the remainder received a test in which the other lever was inserted and S1 and S2 were presented. This design means that half of the rats had the opportunity to respond on a lever that had been paired with the same outcome as the forward S1 and a different outcome to S2, and the remainder had the opportunity to respond on a lever that had been paired with the same outcome as S2 and a different outcome to S1. This design, specifically the fact that only one lever is inserted in a given test session, means that evidence of outcome-selective PIT depends on a between-subjects comparison for both S1 and S2.

3.4.5. Method

Subjects, Apparatus and Procedure. 48 naïve male Lister Hooded rats, which were maintained in the same way as in Experiments 1-4, were used in Experiment 5. The apparatus was the same as that used in Experiments 1-4. After rats had been trained to collect food pellets and sucrose from the food well in the same way as Experiment 4, they then received two instrumental training sessions per day for six days, followed by one Pavlovian session per day for nine days. All sessions were run at the same time every day for each rat. The instrumental training sessions were arranged in the same way as Experiment 4. During each Pavlovian conditioning session, each rat received three S1–O1 trials and three O2–S2 trials that were presented in pseudorandom order (with a mean ITI of eight minutes). The levers were retracted for the duration of the session. For S1–O1 training trials, S1 was presented for 10-s and immediately followed by O1; whereas for the O2–S2 trials, O2 was delivered and the next food well entry triggered a 1-s interval after which S2 was presented for 10-s. This interval was intended to allow the rat to consume O2 before S2 was presented. Before the test session, all rats received four instrumental refresher sessions in the same way as Experiment 4. During the extinction test, a single lever was inserted and the influence of both S1 (Forward) and S2 (Backward) on responding was assessed. For half of the rats, the test consisted of Forward-Same and Backward-Different trials, and for the remainder the test consisted of Forward-Different and Backward-Same trials. That is, the critical comparisons, comparing the levels of same and different lever pressing during (i) the forward S1, and (ii) the backward S2, were both between subjects. The test session consisted of eight 30-s presentations of S1 and S2, that were delivered in a pseudorandom order. The ITI was extended from two minutes, as in Experiments 1-4, to a variable ITI (mean: 8 minutes). This was to reduce the possibility that any residual excitation from preceding CS presentations may be present during successive presentations of a CS. The overall baseline level of instrumental responding was measured during the 30-s periods that immediately preceded S1 and S2. Other details of Experiment 5 that have not been mentioned were the same as in Experiment 4.

3.4.6. Results and discussion

The mean rates of food well entry during S1 (forward) and S2 (backward) during the final two sessions of Pavlovian conditioning, together with the baseline rates of responding during the ITI in the corresponding sessions, are shown in Table 5. Inspection of these scores reveals that the rate of responding during S1 was higher than during S2, and that both were higher than during the ITI. ANOVA confirmed that there was a main effect of trial type (S1, S2, ITI; $F(2, 94) = 47.73$, $p < .001$, MSE = 111.68); and subsequent comparisons showed that S1 (forward) elicited more responding than S2 (backward), ($t(47) = 5.40$, $p < .001$), and that there was more responding during S2 (backward) than during the ITI ($t(47) = 3.35$, $p < .005$). The mean rate of lever pressing on the final day of training is also shown in Table 5.

Figure 15 depicts the critical results from the test in Experiment 5, the mean rates of lever pressing during the Pavlovian stimuli and the ITI[2]. Inspection of this figure suggests that the presence of S1 did not augment the response (R1; i.e., Same responses) associated with the same outcome (i.e., O1) relative to the response (R2; i.e., Different responses) associated with a different outcome (i.e., O2). That is, there was no suggestion of an outcome-selective PIT effect when the Pavlovian stimulus had a forward relationship with the outcome during training. However, inspection of Figure 15 indicates that S2 augmented the response (R2; i.e., Same responses) associated with the same outcome (i.e., O2) relative to the response (R1; i.e., Different responses) associated with a different outcome (i.e., O1). That is, backward training resulted in outcome-selective PIT. ANOVA revealed that there was a main effect of trial type (Same, Different or ITI; $F(2, 92) = 8.74$, $p < .001$, MSE = 2.86), an effect of group (which lever was presented; $F(1, 46) = 3.72$, $p < .001$, MSE = 22.90), and an interaction between these factors, $F(2, 92) = 12.20$, $p < .001$, MSE = 2.86. Subsequent tests revealed that although there was no difference between the

---

[2] For ease of presentation and comparison, the ITI bars in Figure 15 depict the overall mean for both groups. However, in the statistical analysis, the individual ITI scores for each group were included separately (Group 1: S2R1 and S1R2; mean = 6.81 rpm) and Group 2: S2R2 and S1R1; mean = 5.76 rpm), which did not differ, $t(46) = 1.38$, $p > .18$.

rate of R1 and R2 during S1, *t*(46) = -1.66, *p* > .10, the rate of Same responses was higher than Different responses during S2, *t*(46) = 2.21, *p* < .03. The rate of Same and Different responses during S1 was greater than during the ITI, *t*(23) = 3.24, *p* < .005; and *t*(23) = 4.62, *p* < .001, respectively. The rate of Same responses during S2 was also greater than during the ITI, *t*(23) = 2.30, *p* < .03, but the rate of Different responses during S2 was no greater than during the ITI, *t*(23) = .14, *p* > .89. During the test, S1 (mean = 8.07 rpm) elicited more food well entries than S2 (mean = 6.20 rpm; *t(*47) = 2.31, *p* < .05). Although S1 elicited more food well entries than S2, there is no indication that this fact could have influenced their ability to selectively augment instrumental responding: the overall levels of instrumental responding during S1 and S2 were relatively similar. The results of Experiment 5 are striking and are consistent with the suggestion that outcome-selective PIT effect might be based upon a mediated S-R association formed during training, rather than on the integration of S-O and O-R associations during the test. Before I explore the implications of this conclusion I sought to extend and replicate the results upon which it was based.

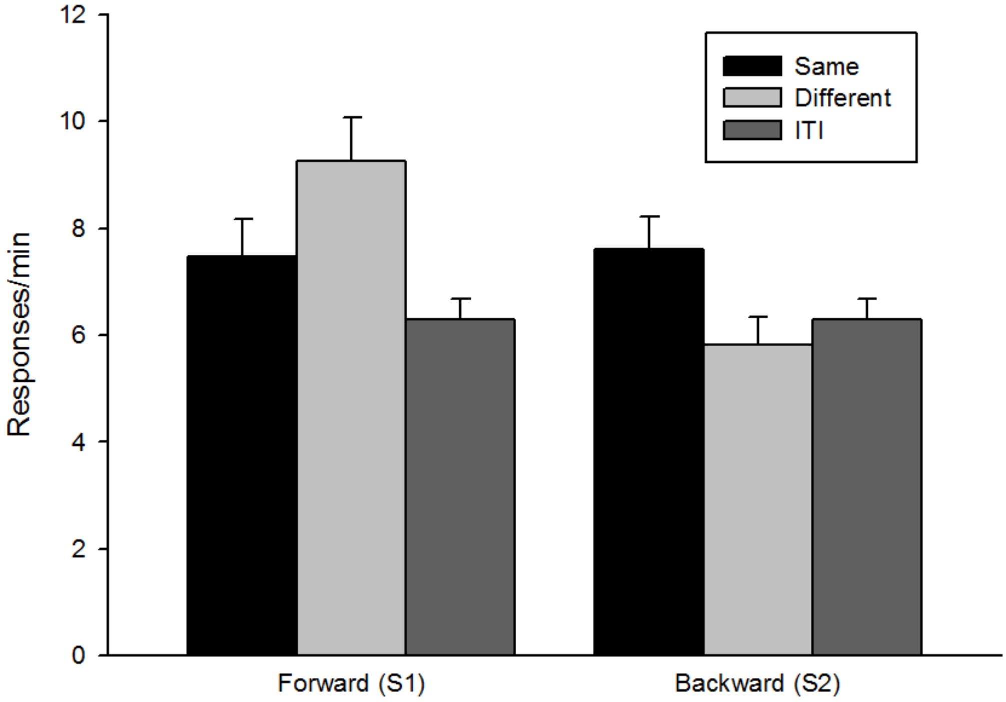


Figure 15. Experiment 5: Test. Mean rates of responding (in responses per minute, rpm; +SEM) on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was presented; and during the inter-trial interval (ITI) when no such stimulus was present. Original Pavlovian training had either involved forward pairings (S1-O1) or backward pairings (O2-S2).

3.5. Experiment 6

Experiment 6 used the experimental design summarized in Table 4 to assess the reliability of the results of Experiment 5 using a fully within-subjects design. As in Experiment 5, rats first received R1–O1 and R2–O2 pairings before receiving S1–O1 and O2–S2 pairings. The principal difference between Experiments 5 and 6 was the fact that in Experiment 6 all rats received tests in which both levers were inserted into the chamber and the influence of S1 and S2 on the frequency of R1 and R2 was assessed.

3.5.1. Method

Subjects, Apparatus and Procedure. The subjects were 48 naïve male Lister Hooded rats, maintained in the same way as in Experiments 1-5. The apparatus was that used in Experiments 1-5. After rats had been training to retrieve the outcomes from the food well, they received instrumental conditioning that was conducted in the same way as Experiment 4. All rats then received nine sessions of Pavlovian conditioning in which they received S1–O1 and O2–S2 training trials. These conditioning trials were arranged in the same way as Experiment 5, and there followed four days of refresher training in which independent training on both levers occurred in an intermixed order (e.g., Left, Right, Left, Right). Finally, all rats received a single extinction test in which both levers were simultaneously available. The test sessions consisted of four 30-s presentations of S1 and S2, that were delivered in a pseudorandom order, with a variable ITI (mean: 8 min). For the purposes of statistical analysis the test was divided into two consecutive blocks with two S1 and two S2 trials.

3.5.2. Results

Table 5 shows the rates of responding during the final day of Pavlovian and instrumental conditioning. Examination of Table 5 shows that the stimulus that had a forward relationship to the outcome (S1) elicited more food well responses than the stimulus that had a backward relationship to the outcome (S2; $t(23) = 2.88, p < .01$); and S2 elicited more responding than was evident in the ITI, $t(23) = 6.36, p < .001$. The mean presented in the right column of Table 5 also confirms that instrumental conditioning had been successful.

The results of the critical test are shown in Figure 16, with the upper panel depicting the first block of testing and the lower block depicting the second block of testing. Inspection of the upper panel reveals that during the forward stimulus (S1), the rate of R1 (Same) responses was similar to both the rate of R2 (Different) responses, and the rate of instrumental responding during the ITI. That is, there was little sign of an outcome-selective PIT effect when S1 was presented. In contrast, during the backward stimulus (S2), the rate of R2 (Same) responses was higher than the rate of R1 (Different) responses; the rate of R1 was similar to the rate of instrumental responding during the ITI. That is, there was an outcome-selective PIT effect when S2 was presented. This pattern of results replicates that observed in Experiment 5. In the second block of testing, the results of which are shown in the lower panel, there was no indication of an outcome-selective PIT effect during S1 or S2, but there was more lever pressing during S1 than during S2. ANOVA with within-factors of trial type and block as factors revealed a significant main effect of trial type (Same Forward, Different Forward, Same Backward, Different Backward or ITI; $F(4, 92) = 4.02$, $p < .005$, MSE = 7.87), block ($F(1, 23) = 24.88$, $p < .001$, MSE = 6.67), and an interaction between these factors ($F(4, 92) = 21.67$, $p < .005$, MSE = 4.68). Separate ANOVAs revealed that there was an effect of trial type in both blocks (Block 1: $F(4, 92) = 2.80$, $p < .03$, MSE = 6.47; and Block 2: $F(4, 92) = 5.78$, $p < .001$, MSE = 6.09). Supplementary analyses of the results from Block 1, revealed that during S1 the rate of R1 (Same) was no greater than the rate of R2 (Different; $t(23) = .79$, $p > .17$), but that during S2 the rate of R2 (Same) was higher than the rate of R1 (Different; $t(23) = 2.14$, $p < .05$). The rates of R1 during S1 was elevated above the baseline ($t(23) = 2.31$, $p < .05$), but the rate of R2 during S1 was not, $t(23) = 1.19$, $p > .25$; and the rate of R2 during S2 was elevated above the baseline, $t(23) = 3.56$, $p < .002$, but the rate of R1 during S2 was not, $t(23) = .73$, $p > .47$. The rates of food well entries during Block 1 were higher during S1 (mean =

16.50) than during S2 (mean = 4.17; $t(23) = 4.31$, $p < .001$). During the second block of testing, there was no difference in the rate of R1 and R2 during either stimulus (largest $t(23) = .31$, $p > .76$). However, the rate of instrumental responding (see Figure 3; $t(23) = 5.82$, $p < .001$) and food well entries (S1 mean = 9.04 rpm and S2 mean = 1.42 rpm; $t(23) = 4.46$, $p < .001$) was higher during S1 than during S2; the rates of R1 and R2 during S1 were elevated above baseline (largest, $t(23) = 3.05$, $p < .01$), but the rates of R1 and R2 during S2 were not (largest, $t(23) = -.52$, $p > .61$.

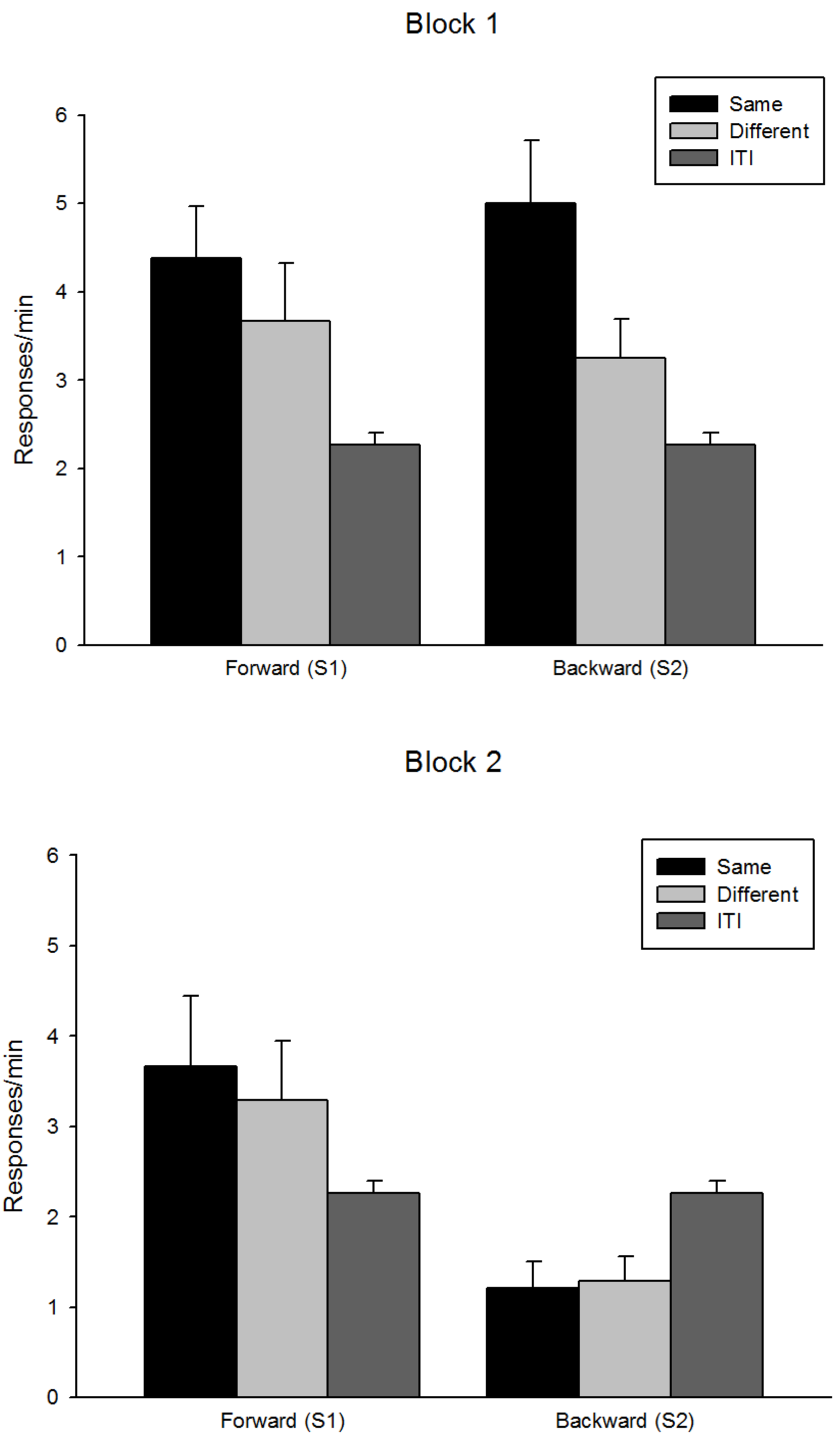


Figure 16. Experiment 6: Test. Mean rates of responding (in responses per minute, rpm; +SEM) during Block 1 (left panel) and Block 2 (right panel) on levers that had either been paired with the Same outcome or Different outcome as the Pavlovian stimulus that was present; and during the inter-trial interval (ITI) when no such stimulus was present. Original Pavlovian training had either involved forward pairings (S1-O1) or backward pairings (O2-S2).

3.5.3. Discussion

The results of block 1 in Experiment 6 confirm the reliability of those of Experiment 5: whereas a stimulus (S1) that had a forward relationship with an outcome (i.e., S1–O1) did not augment a response (R1) paired with that same outcome, a stimulus (S2) that had a backward relationship with an outcome (i.e., O2–S2) did augment a response (R2) that had been paired with a shared outcome (i.e., O2). Two observations make it unlikely that competition between lever pressing and food well entries was responsible for this pattern of results: the overall levels of instrumental responding during S1 and S2 were similar in block 1, and there was both more instrumental responding and food well entries during S1 than S2 in block 2. The theoretical basis of the latter observation, that in block 2 a general PIT effect was more apparent during S1 than S2, will be given further consideration in the general discussion.

My interpretation of the findings from block 1, that outcome-selective PIT was more likely during S2 than S1, is that backward O2–S2 pairings allow O2 to provoke the motor program for R2 during S2, whereas forward S1–O1 pairings will mean that the motor program for R1 will be provoked when S1 is no longer present. It is this difference that should allow the link between S2 and the motor program for R2 to become stronger than the link between S1 and the motor program for R1. Under these conditions, test presentations of S2 should be more likely to produce an outcome-selective PIT effect than should those of S1. This analysis relies on two critical assumptions: First, during Pavlovian conditioning, the outcomes will provoke the responses that they were paired with during instrumental training (i.e., O1 will provoke R1 and O2 will provoke R2). Ostlund and Balleine (2007) have reported results consistent with this suggestion, from an experiment examining the reinstatement of instrumental conditioning after extinction. Presentation of an outcome was found to provoke the response with which it had previously been paired, thus confirming the reciprocal nature of the relationship between an instrumentally trained response and an outcome. Second, backward pairings (i.e., O1–S1 and O2–S2) should allow S1 to be linked to the motor program for R1 and S2 to be linked to the motor program for R2. Experiment 7 examined the validity of these assumptions - assumptions that do not underpin the S–O–R analysis.

3.6. Experiment 7

The design of Experiment 7 is summarized in Table 6. All rats first received R1–O1 and R2–O2 pairings, and in the second stage received backward O1–S1 and O2–S2 pairings intermixed with forward S1–O2 and S2–O1 pairings. During the presentations of S1 and S2, both levers were inserted. If the first assumption described above is accurate, then on backward trials responding after O1 and O2 should be congruent with the responses associated with these outcomes during instrumental training. That is, on backward trials R1 should be provoked after O1 and during S1 (on O1–S1 trials), and R2 should be provoked after O2 and during S2 (on O2–S2 trials). I will refer to these as congruent responses, and contrast them with incongruent responses (i.e., R2 after O1 and R2 after O1). To the extent that these responses become linked to S1 and S2 as a result of temporal contiguity (cf. Honey et al., 1998) then they should also be evident when S1 and S2 are presented alone (on S1–O2 and S2–O1 trials), and in spite of the fact that on these trials the stimuli are paired with the alternate outcomes. Thus, whenever S1 is presented (on O1–S1 and S1–O2 trials) there should be more R1s (congruent responses) than R2s (incongruent responses), and whenever S2 is presented (on O2–S2 and S2–O1 trials) there should be more R2s (congruent responses) than R1s (incongruent responses).

Table 6

Design of Experiment 7 (and designation of instrumental responses)

Instrumental	Test
R1-O1 R2-O2	Backward O1-S1: R1 (Congruent), R2 (Incongruent) Backward O2-S2: R2 (Congruent), R1 (Incongruent) Forward S1-O2: R1 (Congruent), R2 (Incongruent) Forward S2-O1: R2 (Congruent), R1 (Incongruent)

Note: R1 and R2 refer to left and right lever presses; O1 and O2 denote food pellets and sucrose; and S1 and S2 are a tone and a light. After instrumental R1-O1 and R2-O2 training, all rats received backward training trials (O1-S1 and O2-S2) and forward training trials (S1-O2 and S2-O1). On backward trials, after O1 and O2 were delivered, both levers were inserted during S1 and S2, and whether responses were congruent or incongruent with those paired with O1 and O2 during instrumental training was assessed. On forward trials, the levers were inserted during S1 and S2 allowing assessment of the tendency for these congruent and incongruent responses to be evident when S1 and S2 were presented alone (and paired with the alternative outcomes).

3.6.1. Method

Subjects, Apparatus and Procedure. The subjects were 16 naïve male Lister Hooded rats, maintained in the same way as in Experiments 1–6, and the apparatus was that used in those experiments. After rats had been trained to retrieve the outcomes from the food well, they received instrumental conditioning that was conducted in the same way as Experiment 4 with the exception that they received an additional (refresher) training session on each lever. All rats then received three sessions of training in which they received two types of trials: Backward and Forward. The backward (O1–S1 and O2–S2) trials were arranged in the same way as the backward training trials in Experiment 5 and 6, with the notable exception that both levers were inserted into the box throughout S1 and S2 and were withdrawn from the box upon the offset of S1 and S2. The forward (S1–O2 and S2–O1) trials were arranged in the same way as the forward training trials in Experiment 5 and 6, with the notable exception that both levers were again inserted into the box throughout S1 and S2 and were withdrawn from the box upon offset of S1 and S2. In each of the sessions, there were four presentations of the four trial types. The ITI was an average of eight minutes, and the sequences in which the 16 trials were presented in the three sessions were arranged so that each block of four trials contained one trial of each type, which were presented in a pseudorandom order with the constraint that there were no more than two trials involving the same stimulus in succession. The analysis presented here focuses on the first eight trials of each of the three sessions, because the level of lever pressing (which was never reinforced during the test stimuli, S1 and S2) extinguished over the course of the sessions. However, the pattern of responding during the second half of the sessions remained numerically consistent with that observed in the first half.

3.6.2. Results and discussion

The terminal rate of lever pressing on the last day of instrumental training, with a mean of 11.10 rpm, was similar to those seen in Experiments 4–6. The results of principal interest from Experiment 7 are summarized in Figure 17. Inspection of this figure reveals that there was less lever pressing during the stimuli (S1 and S2) on backward than on forward trials, presumably because consuming the outcomes competed with lever pressing during S1 and S2 on the backward trials. More importantly, there were more congruent lever presses (R1 during S1 and R2 during S2) than incongruent lever presses (R2 during S1 and R1 during S2) on both backward (O1–S1 and O2–S2) and forward (S1–O2 and S2–O1) trials. ANOVA with congruent/incongruent and backward/forward as factors revealed an effect of congruence, $F(1, 15) = 6.86$, $p < .05$, MSE = 2.48, and an effect of trial type, $F(1, 15) = 12.08$, $p < .005$, MSE = 10.02, but no interaction between these factors, $F<1$, MSE = 7.86. These results provide direct support for the assumptions that underlie a mediated stimulus-response analysis of the origin of outcome-selective PIT effects observed in Experiments 4–6.

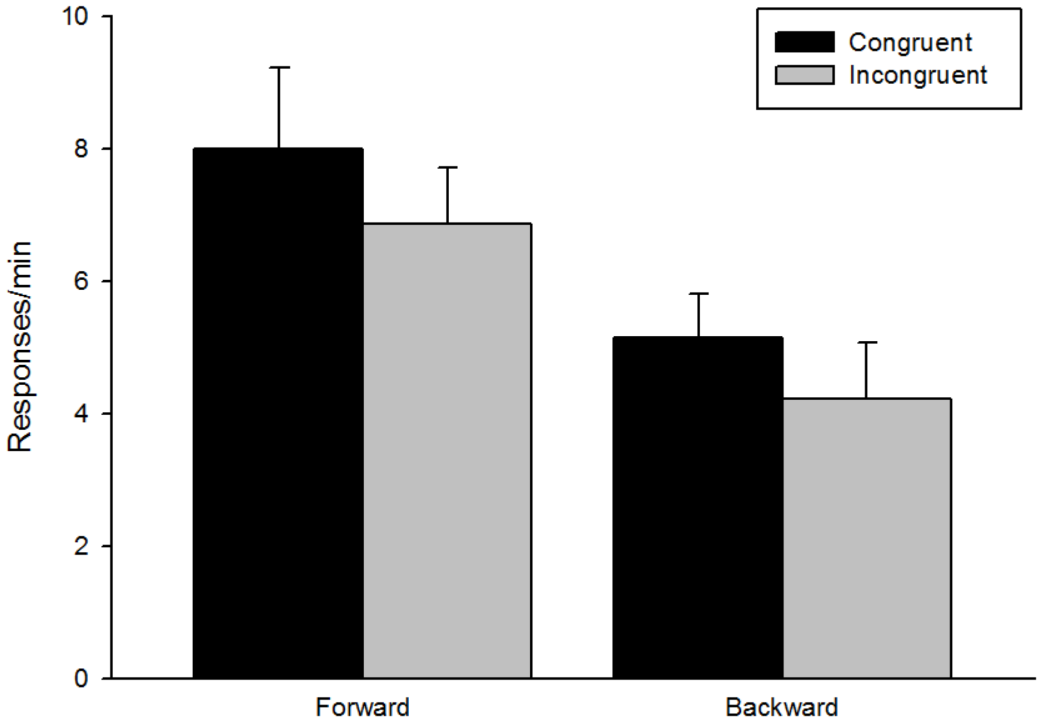


Figure 17. Experiment 7: Test. Mean rates of congruent and incongruent responding (in responses per minute, rpm; +SEM) during S1 and S2 on backward (O1-S1 and O2-S2) and forward (S1-O2 and S2-O1) trials.

3.7. General discussion

The nature of the interactions between Pavlovian and instrumental learning processes is a longstanding theoretical issue (e.g., Rescorla & Solomon, 1967; Trapold & Overmier, 1972) that has been the focus of intensive recent investigation at both behavioral and neuronal levels (e.g., Balleine & Ostlund, 2007; Everitt & Robbins, 2005; Holmes et al., 2010). The observation that a Pavlovian stimulus can selectively augment an instrumental response that has been paired with the same outcome has been taken to reflect a simple sequence of events: instrumental conditioning establishes both response-outcome and outcome-response associations, Pavlovian conditioning results in the formation of a stimulus-outcome association, and when the Pavlovian stimulus is then presented it provokes the response via a stimulus–outcome–response (i.e., S–O–R) associative chain (e.g., Balleine & Ostlund, 2007; Corbit & Balleine, 2005). This is a simple analysis, but there is an equally simple one that helps to address some of the otherwise paradoxical features of outcome-selective PIT that were identified in the introduction to this chapter. According to this alternative analysis, the outcome-selective PIT effect relies on a form of stimulus-response learning brought about by the ability of the outcome to associatively provoke either (i) the instrumental response in the presence of the stimulus (during Pavlovian conditioning) or (ii) the stimulus in the presence of the response (during instrumental conditioning).

Experiment 4 demonstrated outcome-selective PIT under training procedures in which the effect has been reliably observed in the past, when the outcomes occurred throughout the Pavlovian stimulus (for a review, see Holmes et al., 2010). Experiments 5 and 6 then showed that this effect was evident after backward pairings of the outcome and stimulus (i.e., outcome-stimulus), but not after forward pairings (i.e., stimulus-outcome). These results are inconsistent with the S–O–R account of outcome-selective PIT (e.g., Balleine & Ostlund, 2007; Corbit & Balleine, 2005). According to this account, manipulations that should increase the ability of a stimulus to provoke a memory of the outcome (forward rather than backward pairings) should increase the likelihood of outcome-selective PIT being observed. Instead, the pattern of results from Experiments 5 and 6 is consistent with the suggestion that backward outcome-stimulus trials should be particularly conducive to the formation of mediated stimulus–response associations: the presentation of the outcome should associatively activate the response, or its corresponding motor program, at the same time as the stimulus is presented. The resulting stimulus-response association should directly generate performance during the test, which is indicative of outcome-selective PIT, but with a quite different origin than usually envisaged. According to this analysis, this phenomenon does rely on the stimulus and response sharing a common outcome, but the role of the outcome is to mediate a direct stimulus-response association. Finally, the results of Experiment 7 provide direct support for the assumptions upon which this stimulus-response analysis is founded. After pairing a response with an outcome, the outcome was later able to provoke the response during a stimulus with which it had a backward relation; and this arrangement engendered stimulus-response learning. However, even if one accepted that the stimulus–response analysis described above provides a coherent explanation for the conditions under which outcome-selective PIT is observed, then two questions still remain to be considered: Why have other experiments, using backward outcome-stimulus conditioning, produced an outcome-selective suppression of an instrumental response associated with the same outcome (Delamater et al., 2003)? What is the origin of the more general facilitation of instrumental behaviour by a Pavlovian stimulus? These questions will be considered in turn.

Delamater et al. (2003) conducted a series of experiments with rats in which two responses were associated with different outcomes (R1–O1 and R2–O2), and O1 and O2 were followed by different stimuli (S1 and S2) after an interval of 10 seconds (i.e., O1-S1 and O2-S2). During subsequent testing, rats were less likely to perform R1 than R2 during S1 and were less likely to perform R2 than R1 during S2. This is the opposite pattern of results to that reported in Experiments 5-7 and other instances of outcome-selective PIT. Delamater et al. (2003) suggested that the backward trace conditioning trials resulted in the development of inhibitory associations between S1 and O1 and between S2 and O2. They argued that during the PIT test S1 would inhibit O1 (and S2 would inhibit O2) and thereby reduce the likelihood of R1 during S1 (and R2 during S2). This analysis is consistent with the S–O–R account of outcome-selective PIT but relies on the capacity, for example, of S1 to inhibit O1, and for this inhibition to reduce the likelihood of R1. The alternative S-R analysis, which Experiments 5–7 support, suggests a different interpretation. It is possible that during O2–S2 trials the response is activated into a state that allows the development of an excitatory S2–R2 association (cf. Experiments 5-7), but that leaving an interval between O2 and S2 (O2-S2) results in R2 decaying into a refractory state that supports the formation of an inhibitory S2–R2 association (cf. Wagner, 1981). This inhibitory S2–R2 association should be directly able to reduce the likelihood of R2 after S2 during the PIT test. Further research will be required to assess the merit of these two analyses, but considerations of parsimony would suggest that a common explanatory framework should be applied to the results of Experiments 5–7 and those reported by Delamater et al. (2003).

In addition to the outcome-selective influence of Pavlovian stimuli on instrumental performance, such stimuli also exert a general effect on instrumental responding (e.g., Corbit & Balleine, 2005; see also Brandon & Wagner, 1991). Indeed, it is just such a general effect that provides a plausible account for the results observed during the second block of testing in Experiment 6: the presentation of the forward trained stimulus (S1) resulted in more lever pressing than did the backward trained stimulus (S2). One explanation for general PIT is based upon the view that a Pavlovian stimulus can access the motivational components of the outcome during the test and these components will, perhaps by virtue of an outcome-response association (cf. Balleine & Ostlund, 2007), generate responses associated with outcomes having the same motivational value. There is some evidence to support this assertion from an outcome-selective PIT study conducted by Corbit and Balleine (2005). In this study, a third stimulus (S3) was paired with a third outcome (O3) which had not been previously paired with either R1 or R2, but shared a motivational class with outcomes that had been previously paired with both responses. Presentations of S3 resulted in a general increase in responding compared to baseline (i.e., general PIT). This finding suggests S3 activated the motivational components of O3, which generated responding on levers that had been previously paired with outcomes that shared the same motivational properties. The fact that S1 generated a larger general PIT effect than S2, and S2 generated an outcome-selective PIT but S1 did not, raises an obvious possibility: Outcome-selective PIT reflects the influence of stimulus–response associations acquired during conditioning, whereas general PIT reflects a genuine effect of stimulus-outcome associations on test performance.

3.8. Conclusion

To conclude, the results of Experiments 4–6 join a number of other paradoxical findings concerning the conditions under which outcome-selective PIT is observed. Taken together, these findings provide converging evidence that one analysis of this phenomenon, based upon a stimulus-outcome-response associative chain, should not be accepted uncritically: While general PIT might well reflect the kind of interaction between Pavlovian and instrumental processes envisaged by such a chain, outcome-selective PIT might have a quite different origin. I have argued that the existing behavioural evidence, including that from Experiments 4–6, suggests that an alternative view should be given serious consideration. According to this view, outcome-selective PIT can be a direct result of stimulus–response associations, the formation of which is mediated by the presentation of a shared outcome during training. The results from Experiment 7 provide direct support for the assumptions upon which this view is based. Whilst general PIT may result from a genuine interaction of Pavlovian and instrumental processes at test, it is clear that outcome-selective PIT may have a different origin. This phenomenon may result from the interaction between Pavlovian and instrumental associations over time, leading to new mediated associations. Chapter 4 will now explore another instance of interaction between these processes: the contextual control of instrumental performance.

Chapter 4

Contextual control of instrumental behaviour: The role of context-outcome and context-response associations

4.1. Summary

Chapter 4 explores another case where Pavlovian and instrumental processes interact: the contextual control of instrumental conditioning and extinction. This chapter addresses two fundamental issues regarding the renewal of instrumental behaviour: the role of direct Pavlovian associations in this renewal effect, and whether renewed responses are controlled by goal-directed processes. In particular, the source of ABA renewal of instrumental responding in rats was investigated. In Experiment 8, two responses (R1 and R2) were reinforced with one outcome (O1) in contexts A and B, and then R2 was extinguished in A and R1 was extinguished in B. At test, the rate of R1 was higher than R2 in A, and the reverse was the case in B: Renewed responding was independent of Pavlovian context-O1 associations. In Experiment 9, all rats received R1-O1 and R2-O2 in A; and then were placed in B where they were sated on O2, and either received concurrent extinction with R1 and R2 (group Extinction) or not (group No Extinction). At test, there was more responding in A than B in group Extinction, but not in group No Extinction; and renewed responding in A was as sensitive to the current value of the outcome as responding that had not been subject to the extinction (i.e., the rate of R1 was higher than R2): Renewed responding was goal-directed. These results identify contextual inhibition of the response, or its association with the outcome, as a basis for ABA renewal, and the response-outcome association as the source of renewed responding.

4.2. Introduction

The experimental context in which conditioning and extinction occur exerts a profound influence over conditioned behavior. One well-documented example of this influence is the ABA renewal effect, wherein a response that has been conditioned in one context (A), and then extinguished in another context (B), returns when assessed in context A. The overwhelming majority of examples of this effect involve Pavlovian conditioning, where the response in question is one that had been conditioned (and then extinguished) to a conditioned stimulus (CS) presented within the contexts (A and B; Bouton & Bolles, 1979; for a review, see Bouton, 2004). However, analogous effects have also been observed using instrumental conditioning procedures. For example, lever pressing established in context A, and extinguished in context B, is observed again when assessed in context A (e.g., Bouton et al., 2011; Nakajima, et al., 2000; Nakajima et al., 2002; Todd et al., 2012). Instrumental renewal effects have also been demonstrated where drugs of misuse serve as the reinforcer (e.g., Bossert, Liu, Lu, & Shaham, 2004; Crombag & Shaham, 2002; Hamlin et al.,). While the origin of Pavlovian renewal effects has been examined in some detail (e.g., Bouton & King, 1983; Bouton & Swartzentruber, 1986; Rescorla, 2008), the equivalent effect in instrumental conditioning has been the subject of less detailed analysis.

One obvious way in which contexts might exert control over instrumental responding is through the Pavlovian properties that they acquire when paired with reinforcement (during training) and nonreinforcement (during extinction). Consider the case of ABA renewal. The first stage of instrumental conditioning provides the necessary conditions for the context (e.g., A) to acquire an excitatory association with the outcome, and the second, extinction stage provides the conditions that should allow context B to gain an inhibitory association with the same outcome. These Pavlovian associations might well contribute to the renewal of instrumental responding. It is well established that Pavlovian CSs can modulate instrumental performance from studies of PIT (for a review, see Holmes et al., 2010). Although the origin of (outcome-selective) PIT is contentious (see Chapter 3), it is certainly the case that a context paired with an outcome can augment instrumental responding (Pearce & Hall, 1979; see also, Baker et al., 1991; Colwill & Rescorla, 1988; Rescorla & Solomon, 1967). The suggestion that instrumental renewal effects are a direct product of the *excitatory* Pavlovian properties of the training context is undermined, although not precluded, by several observations. The effect can be observed when testing occurs in a context (C) that one could argue is likely to be associatively neutral (in ABC renewal), and ABA renewal is relatively immune to extinction of context A before the test (see Bouton et al., 2011). The latter observations do not, however, exclude the possibility that the extinction context, B, might have gained inhibitory properties, the removal of which increases responding in context C. That is, context B might have come to inhibit the memory of the outcome that was absent when the response was not reinforced during extinction; and this inhibitory association affects the vigour of instrumental responding in context B but not context A (cf. Delamater et al., 2003). To rule out explanations of instrumental renewal based on the (inhibitory or excitatory) Pavlovian properties of the contexts, a procedure is required that equates these properties within a renewal design.

The principal aim of Experiment 8 was to assess whether a renewal effect could be observed under conditions in which the Pavlovian properties of the contexts (i.e., A and B) were equated. The design was modelled on one employed by Rescorla (2008; see also Harris, Jones, Bailey and Westbrook, 2000) to address the equivalent issue in the Pavlovian ABA renewal procedure. This design produced a renewal effect in the Pavlovian case, and if the same was true of instrumental case then it would constrain the range of explanations of instrumental renewal that remain in play. For example, in the case of the renewal of Pavlovian conditioning, contexts have often been described as acting in a hierarchical or occasion-setting fashion, on the link or links between the CS and the outcome or the absence of the outcome (see Bouton & Swartzentruber, 1986; Honey & Watt, 1999). Application of such a hierarchical analysis to the contextual control of instrumental responding assumes that the contexts are acting on the link between the response and the outcome (Rescorla, 1990a); and it is worth remembering that Skinner (1938) referred to discriminative stimuli as setting the occasion for instrumental responding. Set within an associative analysis of instrumental conditioning, ABA renewal could be based on the contexts gating either the excitatory response-outcome link by the test context A or the removal of the gate on the inhibitory response-outcome link provided by context B. The finding that a renewal effect can be observed when conditioning and extinction occur in context A and testing occurs in B is certainly consistent with the idea that the conditioning context gates the inhibitory response-outcome link (Bouton, Winterbauer and Todd, 2012).

There is already some recent evidence showing that renewed instrumental responding can be observed under conditions where the reinforcement histories of the contexts are equated; thus ruling out the possibility that differential context-outcome associations are the sole mechanism for renewal (Todd, in press; see also, Nakajima et al., 2002). Todd (in press) used a procedure where two responses (R1 and R2) were reinforced in separate contexts (A and B), prior to R2 being extinguished in A and R1 being extinguished in B. During tests conducted in both contexts, rats were more likely to perform R1 than R2 in context A, and were more likely to perform R2 than R1 in context B. Renewal was also observed when rats were tested in a novel context (C), thus suggesting that direct context-outcome associations are unlikely to be the source of renewed instrumental responding in the cases of ABA, AAB and ABC renewal. Even when renewed instrumental responding is observed under conditions in which the influence of the Pavlovian associations with the outcome could be excluded, this does not necessarily suggest that renewed responding originates in a response-outcome association. It is also possible that renewed instrumental responding reflects a direct excitatory effect of the conditioning context (A) on the response at test, or the release from an inhibitory effect of the extinction context (e.g., B) on the response. That is, renewed responding might be a product of excitatory or inhibitory context-response associations, rather than being dependent upon response-outcome associations. The source of renewed instrumental responding has not been investigated, and a secondary aim of Experiment 8 and the primary aim of Experiment 9 was to do this. The approach that I adopted was to examine whether the current value of the outcome affected renewed responding (Adams, 1982; Adams & Dickinson, 1981). If renewed responding is not affected by the value of the outcome then it would indicate that the contexts were exerting a direct effect on the response motor program, but if renewed responding is sensitive to the value of the outcome then it would implicate response-outcome associations as a source of renewed responding. This issue is not just of theoretical significance. It has been argued that renewal effects have relevance to our understanding of relapse effects observed following extinction treatments in, for example, drug-seeking behavior (Bouton, 2002; de Wit & Dickinson, 2009; Laborda, McConnell, & Miller, 2011). In the context of instrumental conditioning, which could be aligned to drug-seeking behavior, the nature of renewed responding has distinct relevance to its use as a model of relapse. Is renewed responding (relapse) based on a response-outcome association (and goal-directed) or is it independent of the outcome that engendered original instrumental performance (and habitual)? The main purpose of Experiment 9 was to assess whether or not renewed instrumental responding is goal-directed.

4.3. Experiment 8

The design of Experiment 8 is summarised in the upper panel of Table 7. Rats first received training where pressing the left and right levers (R1 and R2) resulted in the delivery of the same outcome (O1; e.g., a food pellet) in two contexts (A and B). After this, they were given access to one of the levers (supporting R1) in context B and to the other lever (supporting R2) in context A, and lever pressing was not reinforced. The levers were introduced and retracted repeatedly during these sessions. This procedure allowed us to assess lever pressing and magazine entries separately during the two extinction sessions, and to thereby establish the extent to which any decline in the rate of lever pressing was a product of an increase in the tendency of rats to enter the magazine (e.g., as a product of extinction). Aside from this, the use of this discrete-trial procedure, as with a free-operant procedure, allows both the response and the cues associated with that response (e.g., the lever, and its spatial location) to undergo extinction. Finally, rats received a test in which both levers were presented in either context A or B, and the levels of R1 and R2 were assessed. Renewal would be evident if the rate of R1 was higher than the rate of R2 in context A, and the reverse was the case in context B. That is, responding should be more evident when the context in which extinction of R1 and R2 was assessed was different from the context in which extinction occurred (R1 in context A and R2 in B) than when it was the same (R1 in B and R2 in A). It should be noted that the fact that both responses have been reinforced in contexts A and B, means that any renewal effect must reflect something that has been learnt during the extinction stage; as opposed to the retrieval of what was acquired during conditioning (cf. Todd, in press).

I also undertook a preliminary assessment of whether or not any renewed responding was sensitive to the current value of the outcome. To do so, rats were sated with either O1 or a novel outcome (O2) in a third context (C) prior to the tests in contexts A and B. If presentation of O1 resulted in a selective satiation effect, then renewed responding should be less evident after this treatment than after sating rats with O2. While recognizing that this manipulation has some limitations (notably, it confounds value with the familiarity of the outcomes; an issue that Experiment 9 avoids), Experiment 8 should still allow a renewal effect to be observed that is independent of the Pavlovian properties of contexts A and B: the primary goal of Experiment 8.

Table 7

Design of Experiments 8 and 9

Instrumental	Extinction	Satiation	Test
A: R1-O1, R2-O1	A: R2-	C: O1 *or* O2	A: R1 versus R2
&	*&*		*or*
B: R1-O1, R2-O1	B: R1-		B: R1 versus R2
Group Extinction			
A: R1-O1	B: O1…R1-/R2-		A: R1 versus R2
&			*or*
A: R2-O2			B: R1 versus R2
Group No Extinction			
A: R1-O1	B: O1…		A: R1 versus R2
&			*or*
A: R2-O2			B: R1 versus R2

Note: R1 and R2 refer to left and right lever presses; O1 and O2 denote food pellets and sucrose; - denotes no outcome; and A, B and C are contexts.

4.3.1. Method

Subjects, Apparatus and Procedure. The subjects were 16 naïve male Lister Hooded rats, maintained in the same way as in Experiments 1-7. 16 experimental chambers were used that were the same as those used in Experiments 1-7, with the exception that the walls and ceiling were lined with transparent Perspex behind which "wallpapers" were fixed. Four of the chambers had wallpapers consisting of black and white checks, four had white wallpaper with black spots, four had white wallpaper, and the remaining four had black wallpaper. Rats were first trained to retrieve food pellets and sucrose from the food well in two separate 30-min sessions in a chamber with no wallpaper. Half of the rats in each group were trained to collect sucrose in the first session and pellets in the second, and the remainder received the reverse arrangement. The outcomes were delivered on an RT 60-s schedule; each second there was a one in 60 chance that food would become available.

Rats then received two types of instrumental conditioning trials (i.e., R1-O1 and R2-O1) that occurred in separate, alternating sessions in both contexts A and B. For half of the rats, contexts A and B were boxes with spotted or squared wallpaper, and context C was either a box with black or white wallpaper; and for the remaining rats, contexts A and B were boxes with either black or white wallpaper, and context C was either a box with spotted or squared wallpaper. The identity of the box that serves as A or B, and the identity of the box that served as C was counterbalanced within the two sub-groups described above. In these sessions, one lever was inserted and pressing this lever (e.g., R1; pressing the left lever) resulted in the delivery of the designated outcome (O1; for half of the rats, O1 was food pellets and for the remainder it was sucrose). The rats first received four 15-min sessions in which each response (R1 and R2) was separately reinforced on a CRF schedule in each context in a counterbalanced order. Following CRF training, rats received a further 12 sessions of instrumental conditioning over six days, each lasting for 24 min. These sessions were conducted at the same time of day for each rat, and outcomes were delivered on an RI 30-s schedule with only one of the manipulanda present in each session. Rats received one R1-O1 session and one R2-O1 session per day. These sessions were conducted in context A and B on alternating days. The alternating sequence of R1-O1 and R2-O1 training was counterbalanced, as was the order in which rats were placed in contexts A and B. The rate of lever pressing (in responses per min, rpm) was used to assess instrumental conditioning.

Following instrumental training, on a single day, all rats received two 60-min extinction sessions, a satiation treatment and a single test session. During extinction sessions, the rats were placed in context A, where the lever corresponding to R2 was inserted and responses were not followed by reinforcement (food or sucrose); and they were placed in context B, where the lever corresponding to R1 was inserted and responses were not followed by reinforcement (food or sucrose). In these two sessions, the levers corresponding to R1 or R2 were inserted for a 10-s period prior to being retracted again. The lever was presented 120 times with a mean ITI of 30-s. The order in which these sessions occurred was counterbalanced and the sessions were separated by five minutes. Following these extinction treatments, rats were place in a novel context (C) where half were sated on O1 and the remainder were sated on a novel outcome, O2. Rats received 120 presentations of the outcome on an RT-30 schedule. Approximately five minutes after this session, the rats received a single 5-min test session in either context A or context B. Both levers were present and no outcomes were delivered.

4.3.2. Results

As expected, lever pressing increased from the first day of RI-30s training (mean = 4.37 rpm) to the final, sixth day of training (mean = 9.17 rpm; $F(5, 75) = 14.22, p < .001$, MSE = 4.82). The mean rates of lever pressing and magazine entries across the 12 blocks of 10 extinction trials are shown in Figure 18 (pooled across the various counterbalanced factors). ANOVA confirmed the impression that there was a marked reduction in both the rates of lever pressing across the extinction sessions, *F*(11, 165) = 30.89, *p* < .001, MSE = 2.94; and magazine entries, *F*(11, 165) = 15.92, *p* < .001, MSE = 3.33. The results of principal interest from the test in Experiment 8 are summarized in Figure 19, pooled across the various counterbalanced factors. Inspection of this figure shows that the rates of lever pressing on the levers that were tested in the same context as they had been extinguished (R2 in A and R1 in B) was lower than on the levers that were test in a different context than extinction had occurred (R1 in A and R2 in B, respectively). There was also a numerical tendency for the rate of lever pressing to be lower when rats had been sated with the training outcome, O1, than when they had been sated with the novel outcome, O2. ANOVA with test context (Same or Different) as a within-subjects factor and devaluation (O1/O2) as a between-subjects factor confirmed that there was an effect of context, *F*(1, 14) = 10.08, *p* < .01, MSE = 4.67, but there was no effect of devaluation, and no interaction between these factors, *F*s<1. There was a numerical tendency for the rate of magazine entries to be higher in rats sated with O2 than O1 during the first three minutes of the test (see Figure 20). However, ANOVA with devaluation (O1/O2) and minute as factors showed that there was no effect of devaluation, *F*(4, 56) = 1.46, *p* > .23, MSE = 19.04, trial, *F*(4, 56) = 1.03, *p* > .40, MSE = 19.04, and no interaction between these factors, *F*(1, 14) = 1.25, *p* > .28, MSE = 46.21.

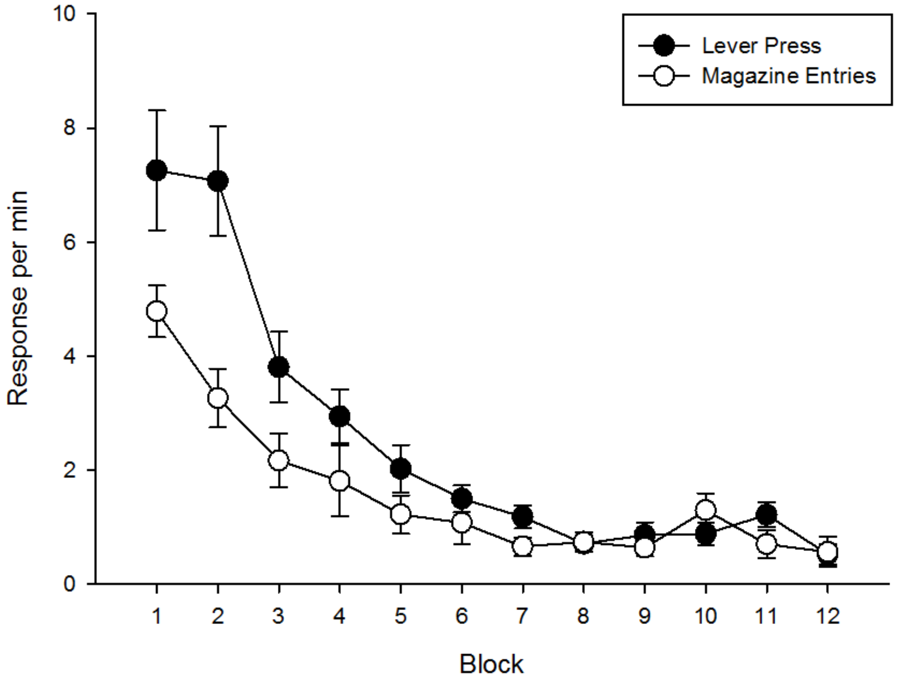


Figure 18. Experiment 8: Extinction. Mean rates (±SEM) of lever pressing (with R1 and R2 pooled) and magazine entries (during periods when the levers were retracted) in responses per minute, rpm, over the course of the extinction sessions.

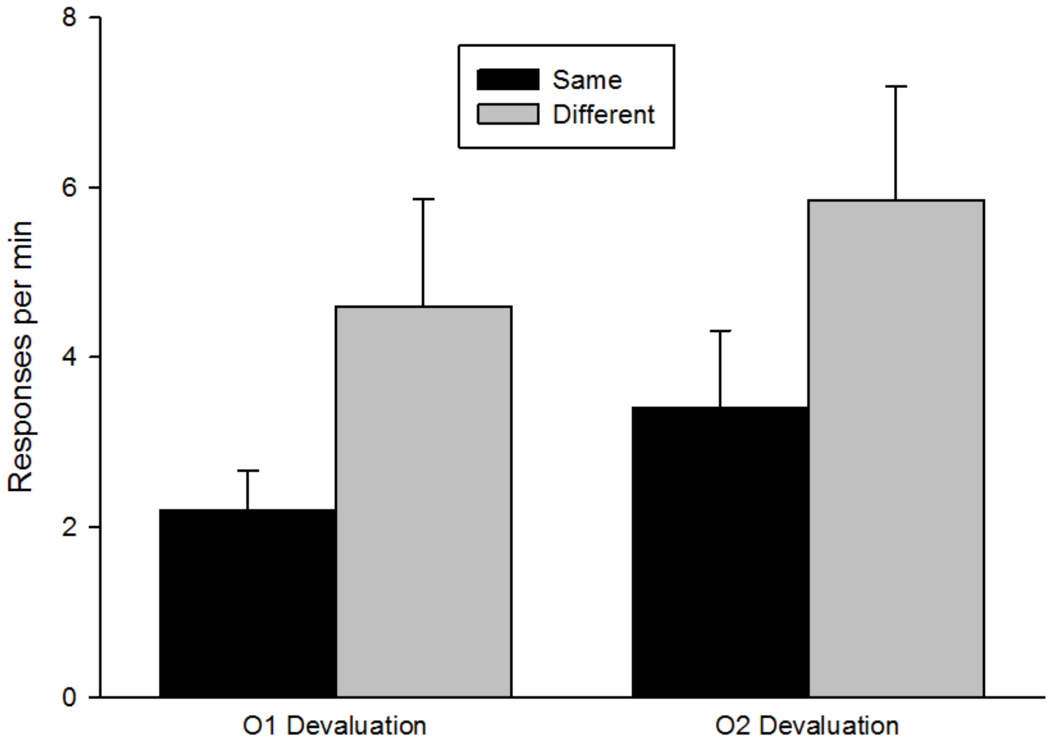


Figure 19. Experiment 8: Test. Mean rates of instrumental responding (in responses per minute, rpm; +SEM) as a function of whether responding was tested in the same context as extinction (i.e., R1 in B and R2 in A) or in a different context (i.e., R1 in A or R2 in B). Rats had either been sated on the outcome associated with R1 and R2 during training (O1) or with a novel outcome (O2).

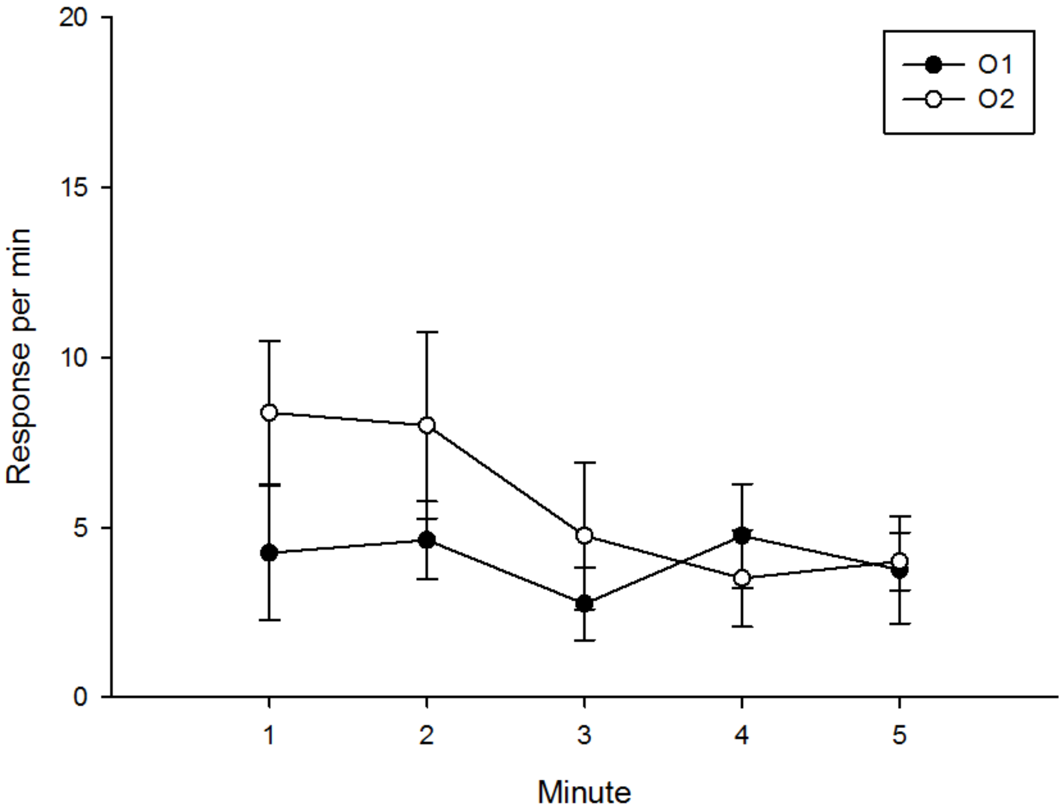


Figure 20. Experiment 8: Test. Mean rates of magazine entries (in responses per minute, rpm; ±SEM) over the 5-min extinction test, as a function of whether the rats had been sated on the outcome associated with R1 and R2 during training (O1) or with a novel outcome (O2).

4.3.3. Discussion

The results of Experiment 8 are theoretically noteworthy, because they establish that an instrumental renewal effect can be observed under circumstances in which the Pavlovian properties of the contexts (A and B) were equated; and serve to establish the generality of the results recently reported by Todd (in press). In this case, unlike in the experiments by Todd (in press), the effect was seen following relatively little extinction (120 extinction trials on each lever) that ended approximately one hour before the test. The contextual control exerted by the extinction context can clearly be acquired very rapidly.[3] The fact that this renewal effect was not influenced by whether rats were sated on the training outcome (O1) or a novel outcome (O2) prior to the test could be taken to suggest that renewed responding is insensitive to the current value of the outcome, and, in this sense, is not goal-directed and is instead habitual. This conclusion would be premature, however. There was a tendency for sating rats with the training outcome to reduced test responding relative to sating them with the novel outcome. Also, the use of an experimental design where the contexts were equated in terms of their Pavlovian properties required that they had the same relationship with a single outcome. This arrangement is not necessarily an optimal procedure for observing a selective satiation effect on instrumental performance. For example, the training procedure might have resulted in the rats not encoding the sensory properties of the outcome; and, as I have mentioned, the satiation procedure necessarily confounded training outcome identity with outcome familiarity. Moreover, the fact that there were two extinction sessions (one with each lever) and a satiation session might have meant that any effects of satiation were difficult to detect.

Given these considerations, in Experiment 9 I used a design in which the rats received training where two responses (R1 and R2) were paired with different outcomes (O1 and O2) in context A, and then received extinction and were sated on O1 in context B, prior to a test in either context A or B. This design closely resembles the instrumental renewal procedure used by Bouton et al. (2011), that was

---

[3] In both Experiment 8 and the experiments reported by Todd (in press) the extent to which extinction of the Pavlovian properties of either the lever or cues associated with the lever cannot be determined, and remain a potential source of the renewed responding that is observed.

itself modelled on many of those used in studies of renewal in Pavlovian conditioning. This design does not equate the Pavlovian properties of contexts A and B in the same way as Experiment 8. To do so, would have required the use of four responses to be used in conjunction with the two outcomes: with each member of a pair of two responses separately associated with the two outcomes in context A, and the remaining pair being associated with the two outcomes in context B. However, the inclusion of further controls should help to establish the generality of the principal conclusion from Experiment 8. Namely, that a renewal of instrumental responding can be observed that is not the consequence of the Pavlovian properties of the contexts.

4.4. Experiment 9

The design of Experiment 9 is summarized in the lower panel of Table 7. In this experiment two instrumental lever press responses (R1 and R2) were paired with the different outcomes (O1 and O2) in context A. After training, rats in group Extinction were placed in context B and received presentations of one of the training outcomes (O1) followed by the opportunity to respond on both levers (R1 and R2). Group No Extinction simply received presentations of one of the training outcomes (O1). The use of a design in which both responses were extinguished in the same session as the rats were sated on one of the outcomes (i.e., O1) allowed us to address some of the methodological concerns raised about Experiment 8. In particular, in Experiment 9 the rats were familiar with both outcomes, and combining the satiation procedure with the extinction procedure in a single session reduced the interval between both operations and the critical test. One potential problem with combining these procedures, however, is that it might result in a form of response reduction that was not the product of conventional extinction. For example, it might result in a reduction in responding that was due to contingency degradation. However, monitoring the rate of R1 and R2 during extinction, in addition to magazine entries, allowed one to assess such explanations of any reduction in responding. For example, a reduction in responding due to contingency degradation should be more apparent in the case of R1 than R2.

Rats then received a 5-min extinction test in either context A or B, with both levers present. For rats in group Extinction, responding overall should be more evident in context A than in context B, and if renewed responding is goal-directed then the level of R2 (which was paired with the outcome that remains valued) should exceed the level of R1 (which was paired with the now devalued outcome). For rats in group No Extinction, the tests in contexts A and B provide an assessment of the extent to which instrumental conditioning is context specific, and allows one to determine whether responding is goal-directed in the absence of any extinction treatment (cf. Adams, 1982; Adams & Dickinson, 1981). It also allows us to gauge whether contexts A and B are having their effects on instrumental responding by virtue of differences in the Pavlovian properties of contexts A and B. If the contexts are exerting an effect of instrumental responding through a difference in their associative strengths, then in group No Extinction the level of instrumental responding (R1 and R2) should be greater in context A than in context B.

4.4.1. Method

Subjects, Apparatus and Procedure. The 32 naïve male Lister Hooded rats were from the same supplier and maintained in the same way as in Experiments 1-8. The apparatus was the same as that used in Experiment 8, with the exception that only 8 chambers were used instead of 16. Half of these chambers were decorated with black and white checked wallpaper and the remainder were decorated with white wallpaper bearing black spots. Rats were first trained to retrieve food pellets and sucrose from the food well in two separate 30-min sessions in the same manner as Experiment 8. Rats then received 12 sessions of instrumental conditioning in context A over 6 days that were conducted at the same time of day for each rat. The two types of instrumental conditioning trial (i.e., R1-O1 and R2-O2) occurred in alternating sessions that were presented in a counterbalanced order. In these sessions, one lever was inserted and pressing this lever (e.g., R1; pressing the left lever) resulted in the delivery of the designated outcome (O1). In the first session of each response-outcome combination, reinforcement was delivered according to a CRF schedule, and on the remaining sessions an RI 30-s schedule was employed. The rate of lever pressing (in rpm) was used to assess instrumental conditioning. Other details of the training stage were identical to those of Experiment 8.

On the final day, rats were randomly assigned to two groups: Extinction and No Extinction. Rats in group Extinction were placed in context B and received 120 presentations of O1 (designed to sate them with this outcome) and 120 opportunities to respond on both levers. To be more specific, the delivery of O1 on an RT 30-s was followed, once the rat had removed its snout from the food well, by the insertion of both levers for 10-s. Response on either lever has no programmed consequences. After this 10-s period, the levers were withdrawn until the next trial. Rats in group No Extinction were also placed in context B and received exactly the same procedure with the exception that the levers were not inserted into the chamber. Approximately five minutes after these sessions, the rats received a single 5-min test session in either context A or context B. Both levers were present and no outcomes were delivered. Other details of the test were the same as in Experiment 8.

4.4.2. Results and discussion

As in Experiment 8, lever pressing increased from the first RI 30-s session (overall mean = 0.83 rpm; group Extinction mean = 0.80 rpm, and group No Extinction mean = 0.86 rpm) to the final, sixth day of training (overall mean = 7.61 rpm; group Extinction mean = 7.83 rpm, and group No Extinction mean = 7.39 rpm). ANOVA with group and day as factors revealed no effect of group, $F < 1$, an effect of day, $F(4, 150) = 67.99$, $p < .001$, MSE = 3.77, and no interaction between these two factors, $F < 1$. Inspection of Figure 21 shows there was a marked reduction in lever pressing across the extinction session for group Extinction, and that there was little difference in the rates of R1 and R2. This description of the results was confirmed by an ANOVA, that revealed that there was no significant main effect of response (R1 versus R2), $F<1$, a main effect of block, $F(11, 165) = 13.56$, $p < .001$, MSE = 1.80, and no interaction between these factors, $F<1$. The fact that the rates of R1 and R2 declined at the same rates suggests that the presentation of O1 was not influencing performance through an outcome-specific contingency degradation effect, which should have resulted in the rate of R1 being lower than the rate of R2 (e.g., Colwill & Rescorla, 1986). Inspection of Figure 22 shows that in both groups the rates of magazine entry also declined over the course of the extinction session (note that levers were retracted during this period for group Extinction). ANOVA confirmed that there was no significant main effect of group, $F < 1$, a main effect of block, $F(11, 165) = 9.31$, $p < .001$, MSE = 7.72, and no interaction between these factors, $F<1$.

The results of principal interest, from the test of Experiment 9, are summarized in Figure 23. This figure reveals that in group Extinction there was more responding in context A than in context B, and this renewed responding in A (and the extinguished responding in B) was sensitive to the current value of the outcome. That is, rats in group Extinction were more likely to perform R2 than R1 in both contexts A and B. In contrast, in group No Extinction the context in which testing occurred (A or B) had no effect on the level of responding (cf. Bouton et al., 2011), but sating rats on O1 resulted in them being more likely to perform R2 than R1. This pattern of results suggests that, using our procedures, rats readily identified the levers that support R1 and R2 in context B, and the ability of R1 and R2 to retrieve their respective outcomes, O1 and O2, was unaffected by being placed in context B. ANOVA with response (R1 or R2) as a within-subjects factor, and context (A or B) and group (Extinction or No Extinction) revealed an effect of response, $F(1, 28) = 6.01$, $p < .05$, MSE = 4.08, group, $F(1, 28) = 9.52$, $p < .005$, and context, $F(1, 28) = 6.04$, $p < .05$, MSE = 3.06. There was also an interaction between group and context, $F(1, 28) = 4.59$, $p < .05$, MSE = 3.06, but no other interactions, *F*s < 1. To analyse the nature of this interaction, separate ANOVAs were conducted on the results from group No Extinction and Extinction. The analysis for group Extinction revealed a main effect of context, $F(1, 14) = 11.22$, $p < .005$, MSE = 2.43, and response, $F(1, 14) = 5.46$, $p < .05$, MSE = 2.89, and no interaction between these factors, $F<1$. The equivalent analysis for group No Extinction revealed no effect of context, $F<1$, no significant effect of response, $F(1, 14) = 1.97$, $p > .18$, MSE = 5.73, and no interaction between these factors, $F<1$. Although the inclusion of response as a factor in these supplementary analyses is not licensed by the results of the omnibus analysis (in which there was an overall effect of response that did not interact with other factors), its inclusion certainly established the critical point that there is a significant effect of response (R1 or R2) in group Extinction: The analysis thereby confirms that responding observed after extinction is goal-directed, whether this responding is subject to the process of renewal (in context A) or not (in context B).

Inspection of Figure 24 shows the overall rates of magazine entries were somewhat lower in group Extinction (upper panel) than group No Extinction (lower panel). It also reveals that while the rate of magazine entries was higher in context A than context B in group Extinction, this was not the case in group No Extinction. ANOVA with minute, group and context as factors confirmed that there was an effect of group, $F(1, 28) = 11.23, p < .005$, MSE = 39.39, no effect of minute, $F(4, 112) = 1.77, p > .14$, MSE = 13.84, and no interaction between these factors, $F(4, 112) = 1.86, p > .12$, MSE = 13.84. Also, while there was no main effect of context, $F(1, 28) = 1.02, p > .32$, MSE = 39.39, or interaction between minute and context, $F(4, 112) = 1.18, p > .32$, MSE = 13.84, and no three-way interaction ($F < 1$), there was a group by context interaction $F(1, 28) = 6.73, p < .05$. Analysis of this interaction showed that while there was more responding in context A than in context B in group Extinction, $t(14) = 3.09, p < 0.01$, there was no difference in responding between contexts A and B in group No Extinction, $t(14) = -.97, p > .35$. This pattern of results mirrors the overall differences in lever press responding among the four groups: in the sense that in the rats given extinction there is more lever pressing in context A than B, whereas this is not the case in the rats that had not received extinction. The relationship between the overall levels of magazine entries and the overall levels of lever press is difficult to disentangle; but the nature of this relationship does not undermine the principle conclusions that one draws from Experiment 9. Namely, that renewed lever pressing is as sensitive to outcome devaluation as lever pressing that has not been subject to an extinction treatment.

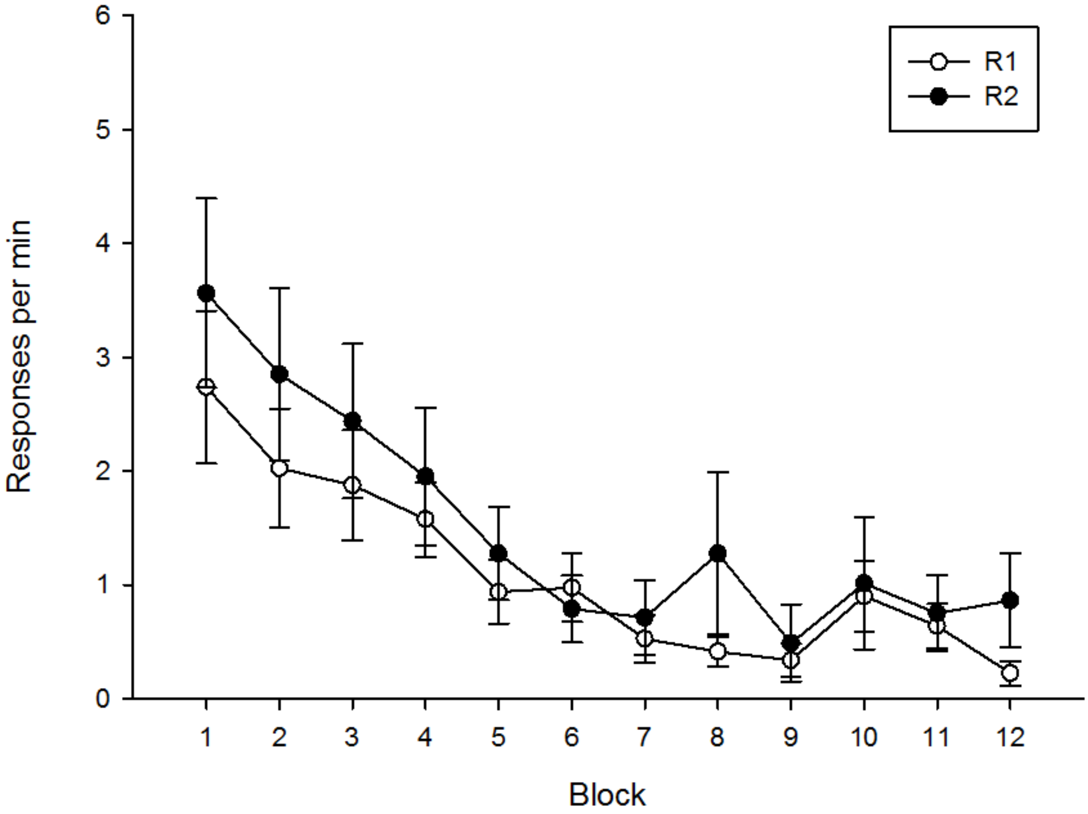


Figure 21. Experiment 9: Extinction of lever pressing. Mean rates of instrumental responding (in responses per minute, rpm; ±SEM) during extinction for group Extinction.

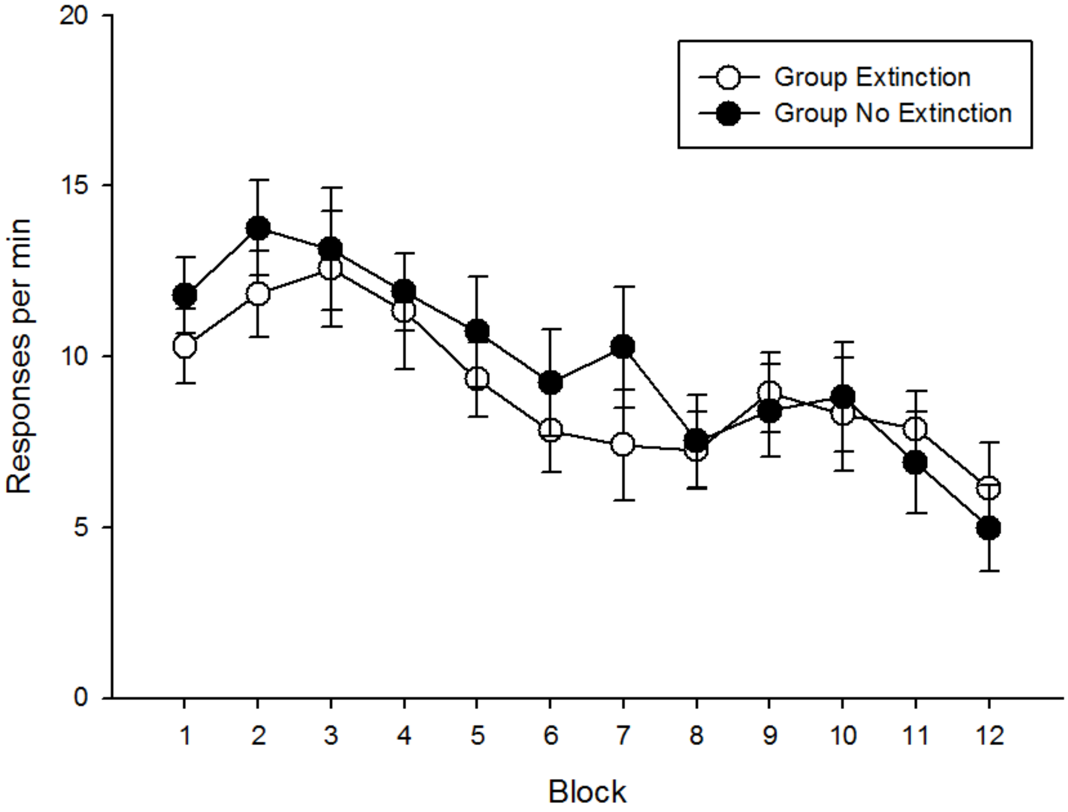


Figure 22. Experiment 9: Extinction of magazine entries. Mean rates magazine entry (in responses per minute, rpm; ±SEM) for group Extinction and group No Extinction. Rats in both groups had been sated with O2 prior to the test, and while rats in group Extinction had received instrumental extinction in context B, those in group No Extinction had not. Magazine entries for group Extinction are shown for the periods when the levers were retracted.

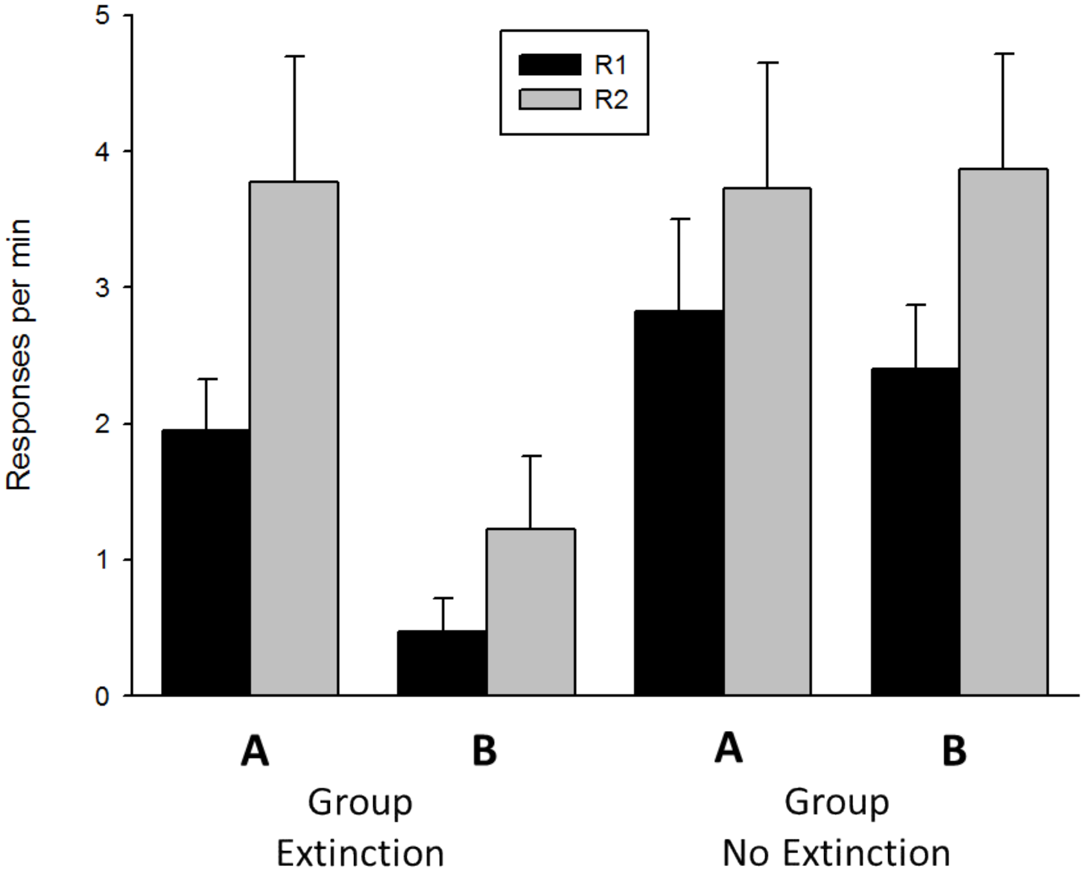


Figure 23. Experiment 9: Test. Mean rates of two responses (R1 and R2; in responses per minute, rpm; +SEM) as a function of whether testing was conducted in the conditioning context (A) or in the extinction context (B). Group Extinction had received extinction with both responses (R1 and R2) and were sated with the outcome associated with R1 (i.e., O1) in context B; and group No Extinction were simply sated with O1 in context B.

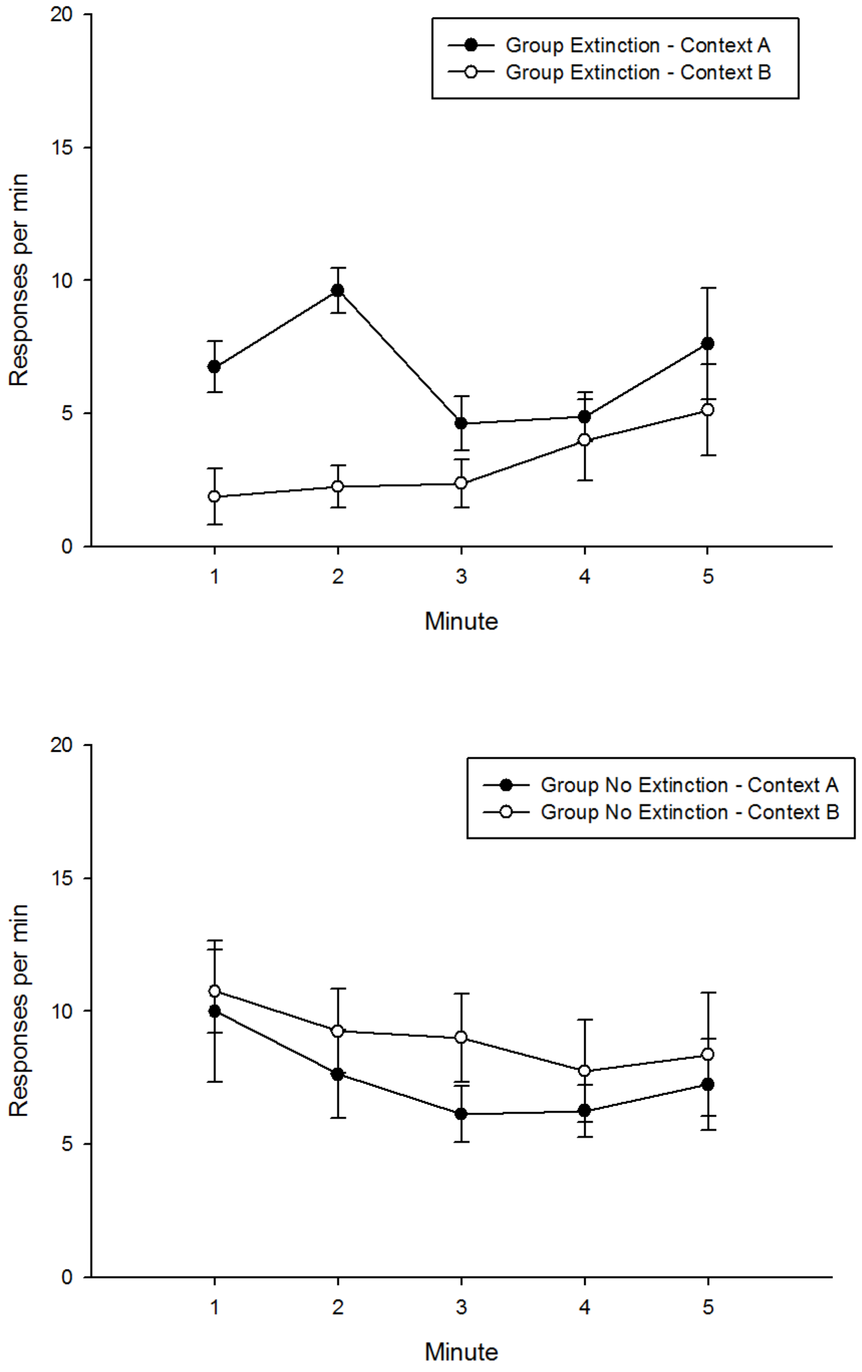


Figure 24. Experiment 9: Test. Mean rates of magazine entries (in responses per minute, rpm; ±SEM) over the 5-min extinction test, as a function of whether rats were tested in context A or B. Group Extinction (upper panel) had received extinction of both instrumental responses (R1 and R2) and were sated with the outcome associated with R1 (i.e., O1) in context B; and group No Extinction were simply sated with O1 in context B.

4.5. General discussion

The processes of conditioning and extinction are fundamental to any understanding of learnt behaviour (Hull, 1943; Spence, 1937). One of the key features of Pavlovian conditioning and extinction is the relative lability of extinction in the face of changes in the context in which it is assessed. For example, in Pavlovian conditioning procedures, a change in context that has little or no apparent effect (cf. Hall & Honey, 1990) on the level of responding generated by reinforced training, markedly attenuates the influence of extinction on performance (see Bouton, 2004). This feature of Pavlovian conditioning is also true of instrumental conditioning. For example, if rats receive instrumental conditioning in context A and are then given extinction in context B, placement in context A results in renewed instrumental responding (e.g., Bouton et al., 2011; Todd et al., 2012). Here, I investigated the source of renewed instrumental responding in two experiments with rats. Experiment 8 showed that renewed instrumental conditioning can be observed without concomitant differences in the excitatory or inhibitory Pavlovian properties of the two contexts, A and B; and Experiment 9 showed that renewed instrumental responding was as sensitive to the current value of the outcome as instrumental responding that has not been subject to an extended period of extinction. These observations begin to allow a more secure interpretation to be given to ABA renewal in an instrumental setting, which has implications for other domains of research (which will be discussed further in Chapter 5).

The fact that ABA instrumental renewal can be observed when the Pavlovian relationship between the contexts, A and B, and reinforcement and nonreinforcement have been equated, leaves open two obvious potential loci for the context in which extinction occurs to exert its influence on instrumental performance. One resides in the capacity for the context to directly affect the response program responsible for instrumental behavior. The extinction of instrumental behavior might result in the extinction context (i.e., B) coming to reduce the likelihood that the motor program for the extinguished response to become active. For example, it has been argued that a stimulus that has undergone extinction might come to inhibit the response that was extinguished in its presence (Rescorla, 1993); and the same might be true of a context in which an instrumental response has been extinguished. Application of this suggestion to the results of Experiment 8 is straightforward: After extinction training involving R2 in context A and R1 in context B, context A might come to inhibit R2 and context B might inhibit R1. Under these conditions, R1 would be released from this inhibitory influence when tested in context A and R2 would be released from this influence when assessed in context B. Of course, this analysis would still leave open the possibility that the current value of the outcome would interact with renewed responding (see Experiment 9) to the extent that instrumental behaviour was also controlled by response-outcome associations.

The second locus is less direct. According to this alternative the extinction context (e.g., B) comes to gate and augment the *inhibitory* R1-O1 association, and placing rats in context A removes this gating influence and allows the excitatory R1-O1 association to be more evident (Bouton et al., 2011). Again, the influence of this excitatory association will be constrained by the current value of O1. Several features of the results of Experiments 8 and 9 are consistent with this hierarchical control of inhibitory associations. In Experiment 8, any ability of contexts A and B to gate the excitatory links between R1 and O1 and R2 and O1 (e.g., Hall & Honey, 1989, 1990; Honey, Willis & Hall, 1990; for a modified configural analysis, see Honey & Watt, 1999) was necessarily equated, leaving the inhibitory R1-O1 and R2-

O1 links acquired during extinction as the sole target for hierarchical contextual control. Moreover, in Experiment 9 there was no effect of a change of context on instrumental conditioning that had not been subject to prolonged extinction (cf. Bouton et al., 2011). The fact that instrumental behaviour was equally evident in contexts A and B in rats that had not received extinction did not appear to reflect a ceiling effect, or some general insensitivity of my response measure: there was no effect of a change in context on R1 (whose outcome has been devalued) or on R2 (whose outcome was not devalued).

4.6. Conclusion

The results of Experiments 8 and 9 provide more and less direct evidence, respectively, that renewed instrumental conditioning can be observed without concomitant involvement of any excitatory or inhibitory associations between the contexts and the outcomes. I also found renewed responding is every bit as sensitive to the current value of the outcome as instrumental responding that has not been subject to an extended period of extinction. Taken together, these results suggest that the extinction context exerts an indirect or a direct inhibitory influence on the effectiveness of the instrumental response-outcome association, the removal of which allows the impact of the response-outcome association of performance to be revealed.

Chapter 5

General Discussion

5.1. Overall summary

The experiments described in this thesis explored the mechanisms by which Pavlovian and instrumental processes interact. The findings suggest that the way in which these processes interact in general and outcome-selective PIT have separate origins. Whilst it is likely that general PIT reflects a genuine interaction of the CS on the instrumental response at test, the interactions that provoke outcome-selective PIT can be the product of mediated S-R associations formed during conditioning. I also found that the Pavlovian properties of contexts in which instrumental responses are extinguished are more influential on instrumental responding than those in which conditioning occur, and exert an inhibitory influence on the effectiveness of the instrumental response-outcome associations. In this concluding chapter, the findings from Chapters 2-4 will be summarised briefly; and their theoretical analyses will then be considered. Chapter 5 will then focus on the implications of these findings in a wider context and future research directions will be presented.

5.2. Summary of results

5.2.1. The elusive nature of outcome-selective Pavlovian-to-instrumental transfer

The PIT phenomenon appears to be somewhat elusive (see Section 2.2., Chapter 2). It was argued that one way in which procedural differences might affect the likelihood of observing PIT is through their effect on the discriminability of the outcomes (see Holmes et al., 2010). Chapter 2 explored this possibility and found no evidence to support it. Briefly, changing the ways in which conditioning was arranged or pre-exposure to the outcomes was delivered (intermixed or blocked), has no effect on whether outcome-selective PIT was observed. In fact, in Experiments 1-3 there was no evidence of this effect, although I did observe general PIT under some conditions.

5.2.2. Pavlovian-to-instrumental transfer: paradoxical effects of the Pavlovian relationship explored

On the basis of the results reported in Chapter 2, I changed the procedures employed in Chapter 3 and demonstrated outcome-selective PIT (Experiment 4). Chapter 3 then proceeded to investigate the associative mechanisms of the outcome-selective PIT effect. Standard outcome-selective PIT procedures embed the presentations of outcomes within a Pavlovian CS, which allow for the formation of both forward stimulus-outcome and backward outcome-stimulus associations. When these forward and backward associations were conditioned separately, outcome-selective PIT occurred after backward pairings of the outcome and stimulus but not after forward pairings (Experiments 5 and 6). These results are consistent with the suggestion that backward outcome-stimulus trials should be especially likely to result in the formation of mediated stimulus-response associations: the presentation of the outcome should associatively activate the response, or its corresponding motor program, at the same time as the stimulus is presented. It was argued that the resulting stimulus-response association should directly generate outcome-selective PIT. The results of Experiment 7 provided direct support for this form of analysis. After pairing a response with an outcome, the outcome was later able to provoke the response during a stimulus with which it had a backward relation; and this arrangement engendered stimulus-response learning. Taken together, these results support the view that outcome-selective PIT can be a direct result of the stimulus-response associations, the formation of which is mediated by the presentation of a shared outcome during training.

It is interesting to note that there was a more sustained general PIT effect with Pavlovian CSs that had been trained in a forward relationship to the outcomes. This observation suggests that general PIT reflects a genuine interaction between Pavlovian and instrumental processes during the test.

5.2.3. Contextual control of instrumental behaviour: The role of context-outcome and context-response associations

Chapter 4 explored the nature of renewed instrumental responding that occurs when the context is changed after extinction. The renewal effect was found using a procedure in which the Pavlovian properties of the contexts (i.e., A and B) were equated (Experiment 8). These results indicated that renewed responding was independent of any excitatory or inhibitory Pavlovian context-outcome associations. In this procedure, there was some tendency for renewed responding to be sensitive to the current value of the outcome, but this effect was not statistically reliable. Using a more conventional ABA renewal procedure (Experiment 9), renewed responding was as sensitive to the current value of the outcome as instrumental responding that had not been subject to an extended period of extinction. This observation suggested that instrumental renewal is based upon response-outcome associations; or in other terms, has a goal-directed nature. Taken together, these results identify contextual inhibition of the response, or its association with the outcome, as a basis for ABA renewal, and the response-outcome association as the source of renewed responding.

5.3. Theoretical analyses of interaction

5.3.1. S-O-R: A two-process theory

Two-process theories predict the rate of an instrumental response will be modified by the presentation of a Pavlovian stimulus (Rescorla & Solomon, 1967). This is based on the assumption that the presentation of a Pavlovian CS paired with an outcome should enhance the expectancy that was created by the instrumental contingency and hence increase responding. The reliance on the motivational qualities of an outcome have long been challenged by the finding that a CS can selectively modulate responding associated with the same outcome (Corbit & Belleine, 2005; Delamater, 1995; 1996; Delamater & Oakeshott, 2007). However, such outcome-selective influences were, to an extent, reconciled by a development to this account, S-O-R theory. This theory suggests that test presentations of the CS activate a representation of the outcome via a bidirectional O-R/R-O association formed during training. At test, these associations promote the selection and initiation of responses that were trained with the same outcome (cf. Balleine & Ostlund, 2007). The findings from Chapter 3 challenge the assumptions on which this analysis of outcome-selective PIT is based, and are discussed at length within Chapter 3. Given the fact that the results presented in Chapter 3 found support for an alternative possible locus for the outcome-selective PIT effect, implications for outcome-selective and general PIT will be considered separately.

When general PIT is considered in isolation, there are several features of the results reported in Chapters 2 and 3 that are consistent with a two-process account of general PIT (Rescorla & Solomon, 1967). Chapter 2 found that when training phases were blocked, with the intention of reducing the distinctiveness of the outcome

(Blair et al., 2003), there was general PIT but no outcome-selective PIT. This suggests general PIT may be controlled by the motivational features of outcome representations, which is predicted by two-process theories. Moreover, using a Pavlovian conditioning procedure that should result in the stimulus activating strongly the motivational aspects of the outcome (i.e., forward conditioning) there was a marked general PIT effect, but no outcome-selective effect.

The S-O-R account is unable to explain the way in which contextual cues affect instrumental behaviours in the case of instrumental renewal (particularly in Experiment 8). According to this account, Pavlovian stimuli at test (in this case the context) would evoke a representation of the outcome, which would in turn elicit responding via the O-R association. The key feature of Experiment 8 is that renewal is observed when associations involving the context and the outcome are equated. Under these circumstances, the ability of the retrieved representation of the outcome to provoke responding will be determined by the strength of the binary O-R associations. In Experiment 8, even if these O-R associations had been undermined by extinguishing R1 and R2, there is no basis upon which the O1-R1 association should have been any more effective than the O1-R2 association. That is, there is no basis for an S-O-R analysis, to predict the instrumental renewal effect observed in Experiment 8.

5.3.2. Associative-Cybernetic model

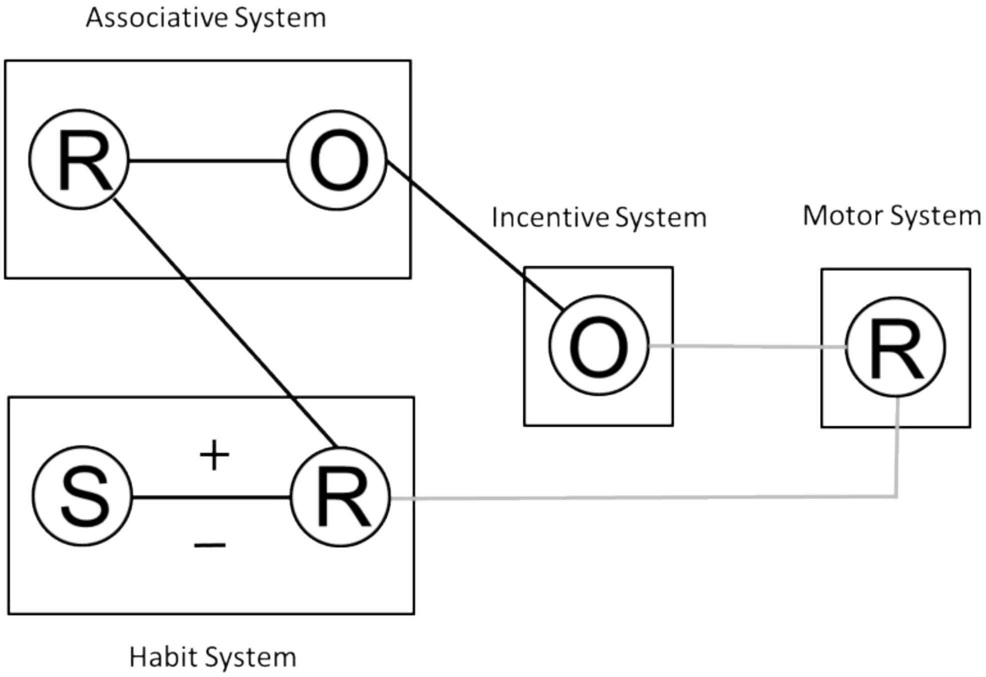


Figure 25. An Associative-Cybernetic model of instrumental responding (black lines represent learned associations; grey lines represent fixed connections). An extension to the current model would be required in order to account for excitatory (+) and inhibitory (-) S-R associations in the habit system.

The implications of the new findings presented in this thesis (specifically those from Chapters 3 and 4) for the Associative-Cybernetic (AC) Model (Dickinson & Balleine, 1993) will now be considered. The modified figure depicting the AC model (see Figure 25) is reproduced in order to aid presentation. According to this model, PIT can be generated in two ways: The CS could act on the outcome memory in the incentive system, that is linked to an instrumental response in the associative system, and this can generate responding through the corresponding motor program; or the CS could trigger the response representation (via a reciprocal O-R association in the associative memory) which then activates the response representation held in the habit system, which in turn activates the motor program. The AC analysis is challenged by the findings of Experiments 5 and 6, where outcome-selective PIT was observed following backward O-S but not forward S-O pairings. In the same way as the S-O-R account, the AC model relies on the outcome representation being triggered via the S-O association. There is no obvious reason why the S-O association should be more effective after backward conditioning than after forward conditioning, and therefore outcome-selective PIT should be less evident after backward conditioning than after forward conditioning. By the same token, the observation that general PIT appeared to be more evident after forward than backward conditioning is consistent with an AC model of general PIT.

The results presented in Chapter 4 cannot be readily explained by the AC model in its present form. Let us first take the finding that renewal can be observed without any concomitant differences in the excitatory or inhibitory Pavlovian properties of the two contexts (Experiment 8). As the associative strengths of the contexts and outcomes were equated, the renewal effect that was observed cannot be explained with the AC system wherein instrumental performance is modulated by the ability of the CSs (here the two contexts, A and B) to activate the outcome representations. In the case of Experiment 8, where a single outcome was used, there is no basis upon which R1 should be more evident than R2 in context A. In order to account for this finding, it would be necessary to extend the AC model to allow, for example, inhibitory S-R associations to form during extinction in one of the three systems. The observation, from Experiment 9, that renewal was sensitive to the value of the outcome, does not constrain the locus of the inhibition. However, at a theoretical level, the obvious location is within the habit system.

5.3.3. Hierarchical theory

As discussed in detail in Chapter 4, hierarchical analyses provide an alternative explanation for the contextual control of renewed instrumental responding to that provided by inhibitory S-R links. I will now consider a hierarchical analysis of PIT. This analysis explains the influence of a Pavlovian CS on instrumental responding via an S-(R-O) response-cueing process (Colwill & Rescorla, 1986; Rescorla, 1990a). As discussed in Chapter 1 (Section 1.4.3), the formation of a reciprocal R1-O1 association, will allow O1 to activate R1 during trials on which S1 is paired with O1. Under these conditions, S1 could be linked to the relationship between the evoked R1 and O1. In the same ways as a mediated S-R analysis predicts that backward conditioning might allow mediated S-R associations to form more readily than would forward conditioning, so too could a mediated hierarchical analysis. The evidence presented in this thesis provides no basis upon which to discriminate between the S-R and hierarchical analyses. However, on the grounds of parsimony one might prefer an analysis based on S-R associations. What is more, the hierarchical analysis provides no obvious account for the fact that outcome-selective PIT is insensitive to the current value of the outcome.

5.3.4. S-R theory

The mediated S-R theory of outcome-selective PIT proposed in this thesis is able to explain the findings reported in Chapter 3, and other features of the conditions under which the effect is observed (see Chapter 1). The merits of this account were discussed in detail in Chapter 3, and I have already described how an S-R analysis (in this case based upon inhibitory learning) provides one account for the renewal effect (see Chapter 3). I will now explore some of the implications of my analysis of outcome-selective PIT for studies that have used this behavioural phenomenon to explore the neural bases of instrumental and Pavlovian conditioning.

5.4. Implications for neuroscience: PIT and renewal

5.4.1. PIT

The dorsal striatum is involved in processes that are crucial for instrumental learning. The dorsolateral striatum (DLS) is involved in S-R learning (e.g., Yin, Knowlton, & Balleine, 2004), whilst the dorsomedial striatum (DMS) has been found to be necessary for some forms of R-O learning (Yin, Ostlund, Knowlton & Balleine, 2005). The role of both of these regions in the integration of Pavolvian and instrumental processes has been explored using standard PIT procedures. However, the interpretation of the results in such studies are based on the principle that an S-O-R analysis underlies the outcome-selective PIT effect. Corbit and Janak (2007) inactivated the DLS and the DMS separately following Pavlovian and instrumental conditioning, prior to assessing responding during a PIT test. They found inactivation of the DLS greatly attenuated both general and outcome-selective PIT, whilst the DMS abolished outcome-selective PIT. Interestingly, this pattern of results was not a consequence of a basic deficit in either Pavlovian or instrumental performance, because separate inactivation of either area failed to affect performance in control tests. Corbit and Janak argued that because S-O and R-O pairings were trained separately then there should have been no opportunity for S-R associations (between the Pavlovian stimuli and the lever response) to develop; and such S-R associations could not have been implicated in the results of the outcome-selective PIT test. On the basis of this argument, they concluded that the reduction in PIT following DLS inactivation must reflect a role for the region in some aspects of goal-directed performance. The results of Experiments 4-7 provide an alternative analysis that allows one to maintain the argument that the DLS is involved in habit (S-R) learning. These results also suggest that one should be cautious, in general, in interpreting outcome-selective PIT effects.

The DMS has been linked to the expression of R-O associations (Yin et al., 2005), and this link provides a natural interpretation for the abolition of outcome-selective PIT following inactivation of this region during the test (Corbit & Janak, 2007). However, there is another possible explanation for this observation. The anterior DMS, that was targeted in the Corbit and Janak (2007) study, has projections from the basolateral amygdala (BLA; see Kelley, Domesick & Nauta, 1982). The BLA has previously been demonstrated to be critical for outcome-selective PIT (Blundell et al., 2001; Corbit & Balleine, 2005). To the extent that lesions of the DMS disrupt the relay of information from the BLA to other structures, then the abolition of outcome-selective PIT might be expected on this basis alone; and does not necessarily implicate the DMS in the integration of Pavlovian and instrumental learning.

5.4.2. Renewal

Chapter 4 explored the source of the contextual control of renewed instrumental responding. In addition to exploring the locus of this mechanism, the results of Experiments 8 and 9 inform the neuroscientific domain. The prelimbic (PL) region of the medial prefrontal cortex has been directly implicated in ABA renewal. Thus, inactivation of the PL, at test, attenuates ABA renewal of extinguished alcohol-seeking behavior (e.g., lever pressing; Willcocks & McNally, 2012; see also, Hamlin et al., 2007; Zironi, Burattini, Aicardi, & Janak, 2006; for a review, see Van den Oever, Spijker, Smit, & De Vries, 2010). The results of

Experiments 8 and 9 established that, at least for the case of instrumental conditioning with food, the extinction context (B) modulates the impact of response-outcome associations on behavior by either directly inhibiting the response or by gating the inhibitory response-outcome association. That is, the results provide support for the view that there is a goal-directed component to renewed instrumental responding (a seeking component) and the context modulates this component. An attenuation of ABA renewal observed after inactivation of the PL, for example, might then reflect a disruption to a contextual inhibitory process or to the interaction between this process and the response-outcome association. However, identifying the specific roles of brain structures in the renewal of extinguished responding awaits the use of procedures that discriminate between the differing ways in which a renewal effect could be generated.

5.5. Future directions

Delamater et al. (2003) conducted a series of experiments with rats using backwards O-S pairings, which found the opposite pattern of results to those reported in Chapter 3 (Experiments 5 and 6). They suggested that the backward trace conditioning trials resulted in the development of inhibitory associations between the stimuli and outcomes. The stimuli would then inhibit their corresponding outcomes at test and thereby reduce the rate at which responses associated with these outcomes would be performed. This analysis is consistent with the S-O-R account. However, the results reported by Delamater et al. (2003) can be explained by the alterative S-R interpretation that I have advanced in this thesis and is supported by the results presented in Chapter 3. Thus, their use of a ten-second delay between presentations of the outcome and the stimulus during training might result in the evoked response representation decaying into a refractory state that supports the formation of a mediated inhibitory S-R association (cf. Wagner, 1981). This inhibitory S-R association should be directly able to reduce the likelihood of this response when the associated stimulus is presented at test.

Given the elusive nature of outcome-selective PIT, as a first step in investigating the analysis presented above one would need to replicate both patterns of results within a single within-subjects experiment (see Table 8). Following instrumental conditioning conducted in the same manner as Experiments 4-7, rats would receive Pavlovian conditioning using two types of backward conditioning trial. For one type, S1 would be delivered 1-sec after the rat had retrieved O1, and for the second type, the corresponding interval between S2 and O2 would be 10-sec. The influence of each stimulus on responding on R1 and R2 would then be assessed with a PIT test: S1 should be more likely to provoke R1 than R2 (cf. Experiment 5 and 6), and S2 should suppress the baseline levels of R2 relative to R1 (cf. Delamater et al., 2003; see Table 8). Given the fact that the predictions for the patterns of responding are the same during the two stimuli (albeit at different levels of performance), then another experiment would be required where a between-subjects procedure is used (where generalization between S1 and S2 is not an issue).

Table 8

Design of proposed Experiment (i; and idealized results)

Instrumental	Pavlovian	Test
R1-O1	O1- (1 sec) -S1	S1: R1 (10) > R2 (5)
R2-O2	O2- (10 sec) - S2	S2: R1 (5) > R2 (0)

Note: R1 and R2 refer to left and right lever presses; O1 and O2 denote food pellets and sucrose; and S1 and S2 are a tone and a light; 10, 5, and 0 indicate levels of instrumental responding, with 5 representing a notional baseline.

To assess my theoretical analysis of the conflicting patterns of results reported in Chapter 3 and those reported by Delamater et al. (2003) another experiment could be conducted of the form outlined in Table 9. Following instrumental training with two responses and outcomes (R1-O1 and R2-O2), rats would receive test sessions where each outcome is presented on separate trials. Following the delivery of one of the outcomes (i.e., O1), both levers would be inserted following a one-second delay, whilst for the other outcome (i.e. O2) the delay would be extended to ten seconds. If the S-R analysis described above is accurate, one would expect rats to respond on the lever associated with the same outcome (i.e. O1 would provoke responding on R1), after a brief delay (cf. Experiment 7). However, with a longer delay between the delivery of O2 and the opportunity to respond, the R2 representation might decay into a refractory state, and the rat should be less inclined to perform R2.

Table 9

Design of proposed experiment (ii; with idealized results)

Instrumental	Test
R1-O1	O1 - 1 sec delay: R1 (10) > R2 (5)
R2-O2	O2 - 10 sec delay: R1 (5) > R2 (0)

Note: R1 and R2 refer to left and right lever presses; O1 and O2 denote food pellets and sucrose; 10, 5, and 0 indicate levels of instrumental responding, with 5 again representing a notional baseline.

Several theoretically noteworthy findings were reported in Chapter 4, namely that instrumental renewal could be observed without any concomitant involvement of excitatory or inhibitory context-outcome associations, and that instrumental renewal has a goal-directed basis. One way in which these results could be strengthened, given the novel discrete-trial procedure used, would be to attempt to replicate these findings using more conventional methods. These could include extinction treatments where the levers are inserted for the duration of session, multiple extinction sessions across several days (cf. Bouton et al., 2011), and separate extinction and devaluation sessions.

While the findings of Chapter 4 demonstrate instrumental renewal has a goal-directed origin, it would be useful to know whether responding that is demonstrably habitual can be (i) renewed, and (ii) remains habitual (i.e., independent of the current value of its outcome). Studies have shown that extended training encourages a transition from goal-directed to habitual responding (e.g., Dickinson, Balleine, Watt, Gonzalez & Boakes, 1995). A procedure of the form employed in Experiment 9 could be replicated, and the duration of instrumental conditioning (either 5, 10 or 25 sessions) manipulated (see Table 10). The results from rats not given extinction would allow the development of habits to be assessed, while those from the rats given extinction would allow an assessment to be made of whether extinction has an effect on the goal-directedness of extinguished and renewed responding.

Table 10

Design of proposed Experiment (iii)

Instrumental	Extinction/Satiation	Test
(5, 10 or 25 sessions)		
A: R1→O1	B: O1..R1-/R2-	A: R1 versus R2
A: R2→O2	or	B: R1 versus R2
	B: O1...	

Note: R1 and R2 refer to left and right lever presses; O1 and O2 denote food pellets and sucrose; - denotes no outcome; A, B and C are contexts. Both satiation (with O1) and extinction (of R1 and R2) procedures would occur within the same session.

5.6. General implications

The results of the experiments presented in this thesis identified different mechanisms that might support the interactions between Pavlovian and instrumental processes. To develop the example discussed in the introduction, the findings presented in this thesis have applied relevance for instances where forms of human psychopathology involve conditioning - most notably drug dependence, anxiety disorders, and over-eating. In Chapter 1, I discussed how Pavlovian cues might prompt overeating in the absence of hunger, which could contribute to the rise in obesity seen in recent years (e.g. Boggiano et al., 2009). In particular, as high calorie foods are readily available, as are the cues that predict them (e.g., advertisements). PIT provides a demonstration of how conditioned stimuli can affect the way one responds, even when the stimuli were not (seemingly) conditioned during the performance of a response. In Chapter 3, I discussed how the ability of such cues to prompt responding might be rooted in mediated S-R associations formed during the time when conditioning occurs. In terms of overeating, this could be taken to suggest that food-cues may form direct associations with the response of eating, even when the cues do not occur during food consumption. This analysis places the origin of the effect in conditioned behavioural mechanisms as opposed to those that rely on higher cognitive processing. There is evidence to suggest that humans are less likely to be aware of external cues than internal sensations, such as hunger (Cohen, 2008; Vartanian, Herman, & Wansink, 2008). However, given the findings presented in Chapter 3, such external cues may have a more important role in overeating than is commonly believed. While some have already proposed strategies aimed at the extinction of food-related cues in the management of overeating (e.g. Sobik, Hutchison, & Craighead, 2005), there may be merit in considering the possibility that the influence of the cues on responding has an S-R nature. Whatever the precise direction of strategies designed to address overeating, greater success may be expected if more consideration is given to the influence of conditioned cues.

Another obvious clinical application that an increased understanding of extinction might have is in the domain of treatments used to eliminate distressing thoughts and behaviours (Bouton, 2002, 2011; Conklin & Tiffany, 2002; Craske, Kircanski, Zelikowsky, Mystkowski, Chowdhury, & Baker, 2008; Myers & Davis, 2002; Thomas, Larsen, & Ayres, 2003). The findings of this thesis reinforce the general view that the efficacy of such extinction treatments might be constrained by the fact that extinction is context specific – with relapse being the consequence of a failure of extinction to generalize to new environments. However, these results also suggest that these failures to generalize do not necessarily reflect any simple relationship between the extinction context (or the test context) and reinforcing events that have occurred in those contexts. Rather they suggest that the extinction context has a more specific role – either setting the occasion for the inhibitory relationship between responses (e.g., drug seeking) and outcomes (drug consumption) or directly inhibiting the response. Behavioural training regimes that enhance either of these processes are clear targets for future basic and applied research.

The results of Experiment 9 reveal that renewed responses are sensitive to the current value of the outcome. The applied significance of this finding is straightforward: for example, renewed drug seeking during relapse is likely to be directed towards the drug (or its effects) that motivated original drug-seeking behaviour prior to any extinction treatment. The implication of this view is that behavioural treatment strategies will stand a greater chance of success if they are combined with other medical treatments that decrease the motivational value of the drug (Volkow & Li, 2004).

5.7. General conclusions

There has been a longstanding interest in how Pavlovian cues interact with instrumental responding to guide behaviour. The main thrust of this thesis was to consider the mechanisms that are involved in the interaction of these learning processes, which are fundamental to both human and animal behaviour. Exploring these interactions using animal models provides an analytically tractable domain in which to gain a more complete understanding of associative processes. The experiments presented in this thesis demonstrated how a variety of associative structures can be implicated in rodents, from mediated S-R associations to hierarchical associations. Understanding the conditions under which such associative structures are acquired, in turn, allows us to begin to identify the circumstances under which cues in our everyday environment may bias the way humans respond. Identification of such biases affords the development of strategies that could mitigate maladaptive responding, such as overeating and drug-seeking.

References

Adams, C.D. (1982). Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, *34B*, 77–98.

Adams, C.D., & Dickinson, A. (1981). Instrumental responding following reinforcer devaluation. *Quarterly Journal of Experimental Psychology*, *33*, 109–121.

Allman, M.J., & Honey, R.C. (2006). Transfer of configural learning between the components of a preexposed stimulus compound: Implications for elemental and configural models of learning. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *32*, 307-313.

Baker, A.G., Steinwald, H., & Bouton, M.E. (1991). Contextual conditioning and reinstatement of extinguished instrumental responding. *Quarterly Journal of Experimental Psychology*, *43*, 199–218.

Balleine, B.W., & Ostlund, S.B. (2007). Still at the choice-point: action selection and initiation in instrumental conditioning. *Annals of the New York Academy of Sciences*, *1104*, 147-171.

Balleine, B.W., & O'Doherty, J.P. (2009). Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, *35,* 48-69.

Birch, L.L. (1991). Obesity and eating disorders: A developmental perspective. *Bulletin of the Psychonomic Society*, *29*, 265–272.

Blair, C.A.J., Blundell, P., Galtress, T., Hall, G., & Killcross, S. (2003). Discrimination between outcomes in instrumental learning: Effects of preexposure to the reinforcers. *Quarterly Journal of Experimental Psychology, 56B*, 253-265.

Blundell, P., Hall, G., & Killcross, S. (2001). Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. *Journal of Neuroscience, 21,* 9018–9026.

Boggiano, M.M., Dorsey, J.R., Thomas, J.M., & Murdaugh, D.L. (2009). The Pavlovian power of palatable food: lessons for weight-loss adherence from a new rodent model of cue-induced overeating. *International Journal of Obesity*, *33*, 693-701.

Bossert, J.M., Liu, S.Y., Lu, L., & Shaham, Y. (2004). A role of ventral tegmental area glutamate in contextual cue-induced relapse to heroin seeking. *Journal of Neuroscience*, *24*, 10726–10730.

Bouton, M.E. (2002). Context, Ambiguity, and Unlearning: Sources of Relapse after Behavioral Extinction. *Biological Psychology*, *52*, 976-986.

Bouton, M.E. (2004). Context and behavioral processes in extinction. *Learning and Memory*, *11*, 485–494.

Bouton, M.E. (2011). Learning and the persistence of appetite: extinction and the motivation to eat and overeat. *Physiology & Behavior*, *103*, 51–58.

Bouton, M.E., Winterbauer, N. E., & Todd, T.P. (2012). Relapse processes after the extinction of instrumental learning: Renewal, resurgence, and reacquisition. *Behavioural Processes*, *90*, 130-141.

Bouton, M.E., & Bolles, R.C. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation*, *10*, 445–466.

Bouton, M.E., & King, D.A. (1983). Contextual control of the extinction of conditioned fear: tests for the associative value of the context. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *9*, 248–265.

Bouton, M.E., & Swartzentruber, D. (1986). Analysis of the associative and occasion setting properties of contexts participating in a Pavlovian discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *12*, 333–350.

Bouton, M.E., Todd, T.P., Vurbic, D., & Winterbauer, N.E. (2011). Renewal after the extinction of free-operant behavior. *Learning & Behavior*, *39*, 57–67.

Brandon, S.E., & Wagner, A.R. (1991). Modulation of a discrete Pavlovian conditioned reflex by a putative emotive Pavlovian conditioned stimulus. *Journal of Experimental Psychology: Animal Behavior Processes, 17,* 299–311.

Cohen, D.A. (2008). Neurophysiological pathways to obesity: below awareness and beyond individual control. *Diabetes*. *57,* 1768–1773.

Colwill, R.M. & Rescorla, R.A. (1986). Associative structures in instrumental learning. In: Bower G.H., (Ed.). *The Psychology of Learning and Motivation*, *20*, 55-104. Academic Press, Orlando, FL.

Colwill, R.M., & Rescorla, R.A. (1988). Associations between the discriminative stimulus and the reinforcer in instrumental learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *14*, 155–164.

Conklin, C.A., & Tiffany, S.T. (2002). Applying extinction research and theory to cue-exposure addiction treatments. *Addiction*, *97*, 155–167.

Corbit, L.H., Muir, J.L., & Balleine, B.W. (2001). The role of the nucleus accumbens in instrumental conditioning: Evidence of a functional dissociation between accumbens core and shell. *Journal of Neuroscience, 21,* 3251–3260.

Corbit, L.H., & Balleine, B.W. (2005). Double dissociation of basolateral and central amygdala lesions on the general and outcome-specific forms of Pavlovian-instrumental transfer. *Journal of Neuroscience*, *25*, 962–970.

Corbit, L.H., & Janak, P.H. (2007). Inactivation of the lateral but not medial dorsal striatum eliminates the excitatory impact of Pavlovian stimuli on instrumental responding. *Journal of Neuroscience*. *27*, 13977–13981.

Craske, M.G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy*, *46*, 5–27.

Crombag, H.S., & Shaham, Y. (2002). Renewal of drug seeking by contextual cues after prolonged extinction in rats. *Behavioral Neuroscience*, *116*, 169–173.

Delamater, A.R. (1995). Outcome-selective effects of intertrial reinforcement in a Pavlovian appetitive conditioning paradigm with rats. *Animal Learning & Behavior*, *23*, 31–39.

Delamater, A.R. (1996). Effects of several extinction treatments upon the integrity of Pavlovian stimulus–outcome associations. *Animal Learning & Behavior*, *24*, 437–449.

Delamater, A.R., & Holland, P. C. (2008). The influence of CS-US interval on several different indices of learning in appetitive conditioning. *Journal of Experimental Psychology: Animal Behavior Processes, 34,* 202-222.

Delamater, A.R., LoLordo, V., & Sosa, W. (2003). Outcome-specific conditioned inhibition in Pavlovian backward conditioning. *Learning and Behaviour, 31,* 393-402.

Delamater, A.R., & Oakeshott, S. (2007). Learning about multiple attributes of reward in Pavlovian conditioning. *Annals of the New York Academy of Sciences*, *1104*, 1–20.

de Wit, S., & Dickinson, A. (2009). Associative theories of goal-directed behaviour: a case for animal–human translational models. *Psychological Research*, 73, 463–476.

Dickinson, A. (1997). Bolles's psychological sylllogism. In M. E. Bouton & M. S. Fanselow (Eds.), Learning, motivation, and cognition: The functional behaviorism of Robert C. Bolles (pp. 345–367). Washington DC: American Psychological Association.

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

Dickinson, A. (1985). Actions and habits — the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London*, *B. 308*, 67–78.

Dickinson, A. & Balleine, B. (1993). Actions and responses: The dual psychology of behaviour. In N. Elian, R. McCarthy, & B. Brewer (Eds),*Spatial* *representation: Problems in philosophy and psychology.* (pp. 277-293). Oxford: Oxford University Press.

Dickinson, A. & Balleine, B. (2002). The role of learning in the operation of motivational systems. In: *Stevens' handbook of experimental psychology*, *learning, motivation and emotion*, Vol. 3, 3rd ed. (Ed. R. Gallistel; pp. 497–533). John Wiley & Sons, Inc., New York.

Dickinson, A., Smith, J., & Mirenowicz, J. (2000). Dissociation of Pavlovian and instrumental incentive learning under dopamine antagonists. *Behavioral Neuroscience, 114,* 468–483.

Doya, K. (2008). Modulators of decision making. *Nature Neuroscience*, *11*, 410-416.

Estes, W.K. (1948). Discriminative conditioning II: Effects of a Pavlovian conditioned stimulus upon a subsequently established operant response. *Journal of Experimental Psychology*, *38*, 173-177.

Everitt, B.J., & Robbins, T.W., (2005). Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nature Neuroscience*, *8*, 1481–1489.

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*, 232-241.

Hall, G., & Honey, R.C. (1990). Context-specific conditioning in the conditioned-emotional-response procedure. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *16*, 271-278.

Hall, J., Parkinson, J.A., Connor, T.M., Dickinson, A., & Everitt, B.J. (2001). Involvement of the central nucleus of the amygdala and nucleus accumbens core in mediating Pavlovian influences on instrumental behavior. *European Journal of Neuroscience, 13,* 1984–1992.

Hamlin, A.S., Clemens, K.J., & McNally, G.P. (2008). Renewal of extinguished cocaine-seeking. *Neuroscience*, *151*, 659–670.

Hamlin, A.S., Newby, J., & McNally, G.P. (2007). The neural correlates and role of D1 dopamine receptors in renewal of extinguished alcohol-seeking. *Neuroscience*, *146*, 525–536.

Harris, J.A., Jones, M.L., Bailey, G.K., & Westbrook, R.F. (2000). Contextual control over conditioned responding in an extinction paradigm. *Journal of Experimental Psychology: Animal Behavior Processes, 26*, 174 -185.

Holland, P.C. (1981). Acquisition of representation-mediated conditioned food aversions. *Learning and Motivation*, *12*, 1-18.

Holland, P.C. (1983). Representation-mediated overshadowing and potentiation of conditioned aversions. *Journal of Experimental Psychology: Animal Behavior Processes*, *9*, 1-13.

Holland, P.C. (1989). Feature extinction enhances transfer of occasion setting. *Animal Learning & Behavior*, *17*, 269-279.

Holland, P.C. (2004). Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*, 104–117.

Holland, P.C., & Gallagher, M. (2003). Double dissociation of the effects of lesions of basolateral and central amygdala on CS-potentiated feeding and Pavlovian-instrumental transfer. *European Journal of Neuroscience, 17,* 1680–1694.

Holmes, N.M., Marchand, A.R., & Coutureau, E. (2010). Pavlovian to instrumental transfer: A neurobehavioural perspective. *Neuroscience and Behavioural Reviews*, *34*, 277-1295.

Honey, R.C., Bateson, P., & Horn, G. (1994). The role of stimulus comparison in perceptual learning: An investigation with the domestic chick. *The Quarterly Journal of Experimental Psychology, 47*, 83-103.

Honey, R.C., Close, J., Lin, T.E. (2010). Acquired distinctiveness and equivalence: A synthesis. In C.J. Mitchell & M.E. Le Pelley (Eds) *Attention and associative learning: From brain to behaviour* (pp. 159-186). Oxford: Oxford University Press.

Honey, R.C., Good, M., & Manser, K.L. (1998). Negative priming in associative learning: Evidence from a serial-habituation procedure. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *24*, 229-237.

Honey, R.C., & Watt, A. (1999). Acquired relational equivalence between contexts and features. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *25*, 324-333.

Honey, R.C., Willis, A., & Hall, G. (1990). Context specificity in pigeon autoshaping. *Learning and Motivation*, *21*, 125-136.

Hull, C.L. (1943). *Principles of behavior*: *An introduction to behavior theory.* New York: Appleton Century Crofts.

Iordanova, M.D., Good, M., & Honey, R.C. (2011). Retrieval-mediated learning involving episodes requires synaptic plasticity in the hippocampus. *Journal of Neuroscience*, *31*, 7156-7162.

Jansen, A., Theunissen, N., Slechten, K., Nederkoorn, C., Boon, B., Mulkens, S., & Roefs, A. (2003). Overweight children overeat after exposure to food cues. *Eating Behaviors, 4,* 197–209.

Johnson, A., Bannerman, D., Rawlins, N., Sprengel, R., Good, M. (2007). Targeted Deletion of the GluR-1 AMPA receptor in mice dissociates general and outcome-specific influences of appetitive rewards on learning. *Behavioural Neuroscience*, *121,* 1192-1202.

Kelley A.E., Domesick, V.B., Nauta W.J. (1982). The amygdaostriatal projection in the rat: An anatomical study by anterograde and retrograde tracing methods. *Neuroscience*, *7,* 615–630.

Konorski, J. (1967). Integrative activity of the brain: An interdisciplinary approach. Chicago: University of Chicago Press.

Kruse, J.M., Overmier, J.B., Konz, W.A., & Rokke, E. (1983). Pavlovian conditioned stimulus effects upon instrumental choice behaviour are reinforcer specific. *Learning & Motivation, 14,* 165-181.

Laborda, M.A., McConnell, B.L., & Miller, R.R. (2011). Behavioral techniques to reduce relapse after exposure therapy: Applications of studies of experimental extinction. In: Schachtman, T.R., Reilly, S. (Eds.), *Associative learning and conditioning theory*: *Human and non-human applications* (pp. 79-103). Oxford: Oxford University Press.

Lin, T.E., & Honey, R.C. (2010). Analysis of the content of configural representations: The role of associatively evoked and trace memories. *Journal of Experimental Psychology*: *Animal Behavior Processes*, *36,* 501-505.

Murray, J.E., Li, C., Palmatier, M.I., & Bevins, R.A., (2007). The interoceptive Pavlovian stimulus effects of caffeine. *Pharmacology, Biochemistry and Behaviour. 86*, 838–846.

Myers, K.M., & Davis, M. (2002). Behavioral and neural analysis of extinction. *Neuron*, *36*, 567–584.

Nakajima, S., Tanaka, S., Urushihara, K., & Imada, H. (2000). Renewal of extinguished lever-press responses upon return to the training context. *Learning and Motivation*, *31*, 416–431.

Nakajima, S., Urushihara, K., & Masaki, T. (2002). Renewal of operant performance formerly eliminated by omission or non-contingency training upon return to the acquisition context. *Learning and Motivation*, *33*, 510–525.

Ostlund, S.B., & Balleine. B.W. (2007). Selective reinstatement of instrumental performance depends on the discriminative stimulus properties of the mediating outcome. *Learning & Behavior*, *35*, 43-52.

Pearce, J.M. (1987). A model for stimulus generalization in Pavlovian conditioning. *Psychological Review, 94,* 61-75.

Pearce, J.M., & Hall, G. (1979). The influence of context–reinforcer associations on instrumental performance. *Animal Learning & Behavior*, *7*, 504–508.

Perks, S.M., & Clifton, P.G. (1997). Reinforcer revaluation and conditioned place preference. *Physiological Behavior*. *61*, 1–5.

Rescorla, R.A. (1990a). The role of information about the response-outcome relation in instrumental discrimination learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *16,* 262.

Rescorla, R.A. (1990b). Instrumental responses become associated with outcomes that differ in one feature. *Animal Learning & Behavior, 18,* 206–211.

Rescorla, R.A. (1992). Associations between an instrumental discriminative stimulus and multiple outcomes. *Journal of Experimental Psychology: Animal Behavior Processes*, *18,* 95.

Rescorla, R.A. (1993). Inhibitory associations between S and R in extinction. *Animal Learning & Behavior*, *21*, 327-336.

Rescorla, R.A. (1994). Control of instrumental performance by Pavlovian and instrumental stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *20*, 44-50.

Rescorla, R.A., (2000). Associative changes with a random CS-US relationship. *Quarterly Journal of Experimental Psychology, 53B*, 325–340.

Rescorla, R.A. (2008). Within-subject renewal in sign tracking. *Quarterly Journal of Experimental Psychology*, *61*, 1793-1802.

Rescorla, R.A., & Solomon, R.L. (1967). Two-process learning theory: Relationships between Pavlovian conditioning and instrumental learning. *Psychological Review*, *74*, 151–182.

Robinson, T.E., & Berridge, K.C. (2008). The incentive sensitization theory of addiction: some current issues. Philosophical Transactions of the Royal Society B: Biological Sciences, *363,* 3137-3146.

Rocha, A. & Kalivas, P.W. (2010). Role of the prefrontal cortex and nucleus accumbens in reinstating methampethamine seeking. *European Journal of Neuroscience*, *31*, 903–909.

Skinner, B.F. (1935). The generic nature of the concepts of stimulus and response. *The Journal of General Psychology*, *12,* 40-65.

Skinner, B.F. (1938). *The behavior of organisms*. NY: Appleton-Century-Crofts.

Sobik, L., Hutchison, K., & Craighead, L. (2005). Cue-elicited craving for food: a fresh approach to the study of binge eating. *Appetite*. *44,* 253–261.

Spence, K.W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, *44*, 430–444.

Symonds, M., & Hall, G. (1995). Perceptual learning in flavor aversion conditioning: Roles of stimulus comparison and latent inhibition of common stimulus elements. *Learning and Motivation*, *26,* 203-219.

Thomas, B.L., Larsen, N., & Ayres, J.J.B. (2003). Role of context similarity in ABA, ABC, and AAB renewal paradigms: Implications for theories of renewal and for treating human phobias. *Learning and Motivation*, *34*, 410–436.

Thorndike, E.L. (1911). Animal intelligence: Experimental studies. New York: Macmillan.

Todd, T.P. (2013). Mechanisms of renewal after the extinction of instrumental behavior. *Journal of Experimental Psychology: Animal Behavior Processes* (in press).

Todd, T.P., Winterbauer, N.E., & Bouton, M.E. (2012). Effects of amount of acquisition and contextual generalization on the renewal of instrumental behavior after extinction. *Learning & Behavior*, *40*, 145-157.

Tolman, E.C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55,* 189–208.

Trapold, M.A., & Overmier, J.B. (1972). The second learning process in instrumental training. In A. Black, & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 427–452). New York: Appleton-Century-Crofts.

Van den Oever, M.C., Spijker, S., Smit, A.B., & De Vries, T.J. (2010). Prefrontal cortex plasticity mechanisms in drug seeking and relapse. *Neuroscience and Biobehavioral Reviews*, *35*, 276–284.

Vartanian, L.R., Herman, C.P., Wansink, B. (2008). Are we aware of the external factors that influence our food intake? *Health Psychology*. *27,* 553–538.

Volkow, N.D., & Li, T.K. (2004). Drug addiction: the neurobiology of behaviour gone awry. *Nature Reviews Neuroscience*, *5*, 963-970.

Watt A., & Honey R.C. (1997). Combining CSs associated with the same or different USs. *Quarterly Journal of Experimental Psychology B. 50,* 350-67.

Wagner A.R. (1981). SOP: A model of automatic memory processing in animal behaviour. In N.E. Spear & R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5-47). Hillsdale, NJ: Erlbaum.

Wheeler, D.S., Sherwood, A., & Holland, P.C. (2008). Excitatory and inhibitory learning with absent stimuli. *Journal of Experimental Psychology: Animal Behavior Processes*, *34*, 247-255.

Willcocks, A. L., & McNally, G. P. (2013). The role of medial prefrontal cortex in extinction and reinstatement of alcohol seeking in rats. *European Journal of Neuroscience*, *37,* 259-268.

Yin H.H., Knowlton B.J., & Balleine B.W. (2004). Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. *European Journal of Neuroscience. 19*, 181–189.

Yin, H.H., Ostlund, S.B., Knowlton, B.J., & Balleine, B.W. (2005) The role of the dorsomedial striatum in instrumental conditioning. *European Journal of Neuroscience, 22*, 513–523.

Zironi, I., Burattini, C., Aicardi, G., & Janak, P.H. (2006). Context is a trigger for relapse to alcohol. *Behavioural Brain Research*, *167*, 150–155.

Zorawski, M., & Killcross, S., (2003). Glucocorticoid receptor agonist enhances pavlovian appetitive conditioning but disrupts outcome-specific associations. *Behavioural Neuroscience, 117*, 1453–1457