Pavlovian-Instrumental transfer: computational models and function

Doctoral candidate
Emilio Cartoni

Tutor
Simona Cabib

Supervisor
Prof. Gianluca Baldassarre

Reviewer
Prof. Mehdi Khamassi

Reviewer
Prof. Giovanni Pezzulo

Academic Year 2015–2016
# Contents

Abstract iii

Acknowledgments v

1 Introduction 1

2 Appetitive Pavlovian-Instrumental Transfer: A Review 4  
2.1 Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4  
2.2 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4  
2.3 The transfer paradigm . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8  
2.4 Behavioral results - variables influencing transfer . . . . . . . . . . . . . 11  
2.5 Neural substrates . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 20  
2.6 Interaction with drugs of abuse . . . . . . . . . . . . . . . . . . . . . . . 28  
2.7 Human transfer . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 32  
2.8 Theoretical aspects of transfer . . . . . . . . . . . . . . . . . . . . . . . . 36  
2.9 Conclusions and future directions . . . . . . . . . . . . . . . . . . . . . . 42

3 The three principles of action: a Pavlovian-instrumental transfer hypothesis 45  
3.1 Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 45  
3.2 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46  
3.3 PIT paradigm . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 49  
3.4 Explaining PIT: current issues . . . . . . . . . . . . . . . . . . . . . . . . 50  
3.5 Three principles of action: context, efficacy, utility . . . . . . . . . . . . . 55  
3.6 A Bayesian model of PIT . . . . . . . . . . . . . . . . . . . . . . . . . . . 58  
3.7 Conclusions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 62

4 The Relationship Between Specific PIT and Instrumental Reward Probability 66  
4.1 Abstract . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 66
Contents

4.2 Introduction ........................................... 67
4.3 Material & Methods ..................................... 69
4.4 Results ....................................................... 72
4.5 Discussion ................................................... 75

5  Pavlovian-Instrumental transfer: a hierarchical model 78
  5.1 Introduction ............................................. 78
  5.2 Methods .................................................... 83
  5.3 Results ..................................................... 84
  5.4 Conclusions .............................................. 89

6  Conclusions .................................................. 91
Abstract

Reward-related cues are an important part of our daily life as they often influence and guide our actions. This thesis focuses on one of the experimental paradigms used to study the effects of cues, the Pavlovian to Instrumental Transfer paradigm (PIT). In this paradigm, cues associated with rewards through Pavlovian conditioning alter motivation and choice of instrumental actions. During the last decade, the PIT effect - the influence of Pavlovian stimuli over instrumental actions - has been subdivided into two types: specific PIT and general PIT, each having its own neural substrates. Specific PIT happens when a conditioned stimulus (CS) associated with a reward enhances an instrumental response directed to the same reward. Under general PIT instead, the CS enhances a response directed to a different reward as well. While important progress has been made into identifying the neural substrates, the function of specific and general PIT and how they interact with instrumental responses, are still not clear. In the experimental paradigm that distinguishes specific and general PIT an effect of PIT inhibition has also been observed and is waiting for an explanation.

In this thesis we propose an hypothesis that links these three PIT effects (specific PIT, general PIT and PIT inhibition) to three aspects of action evaluation. These three aspects, which we call “principles of action” are: context, efficacy, and utility. In goal-directed behaviour, an agent has to evaluate if the context is suitable to accomplish the goal, the efficacy of his action in getting the goal and the utility of the goal itself: we suggest that each of the three PIT effects is related to one of these aspects of action evaluation. In particular, we link specific PIT with the estimation of efficacy, general PIT with the evaluation of utility and PIT inhibition with the adequacy of context.

We then provide a first computational model that exemplifies this hypothesis. The model is a Bayesian generative model with latent variables, based on a Bayesian understanding of conditioning that has been gaining grounds in the latest years. The underlying hypothesis is that animals learn hidden (latent) causes that jointly explain the co-occurrences of several observables (namely, sounds, levers, foods) – as opposed to learning simple associations between these events as more commonly assumed in the
Abstract

animal learning literature. In this scheme, PIT depends on Bayesian inference on the presence or absence of such hidden causes.

We have then tested one part of our hypothesis and its predictions in a human behavioral experiment. In particular, we investigated the hypothesis that cues associated to an outcome elicit specific PIT by rising the estimates of reward probability of actions associated to that same outcome. In other words, cues reduce the uncertainty on the efficacy of instrumental actions. We used a human PIT experimental paradigm to test the effects of two different instrumental contingencies: one group of participants had a 33% chance of being rewarded for each button press, while another had a 100% chance. The group trained with 33% reward probability showed a stronger PIT effect than the 100% group, in line with the hypothesis that Pavlovian cues linked to an outcome work by reducing the uncertainty of receiving it. However, contrary to our prediction, the 100% group also showed a significant specific PIT effect, highlighting additional factors that could contribute to specific PIT beyond the instrumental training contingency.

In the last chapter, we developed a second Bayesian computational model on transfer, to account for the above experimental results. Compared to the previous model, this second model explicitly models Pavlovian and instrumental conditioning into two different components, arranged in a hierarchical fashion. We posited that the key link between the two components is the prediction of food availability by the Pavlovian process, which is then used by the instrumental process to determine which instrumental context is active and subsequently determine the best course of action. The model correctly reproduces the qualitative pattern of the behavioral experiment, albeit it is so far limited to specific transfer only.
Acknowledgments

I would like to thank my tutor Simona Cabib and my supervisor, Gianluca Baldassarre who have made this thesis possible.
1 Introduction

The topic of this thesis is Pavlovian-Instrumental transfer: an interaction between Pavlovian conditioning and instrumental conditioning. Our aim is to gain, through computational modelling, a deeper understanding of the mechanisms of Pavlovian-Instrumental transfer and of its function in particular. The Pavlovian-Instrumental transfer (PIT or simply “transfer” hereafter) is usually measured as the ability of a stimulus associated with a reinforcer through Pavlovian conditioning, to potentiate the instrumental responding towards the same or another reinforcer. For example, hearing a sound associated with the delivery of food through Pavlovian conditioning can rise the frequency of an instrumental response such as lever pressing directed to obtain the same or another food.

In the last decade there has been an increasing number and range of studies on PIT. As a consequence of these studies, PIT has been divided into two types: specific and general transfer, each characterized by a different neural substrate. Specific transfer refers to the ability of Pavlovian stimuli to enhance specific actions associated with the same outcome as the conditioned stimulus, whereas general transfer refers to the ability of conditioned stimuli to enhance also actions paired with different outcomes. Both types of transfer, specific PIT and general PIT are mediated by the amygdala and nucleus accumbens; however, while specific PIT is mediated by basolateral amygdala and the shell of nucleus accumbens, general PIT is mediated by the central amygdala and the core of nucleus accumbens (Corbit & Balleine, 2005, 2011). Even though these two types of transfer have been “localized” in the brain, we still don’t know the neural mechanisms through which these structures give rise to the interaction between Pavlovian and instrumental conditioning. Besides the neural mechanisms, at the functional level things are not clear either. Why does transfer happen? What is its purpose? The explanations in terms of stimulus-outcome-response (S-O-R) chains, where the conditioned stimulus evokes the outcome representation (i.e. some food) and then this food evokes an in-

1Other areas beyond amygdala and nucleus accumbens are involved as well, although most data is on amygdala and nucleus accumbens, see also Chapter 2.
strumenal response are not fully satisfactory. For example, they do not explain why specific and general PIT tend to exclude each other (even though they are apparently parallel mechanisms in terms of structures) or why general PIT is not detected in some conditions.

The objective of this thesis is to give a functional explanation of PIT using computational models. In particular, we will use generative Bayesian models with latent hidden variables: we posit that subjects within a PIT experiment do not simply learn relationships and associations among observable stimuli (i.e. levers, sounds, foods) but they infer the existence of hidden causes which make these stimuli co-occur. The inference on the presence (or absence) of these hidden causes produces the PIT effects.

Through our computational modelling we will try to give a functional explanation of PIT which goes beyond current associationist explanations. From the computational models we derive new behavioral predictions to be empirically tested. As the brain activity may also be understood as probabilistic computing (Nessler et al., 2013; Kappel et al., 2014; Bill et al., 2015) some neural predictions might be derived in the future as well. During the thesis we will present a first model of transfer and an experiment to test its predictions. The results from the experiment are then used to develop a second model.

The importance of modelling PIT is due to its being on the edge of the Pavlovian and Instrumental conditioning, which are both fundamental learning processes. The structures that mediate PIT, amygdala and nucleus accumbens, are key structures in assigning values and in motivated behavior in general (Balleine & Killcross, 2006; Humphries & Prescott, 2010). Beyond the importance of this topic as basic research, this may also have implications in understanding pathological conditions. For example, PIT has been used to explain some aspects of addiction i.e. how the stimuli associated to drugs of abuse elicit drug-seeking (Robinson & Berridge, 2003; Belin et al., 2009). Indeed, it has been observed that drugs of abuse such ampethamine, cocaine or alchool may alter (e.g. potentiate) PIT (Wyvell & Berridge, 2001; LeBlanc et al., 2012; Corbit & Janak, 2007a). PIT has also been associated with relapse detoxified alcohol-dependent patients (Garbusow et al., 2016). Beyond addiction, PIT alterations have also been linked to other pathological conditions, such as schizophrenia and depression (Morris et al., 2014; Huys et al., 2016). A deeper understanding of PIT thus may also lead in the future to a better understanding of these conditions as well.
1 Introduction

Organization

This thesis comprises 1 background chapter, 3 main results chapters and a concluding chapter. The first three chapters have been published as articles in peer-reviewed journals. Chapter 2 is a review of the Pavlovian-Instrumental Transfer literature. Chapter 3 presents a general hypothesis on the function of transfer and a first computational model, along with its predictions. In Chapter 4 presents the results of a behavioral experiment with human participants, aimed at verifying one of the predictions of the model exposed in Chapter 3. Chapter 5 presents a new model refining the ideas exposed in Chapter 3 and accounting for the results of Chapter 4. Chapter 6 draws the conclusions.
2 Appetitive Pavlovian-Instrumental Transfer: A Review

In this chapter, we will present a systematic review of the literature on appetitive Pavlovian Instrumental Transfer.

This chapter has also appeared as an article: Cartoni et al. (2016).

2.1 Abstract

Reward-related cues are an important part of our daily life as they often influence and guide our actions. This paper reviews one of the experimental paradigms used to study the effects of cues, the Pavlovian to Instrumental Transfer paradigm. In this paradigm, cues associated with rewards through Pavlovian conditioning alter motivation and choice of instrumental actions. The first transfer experiments date back to the 1940’s, but only in the last decade has it been fully recognised that there are two types of transfer, specific and general. This paper presents a systematic review of both the neural substrates and the behavioral factors affecting both types of transfer. It also examines the recent application of the paradigm to study the effect of cues on human participants, both in normal and pathological conditions, and the interactions of transfer with drugs of abuse. Finally, the paper analyses the theoretical aspects of transfer to build an overall picture of the phenomenon, from early theories to recent hierarchical accounts.

2.2 Introduction

Predictive cues are an important part of our life that continuously influence and guide our actions. Hearing the sound of a horn makes us stop before we attempt to cross the street. Seeing an advertisement for fast food might make us hungry and lead us to seek out a specific type and source of food. In general, cues can both prompt us towards or stop us from engaging in a certain course of action. They can be adaptive (saving our
life in crossing the street) or maladaptive, leading to suboptimal choices, e.g. making us eat when we are not really hungry (Colagiuri & Lovibond, 2015). In extreme cases they can even play a part in pathologies such as in addiction, where drug associated cues produce craving and provoke relapse (Belin et al., 2009).

One particular paradigm used to study the effect of such cues is the Pavlovian to Instrumental Transfer paradigm. In this paradigm, Pavlovian predictions and instrumental actions are first trained in separate experimental phases. The instrumental actions are then tested in both the presence and the absence of the Pavlovian cues to assess the effect of the latter on the former.

GLOSSARY

- **Pavlovian conditioning** - During Pavlovian conditioning a neutral stimulus, such as a sound, becomes a conditioned stimulus (CS) by pairing its occurrence with an unconditioned stimulus (US) that naturally elicits some response. For example, a sound (CS) might be paired with food (US) by delivering food only when the sound is present. At the end of training, the animal/participant will have learned that the CS predicts the US and so it will approach the site of food delivery when it hears the sounds.

- **Instrumental conditioning** - During instrumental conditioning an animal/participant is trained to make a response by delivering an attractive outcome. For example, a hungry rat might be trained to press a lever that delivers food. This training can lead to two kinds of instrumental behavior: habits, controlled by antecedent stimuli through the formation of stimulus-response (S-R) associations or goal-directed actions, controlled by the consequences of the action through the formation of action-outcome (A-O) associations.

- **Devaluation** - Outcome devaluation is a procedure where the US or the outcome (O) value is altered. For example, the value of a certain food might be altered by feeding it to satiation or by pairing it with illness (the latter induced by lithium-cloride injections).

- **Extinction** - A training session where the US predicted by a CS or the outcome (O) predicted by an action is no longer delivered, thus promoting the extinction of the Pavlovian conditioned response or the instrumental action.

The first Pavlovian to instrumental transfer studies date back to the 1940’s, reporting that stimuli paired with food were able to augment instrumental actions directed
towards food (Estes & Skinner, 1941; Walker, 1942; Estes, 1943). Transfer effects can either promote or discourage actions, with the presence of cues increasing/decreasing the frequency of an action or biasing choice in favour of certain actions. Amongst other factors, the type of effect obtained depends on the valence of the Pavlovian US, i.e., whether it is appetitive or aversive. For example, a Pavlovian cue associated with an aversive shock might promote actions leading to shock avoidance but decrease actions leading to food (Rescorla & Solomon, 1967).

Our understanding of transfer has naturally developed with our understanding of Pavlovian and instrumental conditioning. At the same time, studying the interaction of Pavlovian and instrumental conditioning has often yielded new insights into these individual processes. At the time of the earliest studies, for example, it was not clear if Pavlovian and instrumental conditioning constituted different forms of learning. Gradually, however, two-process theories emerged that separated Pavlovian and instrumental processes (see Rescorla & Solomon 1967 for a review). Transfer effects were, at that stage, understood as the result of Pavlovian cues generating general appetitive or aversive emotional states and, indeed, the transfer paradigm was typically used to study the influence of conditioned emotional responses (Rescorla & Solomon, 1967). Subsequent studies refined this general emotional state finding that, in many conditions, transfer was better characterised as controlled by primary motivational processes than emotional states. So, for example, Dickinson and colleagues demonstrated in studies of the so called irrelevant incentive effect that a cue predicting sugar solution would enhance instrumental actions both when rats were hungry and when they were thirsty whereas a cue associated with dry food pellets would only elevate performance when hungry. These effects were generally interpreted as suggesting that primary motivational processes could modulate the production of conditioned emotional states, much as suggested by Bindra (Dickinson & Dawson, 1987; Balleine, 1994; Bindra, 1974). However, none of these accounts explained the influences of Pavlovian cues on choice: i.e., how Pavlovian cues could sometimes be found to enhance actions tied to a specific outcome, e.g., how a CS associated with grain pellets enhanced lever pressing for grain pellets but not for other food outcomes (such as sugar). One possibility is that both Pavlovian and instrumental conditioning lead to the formation of associations: stimulus-outcome associations (S-O) in one case and response-outcome associations in the other case (R-O), and that the common outcome mediates the interaction (Trapold & Overmier, 1972). In the 1980’s and ’90’s much experimental work was devoted to establishing that instrumental conditioning could be subdivided into two types: habitual actions, controlled by stimulus-response (S-R) asso-
ciations, and goal-directed actions, controlled by response-outcome (R-O) associations (Balleine & Dickinson, 1998). In parallel, a series of articles examined the ability of Pavlovian cues tied to a specific outcome to bias choice between specific actions (Lovibond, 1981; Colwill & Rescorla, 1988; Colwill & Motzkin, 1994; Rescorla, 1991, 1994a,b; Delamater, 1995, 1996). At the beginning of the century, investigation of the neural substrates of transfer began: some initially contrasting results led at that stage to the realisation that transfer effects come in two different forms and so had to be subdivided as well. These studies divided the phenomenon into specific and general transfer, each characterized by a different neural substrate (Corbit & Balleine, 2005, 2011, see Section 2.5). Specific transfer refers to the ability of cues to enhance specific actions associated with the same outcome as the cue, whereas general transfer refers to the ability of cues to enhance also actions paired with different outcomes.

Most data on transfer come from animal studies, however in recent years the transfer paradigm has also been adapted for human participants. In general, human studies have produced similar results to animal studies both in terms of the behavioral factors and the neural substrates (e.g., Bray et al. 2008; Talmi et al. 2008; Watson et al. 2014, see Section 2.7) controlling transfer effects.

2.2.1 Scope and purpose of the review

In the last decade there has been an increasing number and range of studies on transfer, both in animals and humans, examining transfer under both normal and pathological conditions (Corbit et al., 2007; Corbit & Balleine, 2011; Laurent et al., 2014b; Ostlund et al., 2014b; Nadler et al., 2011; Morris et al., 2014; Colagiuri & Lovibond, 2015; Garbusow et al., 2016). Although there has been a relatively recent review on the topic (Holmes et al., 2010) the literature on transfer has essentially doubled in size over the last 5 years providing considerable new information on models and the neural bases of the transfer effect across species. In particular, we will focus on appetitive transfer which is the subject of the large majority of these recent studies. Figure 2.1 summarises our coverage of research and shows some landmarks in the investigation of the transfer effect. In our review we included only articles which follow the standard transfer paradigms: i.e. where the Pavlovian conditioning and instrumental conditioning are conducted in separate sessions.

---

1 Studies mostly involved rats, however other species have been used as well, such as mice (see Lederle et al. 2011 in different mice strains), monkeys (Stebbins & Smith, 1964), dogs (Rescorla & LoLordo, 1965), pigeons (Morse & Skinner, 1958) and even horses (Lansade et al., 2013).
The review is structured as follows: Section 3.3 describes the transfer paradigm and its variants; Section 2.4 then reviews behavioral factors influencing the transfer effect; Section 2.5 describes the neural mechanisms underlying transfer; Section 2.6 reviews studies of the interaction between transfer and drugs of abuse; Section 2.7 reviews transfer studies with human participants; Section 2.8 reviews theoretical aspects of transfer; Section 3.7 then draws overall conclusions. Where possible we will separate experimental findings from their theoretical interpretation using procedural rather than theoretical descriptions or definitions. In this regard, we clarify here our use of what may be seen as a theoretically laden term ‘action’. We use ‘instrumental action’ or just ‘action’ to refer to measures in the instrumental conditioning phase of transfer studies, whether the experiments were designed to encourage the development of goal-directed instrumental actions or of habitual instrumental actions. Measures taken from Pavlovian conditioning instead will always be referred to as ‘Pavlovian conditioned responses’, ‘conditioned responses’ or simply as ‘responses’.

2.3 The transfer paradigm

There are many variations of the transfer paradigm but it is always composed of three phases: Pavlovian training, instrumental training and the transfer test. The two training phases can be conducted in any order (either Pavlovian or instrumental first) with no change in the effect (but see Holmes et al. 2010, where lengthening the first or second phase varied the amount of transfer). In the Pavlovian phase one or more stimuli (usually auditory cues) are paired with the delivery of rewards such as food pellets or sucrose (see Figure 2.2). Pairing stimuli with an aversive event to develop an aversive transfer paradigm has also been conducted (e.g., Lewis et al., 2013; Rigoli et al., 2012; Campese et al., 2013).

During the instrumental training, a contingency is established between one or more actions and the delivery of one or more outcomes - usually involving pairing lever pressing with food delivery. Using one or more actions (e.g. one or two levers paired with different foods) leads to critical differences in what is measured in the final test (see below).

In the last phase, the animal/participant can again perform the instrumental actions, but this time the conditioned stimuli (CS) trained in the Pavlovian phase are presented during the session. The effect of presenting the CS on the instrumental response (the transfer effect) is then assessed by comparing instrumental responding during periods when no CS is presented (baseline) with periods when a CS is presented, or, if two CS’s
Figure 2.1: Articles referenced in this review, grouped in 5 years bins. Letters indicate landmark events in transfer research. The last decade has seen a marked increase in the investigation of transfer. (a) First experiments in the 1940’s. (b) Rescorla & Solomon (1967) review theories of Pavlovian and instrumental conditioning and their interaction, advocating a two-process theory. (c) During the 1980’s and 90’s instrumental conditioning was clearly subdivided into habitual (S-R) and goal-directed (A-O) actions. Around 1990, Rescorla and colleagues produce a series of studies examining how Pavlovian cues affect instrumental responding in an outcome-specific manner. (d) At the turn of the century investigation of the neural substrates of transfer begins, leading to the subdivision of general and specific transfer each with its separate neural substrate. (e) The transfer paradigm is adapted to human participants and neural substrates are investigated using fMRI.

are presented, by comparing responding during the presentation of the different CSs (e.g., one paired with the same food as the lever and another one paired with a different food). The CSs are never presented in the presence of the instrumental manipulanda before the test, so no explicit training of a relationship between the CS and the instrumental action takes place. When multiple actions are used, the test may involve a choice between two levers presented at the same time (e.g., Ostlund & Balleine, 2008) or separate tests of each instrumental action (e.g., Corbit et al., 2001). Test sessions are usually conducted in extinction, i.e., no outcomes are delivered either after the stimuli or after the actions, to avoid changes in performance due to new learning.

In all cases, what is generally found is that a CS paired with an appetitive outcome
(CS+) enhances instrumental responding compared to an unpaired CS (CS-). Usually the CS+ also increases instrumental responses compared to the baseline (CS-free period), however, in some cases, differences between the CS+ and CS- have emerged but with no difference between the CS+ and baseline, producing instead a reduction of CS-induced response suppression. Furthermore, in the two action case, although typically the ‘same’ CS elevates performance of the action delivering the outcome predicted by the CS relative to both the other action and to baseline performance, it has sometimes been found that both CSs reduce lever pressing compared to the baseline but that the CS sharing the same outcome as the instrumental action reduces performance less than a CS associated with a different outcome. This reduction with respect to the baseline could be due to response competition between instrumental and Pavlovian responses. For example, if the CS prompts considerable magazine approach, it will reduce the time spent pressing the lever (see also next section).

As we mentioned in the introduction, there are two kinds of transfer: specific transfer and general transfer. In specific transfer, the CS enhances actions associated with the same outcome as that paired with the CS whereas, in general transfer, a CS can enhance actions directed to other outcomes as well. Studies using a single lever paired with food usually also use one single paired CS in the Pavlovian phase and in the test phase cannot behaviorally distinguish between the specific and general transfer effects. This is because both general and specific transfer effects enhance instrumental performance, something that could be because the action shares the same outcome as the CS or through a general effect of the cue (e.g. motivational). For simplicity, throughout this review we will call these studies “single lever studies” (see Figure 2.2). It must be noted that while it is not possible to behaviorally distinguish between specific and general transfer in a “single lever study”, lesions experiments suggest that using a single lever usually elicits only general transfer. As we will see in section 2.5, studies investigating the neural basis of transfer have found distinct neural substrates for specific and general transfer (Corbit & Balleine, 2005, 2011), and transfer in “single lever” studies is impaired by lesions targeting general transfer substrates and spared by those targeting specific transfer substrates (Hall et al., 2001; Holland & Gallagher, 2003). It is not known why the “single lever” studies elicit general transfer and not specific transfer (as multiple lever studies do). It has been suggested that the different type of transfer elicited by single and double lever procedures might be caused by the more or less detailed representation of the outcome (Holland, 2004). Procedures with multiple levers (and reinforcers) favour the creation of a more detailed and sensory-specific representation of the reinforcers.
used, which in turn may lead to the transfer effect being specific. A single-lever lever procedure does not need a detailed representation of the outcome, so in this case the transfer effect might be conveyed by the more "general" appetitive characteristics of the outcome.

Another category of studies uses two levers, each paired with different outcomes (e.g. food pellet vs. sucrose) and usually two CSs each paired with one of these outcomes. In this case, during the test phase, each CS usually only enhances lever pressing associated with the same outcome as the CS. We will call this variant the "specific transfer paradigm" (see Figure 2.2). In this paradigm, when testing transfer one lever at a time, the CS presentations can be divided into two conditions: in the *same* condition the CS and the lever share the same outcome whereas in the *different* or *diff* condition the CS predicts a different food compared to the one associated with the lever. Usually the different CS does not enhance lever pressing relative to the baseline or it does so to a lesser extent than the specific transfer effect in the *same* condition (see Figure 2.2, bottom graphs). It is still unclear why the different lever is not equally enhanced through a general transfer effect, especially when, in experimental situations able to express both the specific and general transfer effects, the two effects tend to have a comparable size (see Figure 2.2).

This last category of experiments is usually conducted using two levers, each delivering a different outcome, and with three CSs, two paired with the outcomes delivered by the levers and one paired with a third outcome. During the test, this paradigm has been reported to show both the specific and the general transfer effects. The two CSs paired with the same outcomes as the levers enhance responding on the lever sharing the same outcome (specific transfer). Again we note that the CS+ paired with the food associated with one lever does not enhance responding on the other lever although the third CS, paired with an outcome that was not used in the instrumental training, enhances pressing on both levers (general transfer). We will call this the "full transfer paradigm".

### 2.4 Behavioral results - variables influencing transfer

There are many behavioral factors affecting transfer effects. These are analysed in detail in this section.
Figure 2.2: Transfer paradigms. We illustrate three possible experimental setups (“Transfer paradigms”) that have been used to test the transfer effect. On the left (“Single-lever paradigm”), only one CS and one lever are trained with a reward and an unpaired CS is used as control. This first paradigm usually leads to the expression of general transfer. In the centre (“Specific transfer paradigm”), two CS+ and two levers are trained, using two rewards. Each CS+ is associated with a reward used for only one lever, thus enabling the expression of specific transfer. On the right (“Full transfer paradigm”), the specific paradigm is enhanced by adding an additional CS+. This last CS+ is associated with a third outcome that is not used for instrumental training. Using this third CS+ during the test phase provides a test of general transfer. The paradigm on the right is thus a “full transfer paradigm” in the sense that it can test both specific and general transfer. The bottom row provides schematic graphs that exemplify typical results obtained in the three paradigms.

2.4.1 Pavlovian factors

Response competition between Pavlovian responses and instrumental responding can make transfer effects harder to detect (Lovibond, 1983). As we noted in the previous
section, a CS+ usually enhances instrumental responding compared to the baseline. However, the Pavlovian responses elicited by the CS, such as magazine approach, can compete with lever pressing and lead to a reduction of lever pressing compared to the baseline. Using discrete (e.g. a light cue) vs. diffuse (a sound) cues as the CS can favour the development of competing sign-tracking responses (e.g. approaching the CS) which can compete with instrumental responses, or favour them if the cue is located near the manipulanda (Tomie, 1996). The degree of similarity between Pavlovian and instrumental responses might also help or hinder the transfer effect (Baxter & Zamble, 1982). Holmes et al. (2010) showed that extinguishing Pavlovian responses can enhance transfer in a subsequent test. This result might be seen to conflict with previous experiments by Delamater (1996) in which the transfer effect was shown to be unaltered after various types of CS extinction such as nonreinforcement, pairing with an alternative outcome and exposure to random or explicitly unpaired S-O contingencies. The difference might lie in the length of previous CS training, which in the case of Holmes et al. (2010) was deliberately extended to create strong Pavlovian approach responses. In other words, Pavlovian extinction might be beneficial to the transfer effect only when the Pavlovian training is sufficiently strong that it interferes with instrumental responding (in the ways described above) but has no effect otherwise. On the relationship between transfer and Pavlovian extinction, we also note that in a single-lever human transfer paradigm Lovibond et al. (2015) found, in contrast to the animal studies, that Pavlovian extinction of the CS affected transfer. Although the Pavlovian extinction did not completely eliminate the transfer effect, its efficacy contrasted with the results obtained in animals by Delamater (1996). Lovibond et al. (2015) suggested that one of the critical differences might lie in the fact that they focused on absolute response rate (using a single-lever paradigm) whereas in Delamater’s (1996) study the transfer test involved a response choice (specific-transfer paradigm). This suggestion predicts that Delamater’s extinction procedure would have been effective if he had tested his actions individually rather than in a choice test. However, given that single-lever paradigms seem to evoke general transfer, it remains possible that CS extinction differentially affects specific and general transfer. And, of course, the same is likely to be true of conflict between the CR and the instrumental action; such conflict could directly alter response vigor and so the size of any general transfer effect but has difficulty explaining variations in specific transfer because any general effect on performance should have similar effects across both actions.

The duration and timing of the CS can also affect transfer; Crombag et al. (2008a)
found in mice that a CS+ lasting 10s (with reward delivered during the last 5s of the stimulus) produced strong conditioned reinforcement but no transfer, whereas a CS+ lasting 2 minutes (with rewards delivered randomly during the interval) produced robust transfer but no conditioned reinforcement. Delamater & Holland (2008) also examined the effects of varying the CS-US interval: results confirmed that sensory-specific stimulus-outcome associations (i.e., those underlying specific transfer) were established over a wide range of long but not short intervals. These results are also consistent with older studies, e.g. Meltzer & Brahlek (1970) where a long CS (120s) led to positive transfer whereas a short CS led to suppression of responding, although in that case the CS’s were trained during instrumental conditioning and so it was not a canonical transfer design as defined here. Delamater & Holland (2008) also found that conditioned responses (magazine approach) had an inverse relationship to the CS–US interval, with longer intervals leading to a lower magazine approach performance. As such, measures of conditioned responding such as magazine approach do not necessarily correlate with measures of specific transfer. This has also been confirmed by the finding that it is possible to have a CS which reduces magazine approach but still increases instrumental actions, as we will see below (Shiflett, 2012). Delamater & Oakeshott (2007) furnishes additional support for a dissociation between magazine approach and transfer. In a specific transfer paradigm they gave rats different amounts of Pavlovian training, ranging from 4 to 112 presentations of a 60 sec CS. During the test sessions, the amount of magazine approach varied greatly between the different groups, with more training leading to more magazine approach. Also, a tendency to concentrate approaching during the last part of the CS (when the US was previously delivered) only developed for the group with most Pavlovian training. In contrast, the amount of transfer displayed was less influenced by length of training and the increase in lever pressing was more pronounced towards the end of the CS for all but the shortest training group. So both the amount and timing of magazine approach and transfer appear to develop during training at different rates.

Backward conditioning, in which the US precedes the CS, can also support transfer (Delamater et al., 2003; Shiflett, 2012; Cohen-Hatton et al., 2013); however depending on the US-CS interval, backward conditioning can produce either Pavlovian excitors or Pavlovian inhibitors, i.e. CSs that, respectively, enhance or suppress Pavlovian responses such as conditioned approach. As a consequence, the influence of backward conditioning on transfer is complex and its application in a transfer paradigm can re-

---

2But see also Van Dyne (1971) or Lovibond (1981) for short CS leading to suppression.
result in either positive (Shiflett, 2012; Cohen-Hatton et al., 2013) or ‘negative’ transfer effects (Delamater et al., 2003). The latter effects, observed in outcome specific transfer, often reflect the opposite of the standard excitatory specific transfer; i.e., rather than elevating the action that delivered the same outcome as that predicted by the stimulus, a backwardly paired CS can bias choice towards other actions that animals have learned do not deliver the backwardly paired US. This was explicitly investigated by Laurent & Balleine (2015) who compared the effect of a zero delay and of a 10-sec delay between US and CS finding standard specific transfer with the former and the ‘negative’ or reversed transfer effect with the latter. Such effects suggest, therefore, that it is the information conveyed by the backwardly paired CS that is important for the direction of choice; i.e., that the CS can provide both information about the likelihood of a forthcoming outcome for use in action selection, but can also provoke direct changes in conditioned reflexes and that these need not be the same; indeed it is important to note in this context that the effect of backward conditioning on transfer can be dissociated from its effects on conditioned approach. Shiflett (2012) managed to train a backward CS that exhibited positive transfer but suppressed conditioned approach responses (i.e., CS presentation increased lever pressing but reduced magazine approach compared to baseline). Interestingly, in Laurent et al. (2015) and in Laurent & Balleine (2015) a backward conditioning procedure was used to obtain outcome-specific Pavlovian inhibitors for two different rewards. When testing specific transfer, these backward-CSs did not reduce pressing on the lever sharing the same reward as the CS, instead, they increased pressing on the lever paired with the other reward. This was used in Laurent & Balleine (2015) to demonstrate that rats can engage in a form of “counterfactual reasoning”, meaning that they can use the information about an absent reward (furnished by the backwardly paired CS) to promote the selection of actions associated with other outcomes (in this case increasing responding on the other lever). One important aspect of these studies was the finding that specific inhibitory predictions mirrored the effects of excitatory predictions and altered action selection quite specifically away from the action delivering the outcome that the inhibitory stimulus predicted would be withheld. To confirm that this effect of backwardly pairing of CS and US was due to inhibition, Laurent & Balleine (2015) went on to assess the effects of conditioned inhibitors established using other methods; i.e., a feature-negative procedure and an overexpectation procedure. Both of these procedures produced identical effects to those induced by backward conditioning; i.e., whereas the conditioned excitors elevated performance of the action with which they shared an outcome, the conditioned inhibitors produced a shift away from the specific
action delivering the outcome the inhibitor predicted would be withheld. In summary, excitatory and inhibitory conditioning exert symmetrically opposing effects on specific transfer; biasing choice towards or away from an action based on the information provided by the cues regarding the relative likelihood of earning some specific outcome or other. In humans, Alarcón & Bonardi (2016) used a specific transfer paradigm in which they also trained a Pavlovian inhibitor using a feature-negative design (i.e. $A \rightarrow O_1$, $AX \rightarrow \emptyset$, with $X$ being the stimulus trained as an inhibitor). When the inhibitor was paired with another CS associated with the inhibited outcome during the transfer test, it abolished specific transfer and the participants instead showed a tendency to respond on the other available response, paralleling Laurent & Balleine’s (2015) results.

Lastly, we note that van den Bos et al. (2004) tried to vary the amount of reward associated with the CS (1 or 3 pellets) but found no effect of reward magnitude on transfer.

2.4.2 Instrumental factors

The amount of training and type of reinforcement schedule can alter the amount of transfer at test. Testing different variable-interval (VI) schedules Meltzer & Hamm (1974) found that VI’s with longer intervals, which lead to lower response rates, produced stronger transfer effects. This might be due to transfer being easier to detect on a lower baseline and indeed many transfer experiments include a period of instrumental extinction prior to the test on the view that it makes positive transfer easier to detect (Dickinson et al., 2000). Lovibond (1981) used discrimination training on a single lever in which animals had to discriminate between two alternating periods of reinforcement (on a VR15 schedule $S^+$, chamber fans on) and non-reinforcement ($S^-$, chamber fans off). During the transfer test, the presence of the CS suppressed responding in the $S^+$ period whereas during the $S^-$ period the CS produced positive transfer which lasted beyond the duration of the CS itself (10s). This is consistent with the above reported findings because the $S^-$ period was indeed a period with lower baseline. Lovibond (1981) also verified that the enhancement of lever pressing occurred when the baseline was lowered by other means. In a second experiment, he tested the CS after rats had achieved a low baseline due to satiety: in this case, the CS tended to suppress responding rather than enhance it. This result tells us that it was probably not a lower baseline per se that produced the stronger transfer effects but rather the interaction between the expectancy of food generated by the CS and the expectancy of food controlled by the instrumental schedule (Lovibond, 1981). In other words, the CS is more effective on
an extinguished baseline because it brings an expectation of food when the current expectation is low. The reduction in transfer observed when testing under satiety can also be seen as consistent with subsequent results by Corbit et al. (2007) and Aitken et al. (2016) which showed that satiety can abolish general transfer (see next section).

Beyond this, Holland (2004) showed that longer training on a VI schedule leads to an increase in the transfer effect when assessed in a single-lever paradigm. Subsequently, Wiltgen et al. (2012) showed that mice trained under a VI schedule exhibited more transfer than those trained under a random-ratio (RR) schedule using an outcome specific transfer design. However, although the group trained on the RR schedules was sensitive to devaluation - whereas a group trained on the VI schedules was not - both groups showed similar rates of performance during training. This is unusual; it is commonly found that ratio schedules promote far higher rates of responding than interval schedules (see Dawson & Dickinson 1990) suggesting that, for some unspecified reason, the mice in the RR group may have failed to detect the full impact of the ratio contingency. Furthermore, what transfer Wiltgen et al. (2012) reported was clearly not specific transfer; both the same and different actions were elevated relative to baseline but did not differ from each other. The authors concluded that responding that is insensitive to devaluation, and so habitual, might be more sensitive to transfer effects than goal-directed actions and, while this could be the case, it is important to qualify this statement; the evidence suggests that habitual actions are more sensitive to general transfer than goal-directed actions. If this is true, this effect would also explain Holland’s (2004) results, in which transfer correlated with the amount of VI training, with the effect growing larger as instrumental performance shifted from goal-directed to habitual control. Furthermore, in a number of studies, Balleine and colleagues have trained rats using ratio schedules and found very clear evidence of specific transfer (Corbit & Balleine, 2005, 2011; Laurent et al., 2012, 2014a; Laurent & Balleine, 2015). As such, Wiltgen et al.’s (2012) results may be anomalous, resulting from an effect specific to instrumental training rather than to the transfer test per se, or may reflect an effect on general as opposed to specific transfer. On balance it seems more likely that the latter is the correct conclusion.

2.4.3 Other factors and results

In a meta-analysis of 30 transfer studies, Holmes et al. (2010) found a correlation between the degree of transfer and the amount of phase 1 and phase 2 training, regardless of whether the Pavlovian or Instrumental training was conducted first. Increasing phase 1 training seems to increase transfer, whereas increasing phase 2 training seems to decrease
It is not clear why this training effect emerged but, presumably, it has something to do with the stabilisation of Pavlovian and instrumental learning and, therefore, the relatively consistent impact of the outcome prediction on instrumental performance. Perhaps with less training the associative strength of the CS is still below asymptote and its associative status relatively ambiguous compared to better trained CS’s. Specific transfer has been shown to be immune to devaluation: that is, pairing a food outcome with illness (LiCl-induced) prior to the test does not reduce the size of the effect in a subsequent test (Holland 2004, see also Rescorla 1994c using discriminative stimuli). However, Corbit et al. (2007) found that devaluation by satiation was able to eliminate general transfer while sparing specific transfer. This is in accord with earlier results that found no effect of devaluation on specific transfer (Holland, 2004) but apparently in contrast with Holland’s (2004) results in which devaluation did not appear to affect transfer in a single-lever paradigm. Nevertheless, as suggested above, it is difficult to distinguish specific and general transfer using this kind of design and so this failure to find an effect could be due to any specific transfer component engaged by Holland’s task. It is also possible that this discrepancy is due to the devaluation method used (satiation vs illness) or perhaps other variables. More recently Dailey et al. (2016) failed to detect transfer in a single-lever paradigm with the test conducted under satiety. Conversely, transfer was observed under satiety after administration of an antagonist of ghrelin, a peptide related to appetite signalling. Aitken et al. (2016) using a single-lever design with different foods for the Pavlovian and instrumental training, also confirmed that satiety can abolish general transfer.

Corbit & Balleine (2003a) showed that transfer can differentially affect components of an instrumental chain. In particular, they employed a paradigm where pressing a lever led to the appearance of a second lever and pressing this latter lever delivered the food reward. Devaluation and transfer differentially affected responding on these two levers. Transfer only enhanced responding to the proximal lever (i.e., the lever closest to reward delivery) whereas devaluation depressed responding on the more distal lever.

In a recent experiment, Gilroy et al. (2014) demonstrated that specific transfer can be affected by test context. They trained rats to press two levers for two different foods using two different contexts. In one group of rats (Group Differential), each lever-food pairing was trained in a specific context, whereas in the other group (Group Non-Differential) both lever-food pairings were trained in both contexts (alternating contexts in different training sessions). Pavlovian training was conducted in a third context for both groups. When tested for transfer in each of the three contexts, group Non-differential always
showed specific transfer. In contrast, Group Differential failed to show specific transfer in the Pavlovian training context but, when tested in the instrumental contexts, showed specific transfer when the CS and the context where not associated with the same food reward. Thus, Group Differential exhibited less specific transfer overall than Group Non-differential. It can be speculated that the differential training reduced the effect of CS presentation by making the contexts more informative and thus rendering the cues redundant. We will discuss this result further in Section 2.8.

Some studies have explored the relationship between transfer and stress both acute and chronic, and between transfer and stress-related molecules such as corticotropin-releasing factor (CRF) or glucocorticoids. Zorawski & Killcross (2003) tested the effects of dexamethasone, a glucocorticoid receptor agonist, in a specific transfer paradigm. They found that administering dexamethasone at the end of Pavlovian sessions impaired the ability of the CS trained during those sessions to evoke specific transfer. A similar result was also found by Pielock et al. (2013b). In 2006, Peciña et al. examined the effects of CRF microinjections in medial shell and found that the highest dose of 500 ng CRF (but not the 250 ng dose) enhanced transfer in a single-lever paradigm. Later, Morgado et al. (2012) has found that chronic stress can reduce specific transfer. A stress-free period reversed the effect. In a human transfer study (Pool et al., 2014) acute stress was also found to enhance transfer although Pielock et al. (2013a) failed to find an effect of acute stress on transfer in rodents, with both these studies using “single lever” paradigms. Soares-Cunha et al. (2014) also found that in utero exposure to elevated levels of glucocorticoids impaired both specific and general transfer in rats. Reduced levels of dopamine were observed in prefrontal and orbitofrontal cortices and normalization of these levels (using either L-DOPA or a D2/3 agonist, but not a D1 agonist) restored transfer. In humans, Quail et al. (2016) examined the relationship between scores on the Depression Anxiety and Stress Scale (DASS) and transfer, using a full transfer paradigm which also included a fourth cue associated with no reward. Participants with higher scores on the Anxiety and Stress subscale showed higher cue-driven response vigour, meaning that they showed increased instrumental responding even in the presence of the fourth cue associated with no reward. High anxiety participants also seemed to show a somewhat blunted specific transfer, with the cues paired with the two instrumental rewards enhancing both the same and different instrumental actions to a similar degree, albeit this trend compared to low anxiety participants was not statistically significant.

Using a chronic unpredictable stress paradigm, composed of daily exposures of 1 hour to one the following stressors: cold water, vibration, restraint, overcrowding, and a hot air stream.
2.5 Neural substrates

2.5.1 Amygdala and nucleus accumbens

Starting from the beginning of this century, lesion studies on rats have begun to uncover the neural basis of the various transfer effects, reporting that both nucleus accumbens (Hall et al., 2001; Corbit et al., 2001; de Borchgrave et al., 2002) and amygdala (Blundell et al., 2001; Hall et al., 2001; Holland & Gallagher, 2003) are necessary for it to take place. During the first studies: an initial disagreement arose about which parts of amygdala (BLA or CeA) and which parts of nucleus accumbens were involved (Nacc Core or Shell). It was reported that CeA and Core, but not BLA or Shell were involved in transfer (Hall et al., 2001; Holland & Gallagher, 2003) and also the opposite pattern of results (Blundell et al., 2001; Corbit et al., 2001). These data were later reconciled when Corbit & Balleine (2005) introduced the full transfer paradigm which was able to distinguish specific transfer from general transfer. Specific transfer is abolished by BLA and Shell lesions whereas general transfer is abolished by CeA and Core lesions (Corbit & Balleine, 2005, 2011). Muscimol-induced inactivation of core and shell have similar effects as lesions of these structures (Corbit & Balleine, 2011). As a further confirmation, Shiflett & Balleine (2010) used a disconnection procedure between BLA and either medial shell or core and found that only the disconnection involving shell abolished specific transfer. Lingawi & Balleine (2012) found that both anterior and posterior CeA lesions abolished transfer in a single-lever paradigm (therefore presumably general transfer). Human fMRI studies have also confirmed the involvement of amygdala and ventral striatum in transfer (Talmi et al., 2008; Bray et al., 2008; Prévost et al., 2012; Mendelsohn et al., 2014).

Studies investigating the role of glutamate receptors are also in line with the dissociation of neural substrates involved in single and two-lever studies and furnish some details on the mechanisms working inside amygdala that mediate transfer. Mead & Stephens (2003a,b) used a single-lever paradigm to investigate the effects of AMPA receptor subunits GluR1 and GluR2 deletion in mice. Neither deletion impaired Pavlovian or instrumental training, however GluR1 deletion impaired conditioned reinforcement (usage of CS+ as reinforcers) leaving transfer intact, whereas GluR2 deletion impaired transfer while leaving conditioned reinforcement intact. The authors suggested that this dissociation might be explained by GluR1 and GluR2 deletion impacting learning in BLA and CeA respectively, as the behavioral consequences of these deletions mimicked the effects of lesions on these structures (Mead & Stephens, 2003b). This hypothesis was
supported by a later finding by Johnson et al. (2007) who found that GluR1 deletion impaired specific transfer. In particular they found that mice without GluR1 expressed non-selective transfer when trained in a specific transfer paradigm: i.e. they increased pressing on both levers to similar degree when presented with a CS paired with the outcome delivered by one of the levers, whereas wild-type mice showed specific transfer. All these results are thus consistent with the view that single-lever studies evoke general transfer (mediated by CeA) whereas the specific transfer in two-lever studies is mediated by BLA, without the necessary presence of CeA. However, subsequently Crombag et al. (2008b) found that mutations on GluR1 phosphorylation sites can abolish single-lever transfer in mice. This is at odds with previous results and it may be due to GluR1 deletion triggering different compensatory effects to those induced by altering phosphorylation sites in Crombag et al. (2008b).

Malvaez et al. (2015) further characterized specific transfer processes in the BLA. They monitored glutamate concentrations during the transfer test and found that glutamate transients were time-locked to and correlated with only the instrumental pressing directed to the lever sharing the same outcome as the CS (i.e., in a specific transfer test). In addition, local blockade of AMPA receptors, but not NMDA receptors, abolished specific transfer. Related results were also obtained by George et al. (2009) who showed that the selective mGluR5 antagonist 2-methyl-6- (phenylethynyl)-pyridine (MPEP) reduced transfer in a single-lever study and by Murschall & Hauber (2005) who did not detect any impact on single-lever transfer using systemic AMPA/KA and NMDA blockade.

Leung & Balleine (2013) characterized the circuit downstream from the Nacc shell by investigating one of its main projections, the medial ventral pallidum (VP-m). Rats exposed to a specific transfer test showed higher expression of the cellular activity marker c-fos in both shell and VP-m compared to controls. Also, both VP-m inactivation and shell-VP-m disconnection procedures abolished specific transfer. In a previous study, lesions of mediodorsal thalamus (MD), which is further downstream as it receive VP outputs, were found to impair specific transfer (Ostlund & Balleine, 2008). As a consequence, Leung & Balleine (2015) examined the functional contributions of both MD and VTA, which is another target of VP-m. Results showed that VP-m neurons projecting to MD were more active (c-fos) during the transfer test than those projecting to VTA, but it was the activation of these latter neurons that correlated with the absolute size of the transfer effect. Interestingly, by using a disconnection procedure Leung & Balleine (2015) then demonstrated that disrupting the VP-m to VTA pathway reduced the overall rate of responding (similar to a finding by Corbit et al. (2007) following VTA
inactivation) whereas disconnecting the VP-m to MD pathway removed the bias of the specific predictive cues on choice (i.e., the specific transfer effect). In this latter VP-m to MD pathway disconnection, the CS elevated the performance of both actions; i.e. both the lever delivering the same outcome as the stimulus and the different action, suggesting that the VTA mediates the motivational component of specific transfer (overall rate of lever press performance) whereas the MD mediates the cognitive component (the effect of predictive learning on choice).

2.5.2 Molecular processes in nucleus accumbens

A series of studies has attempted to uncover the molecular underpinnings of transfer. Remus & Thiels (2013) investigated ERK kinase activation during transfer in core and shell. They found that CS presentations caused a significant increase in ERK activation in both subregions, with the effect being slightly more robust in the core than the shell. An in-depth review of molecular mechanisms involving ERK in transfer and instrumental conditioning can be found in Shiflett & Balleine (2011). Lex & Hauber (2008) examined instead the effects of D1 and D2 receptor antagonism using SCH-23390 and raclopride injections, respectively, into both core and shell. Both core and shell D1 antagonism abolished transfer in a single-lever paradigm, with D2 antagonism also reducing transfer but to a lesser extent. Similarly Peciña & Berridge (2013) tested the ability of dopamine stimulation (amphetamine microinjection) versus µ-opioid stimulation (DAMGO microinjection) in either core or shell to amplify transfer. Both amphetamine and DAMGO augmented the transfer effect in a single-lever design and did so when infused in both core and medial shell, excluding only a small far-rostral strip of shell. Consistently with this result, in Weber et al. (2016) administering a µ-opioid antagonist (naltrexone) reduced transfer in humans. In contrast to the effects of Lex & Hauber (2008), Laurent et al. (2014a) found that SCH-23390, but not raclopride, injections specifically in the Nacc shell abolished transfer in a specific transfer paradigm. Furthermore, in a complex series of experiments, Laurent et al. (2014a) showed that delta-opioid receptors (DORs) on cholinergic interneurons (CINs) mediate specific transfer in Nacc shell by altering CIN firing and their effect on D1-expressing medium-spiny neurons (MSNs). First, they confirmed the involvement of shell D1-expressing MSNs by measuring ERK phosphorylation (pERK) after the transfer test and by infusion of SCH-23390 and raclopride. Results showed enhanced pERK in D1 but not in D2-expressing MSNs confirming the effect of SCH-23390 relative to raclopride. They then provided data to support the interaction between processes involving DORs and D1Rs using asymmetrical infusions in the
shell. The results showed that rats with infusions of SCH-23390 in one hemisphere and naltrindole (a DOR antagonist) in the other, exhibited no specific transfer. In addition, electrophysiological recordings of CINs in Nacc shell slices taken after the transfer test confirmed alterations in their firing patterns when exposed to deltorphin (DOR endogenous ligand) compared to CINs in slices taken from rats exposed to non-contingent CS training. It was hypothesized that the effect of DOR on CIN firing altered acetylcholine release and that changes in acetylcholinergic release affected the activity of D1 neurons. With regard to this latter effect, it was further hypothesized that acetylcholine alters D1 MSN activity through the activity of Gi-coupled M4 receptors recently found to be uniquely expressed on post-synaptic D1 MSNs. Increased activity at the M4 receptor has been found to inhibit D1 activity and reduced M4 binding to increase D1 activity. If DOR activity inhibits acetylcholine release and naltrindole blocks that reduction then a ready explanation for Laurent et al’s results could be provided. To test this, Laurent et al. (2014a) sought to block specific transfer using peripherally administered naltrindole and then to release that blockade by infusing the M4 antagonist MT3 into the Nacc shell. Although MT3 had no effect on its own, its infusion rescued specific transfer after it was abolished by naltrindole, To summarize, Laurent et al. (2014a) suggest that, in the Nacc shell, specific transfer is mediated by a complex interaction involving opioid, cholinergic and dopaminergic systems (see Figure 2.3). The basis for this interaction is formed during Pavlovian training, with the increased expression of DORs on CINs at the cell membrane (Bertran-Gonzalez et al., 2013). DORs are then activated during the transfer test by enkephaline (ENK), the latter possibly released by D2 MSNs, altering CIN activity and causing lower acetylcholine release overall and thus a lower level of M4 binding, resulting in increased D1 neuron activity and increased specific transfer as a consequence.

A confirmation of the involvement of acetylcholine in specific transfer can also be found in Ostlund et al. (2014a), were the effects of scopolamine (muscarinic receptor antagonist) and mecamylamine (nicotinic receptor antagonist) were tested. In this case, both muscarinic and nicotinic acetylcholine systemic antagonism impaired specific transfer. Collins et al. (2016) also examined the role of acetylcholine and its interaction with dopamine, however they focused on Nacc core, using a single-lever paradigm. Using local infusions of scopolamine and mecamylamine combined with fast-scan cyclic voltammetry they showed that acetylcholine antagonism can modulate both transfer and cue-evoked

4In Laurent et al. (2015) naltrindole infusions into shell were also shown to abolish the reversed specific transfer induced by backward conditioning
dopamine release in Nacc core. In particular, they showed that muscarinic receptor antagonism suppresses both transfer and cue-evoked dopamine release, while nicotinic receptor antagonism augments transfer and cue-evoked dopamine release. The discrepancy between these latter results and Ostlund et al.’s (2014a) results in which both scopolamine and mecamylamine impaired transfer may be either due to the type of infusions (systemic vs. local) or due to the paradigm used (specific transfer vs. single-lever).

Figure 2.3: Model of the complex interactions of opioid, cholinergic and dopamine systems in the Nacc shell during specific transfer, from Laurent et al. (2014a). DOR receptors on CIN cell bodies are activated by ENK. ENK alters the firing pattern of CINs, leading to a lower Ach release. Lower Ach release leads to less activation of M4 receptors on D1R-expressing neurons. In turn, the lower activity of M4 permits cAMP pathway signalling, increased D1 neuron activity and so increased specific transfer expression. See text for details of the experiments leading to this model. Reprinted with permission.
2.5.3 Dorsal striatum

Corbit & Janak (2007b) found that inactivation of either dorsomedial (DMS) or dorso-lateral (DLS) striatum in a specific transfer paradigm had different effects on transfer expression: DLS inactivation abolished specific transfer whereas DMS rendered transfer non-specific, with the CS increasing performance on both the same and different outcome lever. In a later study (Corbit & Janak, 2010a), DLS, anterior DMS (aDMS) or posterior DMS (pDMS) were inactivated during Pavlovian and instrumental training: inactivation of DLS and pDMS appeared to impair the acquisition of stimulus-outcome (S-O) associations whereas aDMS and pDMS inactivation impaired response-outcome (R-O) associations, as revealed by Pavlovian and instrumental devaluation tests respectively. In all of the inactivation groups - aDMS, pDMS, and DLS - specific transfer was impaired in a subsequent transfer test. This was expected because specific transfer requires the integration of both S-O (Pavlovian) and R-O (instrumental) associations and any loss of these should, quite naturally, be predicted to impair this integrative process.

In contrast, Pielock et al. (2011) examined the effects of 6-OHDA induced DA depletion in aDMS and pDMS on transfer using a single lever transfer paradigm. Neither depletion had any effect in the single-lever PIT paradigm. Furthermore, in a second experiment that used the specific transfer paradigm, aDMS 6-OHDA again had no effect whereas in the pDMS it had only a minor, if any, effect, suggesting that the dopamine innervation of dorsal striatum is not necessary for transfer.

2.5.4 Midbrain structures

In a single-lever paradigm, Murschall & Hauber (2006) found that ventral tegmental area (VTA) inactivation by baclofen/muscimol can abolish transfer (supposedly general transfer). However, Corbit et al. (2007) using the same dose effective in Murschall & Hauber (2006) found in the full transfer paradigm that VTA inactivation only attenuated specific and general transfer; essentially reducing but not abolishing these effects.

El-Amamy & Holland (2007) used instead a disconnection procedure to separate CeA from either substantia nigra pars compacta (SNpc) or VTA. In a single lever paradigm, both CeA-SNpc and CeA-VTA unilateral lesion reduced transfer; a CeA-SNpc contralateral lesion (disconnection) also reduced transfer, with no additional effect to the unilateral lesion. Puzzlingly, CeA-VTA disconnection was found to restore transfer to control levels. The result was explained by reference to possible cross-hemispheric inhibitory connections between the two VTAs and CeA output to the SNpc (see also Lee
Dopamine

Other results have, however, pointed to the involvement of dopaminergic areas such as VTA and SNpc in transfer. Dickinson et al. (2000) showed that dopamine antagonists (such as flupenthixol) reduced transfer using a single-lever paradigm, a result also replicated in Wassum et al. (2011). Later, Lex & Hauber (2008) used D1 and D2 receptor antagonism in core and shell and confirmed the ability of dopamine antagonism to reduce transfer in a single-lever paradigm. Conversely, transfer is potentiated by indirect dopamine agonists such as amphetamine (i.e. Wyvell & Berridge (2000), see also section 2.6). Using fast-scan cyclic voltammetry Wassum et al. (2013) monitored dopamine release in real time in nucleus accumbens core and showed that during a single-lever transfer test phasic dopamine release evoked by the CS correlated with the increase in lever pressing. This result was also replicated in Aitken et al. (2016). Ostlund & Maidment (2012) instead examined the effects of flupentixol in a specific transfer paradigm, presenting both levers during the test session. They found that although flupentixol reduced the response invigoration generated by the CS, it did not influence the ability of the CS to bias action selection towards the lever sharing the same outcome as the CS. In humans, Weber et al. (2016) administered amisulpride (a selective D2/D3 antagonist) and found reduced transfer in a single-lever paradigm. All these results (and also Soares-Cunha et al. (2014); Laurent et al. (2014a) described elsewhere in this review) indicate that the increase in lever pressing during the transfer effect is mediated by dopamine (and likely by VTA); interestingly, the bias towards one lever in a choice situation might be dopamine-independent instead (Ostlund & Maidment, 2012).

2.5.5 Prefrontal cortex

Basal ganglia structures such as the Nacc form a closed loop with prefrontal cortex and so, given the previous results, the involvement of PFC structures in transfer would be expected. Indeed, Ostlund & Balleine (2007a) found that post-training lesions of OFC abolished specific transfer, while pre-training lesions had no effect. Subsequently

---

3We note that in Ostlund & Balleine (2007a) and in Ostlund & Balleine (2008) lesions of OFC and mediodorsal thalamus were reported to abolish specific transfer although rats actually kept pressing more in the same and diff conditions compared to the baseline. Although the effect was not statistically significant in those experiments, it is possible that these lesions do not abolish specific transfer but render it “non-specific”. Indeed, in Leung & Balleine (2015), disconnection of VP-m from the MD produced a similar “non-specific” transfer effect such that the CS significantly elevated
Balleine, Leung and Ostlund (2011) reported that this post-training effect was likely a product of sparing aspects of ventral OFC. Complete ventral and lateral OFC lesions produced pre-training deficits in specific transfer (Balleine et al., 2011). Similarly, Scarlet et al. (2012) found a reduction in specific transfer using pre-training OFC lesions, albeit using a very different paradigm. In Bradfield et al. (2015) lesions targeting the medial OFC did not eliminate transfer but instead made it non-specific: i.e. the CS no longer enhanced lever pressing in the same condition, but also in the diff condition. Despite both ACC and PL having connection with Nacc, Cardinal et al. (2003) found no effects of ACC lesions on transfer using a single-lever paradigm, whereas Corbit & Balleine (2003b) found no effects of PL lesions in a specific transfer paradigm. More recently, Keistler et al. (2015) found that bilateral IL lesions abolish specific transfer. They also employed an IL-Shell disconnection procedure and confirmed that IL mediates this effect via functional connectivity with Nacc shell, also part of the specific transfer circuitry.

2.5.6 Neural correlates

Finally, we briefly mention two studies that have investigated the neural correlates of transfer using electrophysiological recordings: Homayoun & Moghaddam (2009) and Saddoris et al. (2011). Both studies used a single-lever transfer paradigm. Homayoun & Moghaddam (2009) recorded from medial (mPFC), orbital prefrontal cortex (OFC) and dorsal striatum (DS) of freely moving rats during transfer test sessions. They found that the CS+ amplified neuronal responses to both instrumental nosepokes and Pavlovian approaches in all of these structures. However, the amplification of the instrumental responses correlated with the strength of transfer only in the mPFC and OFC but not DS. Moreover, DS neurons represented transfer and approach behavior through mostly-segregated populations, whereas in mPFC and OFC they were represented in overlapping populations of neurons.

In Saddoris et al. (2011) electrophysiological recordings of the single lever transfer test session were conducted in the Nacc core and shell. Multiple groups were used with one exposed to separate cocaine self-administration training. In all groups, neurons in both core and shell encoded information about cues, rewards and responses. In control animals, core neurons were more likely to encode this information, which correlated with behavioral performance in the transfer test. However, neurons that expressed transfer-specific encoding (their lever-press related activity was increased during CS periods)

responding on both the same and diff lever.
correlated with transfer performance in the shell, but not in the core. The group with a history of cocaine self-administration showed increased transfer and increased neural encoding of task events in the shell. Generally, these studies are consistent with what has been found in lesion and inactivation studies; however, the possibility of mixed transfer effects emerging in single action studies likely accounts for the breadth of the observed effects.

2.6 Interaction with drugs of abuse

The importance of conditioned stimuli in promoting drug-taking and relapse is widely recognised in addiction research (Belin et al., 2009). Conditioned stimuli associated with drugs of abuse can promote drug use through various mechanisms, such as conditioned approach, conditioned reinforcement as well as Pavlovian-instrumental transfer (Belin et al., 2009). Despite this, relatively few studies have been conducted using the transfer paradigm with drugs of abuse. As an example, searching Pubmed with terms referring to transfer (e.g. PIT, transfer) and addiction (e.g. addiction, cocaine, ethanol, opiate) we found only 27 articles with experiments investigating the relationship between transfer and drugs of abuse. This is a small number compared to other paradigms used in addiction research, e.g. conditioned place preference, for which it is possible to find hundreds of results. As expressed by LeBlanc et al. (2012), this might be due to “experienced or perceived difficulties in generating the PIT [transfer] effect using conventional drug self-administration procedures”.

Studies investigating transfer with drugs of abuse have involved different substances, including amphetamine (Wyvell & Berridge, 2000, 2001; Peciña et al., 2006; Peciña & Berridge, 2013; Hall & Gulley, 2011; Shiflett, 2012; Shiflett et al., 2013), cocaine (Kruzich et al., 2001; Saddoris et al., 2011; LeBlanc et al., 2012, 2013, 2014; Ostlund et al., 2014b), and ethanol (Krank, 2003; Ripley et al., 2004; Glasner et al., 2005; Corbit & Janak, 2007a; Krank et al., 2008; Milton et al., 2012; Depoy et al., 2014; Corbit et al., 2016) in rodents and tobacco (Hogarth et al., 2007; Hogarth & Chase, 2011, 2012; Hogarth, 2012; Hogarth et al., 2013b, 2014, 2015) and beer (Martinovic et al., 2014; Garbusow et al., 2014, 2016) in humans.

These studies show that transfer can be observed with drugs of abuse just as it is observed with natural rewards: a CS associated with drugs of abuse can enhance instrumental responding directed to the drug itself, just as a CS associated with food enhances food responding (e.g. Krank 2003; LeBlanc et al. 2012; Hogarth et al. 2007). However,
there are also some peculiarities of transfer when the subjects are exposed to drugs of abuse. In fact, some studies have found a stronger transfer effect in subjects that have been exposed to drugs of abuse compared to controls (Wyvell & Berridge, 2000, 2001; Saddoris et al., 2011; LeBlanc et al., 2013). This enhancement of transfer can be found even when the subject has been previously sensitized and then given the transfer test when drug-free (Wyvell & Berridge, 2001; Saddoris et al., 2011). In other cases, drugs of abuse led to the CS enhancing instrumental responding for all rewards in conditions that usually lead to specific transfer instead (Glasner et al., 2005; Corbit & Janak, 2007a; Shiflett, 2012); i.e., the CS enhanced lever pressing similarly in both the same and different conditions in a specific transfer paradigm. Apparently, exposure to drugs of abuse (such as alcohol in Glasner et al., 2005; Corbit & Janak, 2007a or amphetamine in Shiflett, 2012) promoted the expression of general instead of specific transfer. A recent experiment (Corbit et al., 2016) confirmed that this is the case by showing that the transfer effect evoked by an alcohol paired CS on both an alcohol paired and a sucrose paired lever is mediated mainly by nucleus accumbens core, a structure involved in general (but not specific) transfer. In Corbit et al. (2016) rats were trained in a specific transfer paradigm using both alcohol and sucrose as rewards. An alcohol-paired CS increased responding to both alcohol and sucrose paired levers, while sucrose-paired CS increased responding more selectively to the same reward lever. Temporary inactivation of either nucleus accumbens core or shell using baclofen/muscimol injections altered these transfer effects: core inactivation reduced transfer induced on both levers by the alcohol CS while leaving transfer induced by a sucrose CS mostly intact directing performance towards the sucrose lever. In contrast, shell inactivation reduced the specificity of transfer while leaving both CS’s capable of enhancing lever pressing on both levers.

As we said above, some of these studies have revealed increased transfer after exposure to drugs of abuse compared to drug-naive controls, whereas other have shown that drugs of abuse can lead to the expression of general transfer under conditions in which specific transfer is usually expressed. For brevity, in the following, we will refer to these findings as potentiation and generalization effects.

Interestingly, studies of transfer with drugs of abuse have so far shown either a potentiation of transfer by drugs of abuse or generalization effects but not both at the same time, with psychostimulant studies usually finding the former (Wyvell & Berridge, 2000, 2001; Peciña et al., 2006; Peciña & Berridge, 2013) and ethanol studies the latter (Glasner et al., 2005; Corbit & Janak, 2007a; Corbit et al., 2016). However, it seems likely that this difference is due to the paradigms used and not the substance. Indeed,
Shiflett (2012) reported a *generalization* effect when testing amphetamine-sensitization effects on transfer. Studies finding *potentiation* have generally done so in a situation using drug-naive controls but only one rewarding outcome (thus they could not detect *generalization*); whereas studies finding *generalization* have used two rewards but no drug-naive controls (which prevented them from assessing any *potentiation*). It would be interesting to confirm this in a study that aimed to assess both *potentiation* and *generalization* effects at the same time. Nevertheless, it should be noted that not all studies have found these effects.

In some cases this can be attributed to the procedure: Krank (2003), Krank et al. (2008) and Kruzich et al. (2001) used only one reward for all subjects, thus making it impossible to detect either *generalization* or *potentiation*. However, Depoy et al. (2014) compared two groups of rats, one exposed to a prolonged chronic intermittent ethanol exposure (CIE, 16 daily bouts, using alcohol vapors) and a control group. The CIE group showed less transfer in a single-lever paradigm using food rewards. The authors suggested that this might related to the length of CIE, with shorter CIE promoting dorsal striatal-mediated behaviors whilst longer exposures might impair them (DePoy et al., 2013; Depoy et al., 2014). Ripley et al. (2004) also found less transfer in groups exposed to ethanol-withdrawal compared to controls, using a single-lever paradigm with sucrose as reward. However, rats exposed to ethanol had higher baseline levels of responding and the impairment in transfer was significant only when looking at the ratio between responding during the CS and the baseline. As we noted in Section 2.4, transfer is harder to detect when baseline responding is higher.

Drugs of abuse can alter dopamine transmission and dopamine is involved in the expression of transfer as we have seen in section 2.5. Hence, a possible neural mechanism of *potentiation* might involve dopamine: Ostlund et al. (2014b) monitored phasic DA release in Nacc core using fast-scan cyclic voltammetry and reported that prior cocaine exposure enhanced both transfer and mesolimbic DA signalling, with both measures being correlated. In contrast, *generalization* might be related to the ability of drugs to promote habitual actions (LeBlanc et al., 2013). Given that specific transfer requires action-outcome encoding, it might be argued that a switch from goal-directed to habitual actions should also promote a switch from specific to general transfer, providing a possible explanation of the *generalization* effect. However, results from Hogarth et al. (2013b) suggested that a drug-induced switch to habitual actions did not cause a loss of specific transfer, at least in human subjects.

This latter study was part of a series by Hogarth and colleagues assessing habitual
smokers in which transfer was used to understand the role of cues in drug-seeking (Hogarth et al., 2007; Hogarth & Chase, 2011, 2012; Hogarth, 2012; Hogarth et al., 2013b, 2014, 2015). In this series of studies, it was shown that a CS associated with tobacco biased choice towards tobacco-related actions in an instrumental choice test and that this effect was not diminished when tobacco was devalued by showing health warnings or through satiation (i.e. after a smoking session). In these studies, neither generalization nor potentiation effects were detected. It is possible that the presence of these effects might have been masked; for example the potentiation effect might be masked by the fact that the participants were all smokers, so there was no drug-naive control. However, there was a distinction between daily and non-daily smokers and no correlation was found between this factor and transfer. As for the generalization effect, we point out that the transfer test was always conducted as a choice test; thus the ability of the drug CS to prompt both choices may not be detected because an enhancement in both choices would balance each other out when analyzing proportion of choices.

As already mentioned, transfer and other Pavlovian conditioning effects are considered to play an important role in drug addiction (Belin et al., 2009). Milton and colleagues have, therefore, studied how pharmacological intervention in memory reconsolidation processes could disrupt maladaptive Pavlovian associations. In particular, Milton et al. (2012) found that both transfer and Pavlovian approach effects can be disrupted using the NDMAr antagonist MK-801 in conjunction with Pavlovian memory reactivation in ethanol self-administering rats; in contrast, the β-adrenergic antagonist propranolol had no effect (although it was previously shown to affect conditioned reinforcement by Milton & Everitt 2010).

Recently, studies have investigated the transfer effect in alcohol dependent subjects. In a first pilot study, Garbusow et al. (2014) tested transfer in both detoxified alcohol-dependent patients and controls. In their paradigm, instrumental actions led to either monetary rewards or losses and, similarly, Pavlovian CSs were associated with either monetary reward, with losses, or with pictures of drinks (alcohol or water). The patient group showed a stronger transfer effect on the negative transfer part of the experiment; i.e., they showed stronger suppression of instrumental responses by the CS associated with monetary losses. In a subsequent study this stronger transfer effect in the detoxified alcohol-dependent group was confirmed compared to controls (Garbusow et al., 2016). Indeed, in this study a stronger overall transfer effect was detected, not just an effect specific to the negative CS. In addition, participants underwent fMRI scanning which found that BOLD activation in the left Nacc was related to the transfer effect. This
activation was also predictive of relapse and alcohol intake during a 3 month followup, thereby establishing the clinical significance of transfer and its neural correlates for alcohol-dependence treatment. Martinovic et al. (2014), instead, tested transfer using beer and chocolate cues on a group of social drinkers but in this study no correlation was found between transfer effects and individual differences of drinking behavior. However, in this latter study, as in Hogarth studies of smokers, the participants were drug users (but not necessarily dependent) and no control group naive to the substance was used, which might account for their failure to find a potentiation effect.

Readers interested in the relationship between transfer, drugs of abuse and addiction might also refer to the recent review by Lamb et al. (2016) which analyzed some of the studies reported here in terms of their support for theories of addiction.

2.7 Human transfer

In the last decade, transfer has also been investigated in humans, with results similar to those previously shown in rodents. The presence of both specific and general transfer in humans has been confirmed (Nadler et al., 2011; Prévost et al., 2012; Watson et al., 2014). Studies using fMRI have also confirmed roughly the same main neural structures underlying transfer with the involvement of the amygdala (Talmi et al., 2008; Prévost et al., 2012; Mendelsohn et al., 2014) and NAcc (Talmi et al., 2008; Mendelsohn et al., 2014) or the nearby ventrolateral putamen (Bray et al., 2008; Prévost et al., 2012). More detailed confirmation such as the involvement of shell versus core in the two types of transfer might follow as fMRI resolution increases.

The tools and experimental approaches vary greatly among human transfer studies, with some employing a “game-like” paradigm with abstract rewards (Paredes-Olay et al., 2002; Allman et al., 2010) whilst others have employed paradigms similar to rodent studies using food rewards (Bray et al., 2008; Watson et al., 2014). In particular, Lovibond & Colagiuri (2013) developed a human paradigm very close to the rodent version (even using a food dispenser). Morris et al. (2014), on the other hand, have used perhaps the most ecologically valid paradigm involving the interaction with a simulated vending machine. Participants learned to associate different colors on the vending machine with different snack food outputs (the Pavlovian phase), whereas the instrumental phase involved choosing between two actions that tilted the machine and liberated different snack foods. As another interesting variant, one fMRI study confirmed the involvement of the amygdala and Nacc in transfer using motor imagery rather than the motor ac-
tion: the instrumental action thus consisted in the participants imagining themselves throwing a ball or a rock with their right hand without actually moving it (Mendelsohn et al., 2014).

Some studies have investigated the role of devaluation in transfer, finding conflicting results: Allman et al. (2010) reported that devaluation was able to eliminate specific transfer whereas Watson et al. (2014) reported devaluation having an effect only on general transfer with no effect on choice bias (specific transfer). These differing results might be explained by the different paradigms used. In Allman et al. (2010) the participant played a “stock market game” in which devaluation was induced by ensuring one of the outcomes no longer had any monetary value; whereas in Watson et al. (2014) devaluation was generated by food satiation. The devaluation in Allman et al. (2010) was, therefore, of a more cognitive nature and it might have worked as an additional rule of the game instead of a “motivational” change, such as satiation, thus leading to the different result. Furthermore, Eder & Dignath (2016b) also obtained specific transfer after devaluation and suggested that the previously contrasting results might also be due to using a primary reinforcer (food) vs. a secondary one (money). However, Eder & Dignath (2016b) found that specific transfer on primary rewards can be eliminated by devaluation if the participants had to consume the reward that was made aversive. In particular, they ran two experiments using a liquid reinforcer that was devalued by pairing its consumption with bad tasting Tween20. In one case, participants had to drink the earned liquid rewards immediately after the transfer test, whereas in the other case they could just take the bottles home. Only in the first case, when they were asked to consume the devalued reward, did devaluation eliminate the specific transfer effect during the test. It is interesting to note that Eder & Dignath’s (2016b) results were obtained using aversive devaluation, which might lead to different results compared to devaluation through satiation (see also Corbit et al. (2007) vs. Holland (2004) results in animal studies, Section 2.4). Subsequently Eder & Dignath obtained again a reduction of specific transfer through devaluation using a “stock-market game” similar to the one used by Allman et al. (2010). In this case, they also tested the effects of an “upvaluation” of the outcomes (an increase of their monetary value in the game) but observed no effect of the upvaluation on transfer (Eder & Dignath, 2016a). Instead, Colagiuri & Lovibond (2015) tested transfer after variable amounts of instrumental training to manipulate satiation and provided data in support of transfer as a mechanism of food

---

6On a similar note, Colwill & Rescorla (1990) showed that whereas devaluation usually leaves some residual instrumental responding, this responding could be eliminated if the reward was delivered intraorally instead of into a magazine.
over-consumption. The overall results showed that participants with a low-baseline response rate, which had supposedly reached a higher level of satiation during the instrumental training, still showed a significant transfer effect. This further confirms that transfer is present after devaluation, at least when conducted using satiation. In contrast, high-baseline responders, which where supposedly less sated and more actively seeking the chocolate outcome, showed no significant transfer effect when tested with the CS+ but did show an inhibitory transfer effect when tested using the unpaired CS-. This latter finding is in line with the fact that positive transfer effects are harder to find when baseline responding is higher (see Section 2.4).

Beyond Colagiuri & Lovibond (2015), transfer as a mechanism of food over-consumption was also hypothesized in Watson et al. (2016). In this latter work, transfer using food rewards was tested and found in adolescents, as a part of an investigation about possible factors of an obesogenic environment.

Rosas et al. (2010) studied the relationship between specific transfer and extinction and found that extinction does not alter the ability of cues to promote specific transfer. This is consistent with the results of Delamater (1996) in rodents. In a recent study, Hogarth et al. (2014) further confirmed that extinction by non-reinforcement does not alter specific transfer. However, these authors also found that discriminative extinction training (pairing the CS with the extinction of an instrumental response) can abolish specific transfer, a result also found earlier by Gámez & Rosas (2005). In a third experiment, Hogarth et al. (2014) also managed to abolish specific transfer using explicit verbal instructions stating that the CS did not signal a more effective response-outcome contingency. These results suggest the hypothesis that the specific transfer effect might be of a hierarchical nature, with the CS signalling the efficacy of an action-outcome (A-O) contingency (Hogarth et al., 2014; Hogarth & Troisi, 2015). Thus, extinction procedures such as discriminative extinction or the verbal instructions targeting directly the S-(A-O) hierarchy are effective in blocking transfer whereas those targeting the simple S-O association are not. Of course such procedures may also be effective because they produce a form of instructed extinction of the instrumental contingency. Lastly, as already mentioned in Section 2.4, using a single-lever paradigm Lovibond et al. (2015) found a reduction in transfer after Pavlovian extinction of the CS. As the authors point out, one of the critical differences between their results and those previously reported was the fact that they focused on absolute response rate (using a single-lever paradigm), whereas other studies employed a choice test (specific transfer). Given that single-lever paradigms usually measure general transfer, we thus speculate that specific and general
transfer might be differentially affected by CS extinction.

As already reported (see Section 2.6), human studies have also investigated transfer in the context of drugs of abuse, such as tobacco and alcohol (Hogarth et al., 2007; Hogarth & Chase, 2011; Hogarth, 2012; Hogarth et al., 2013b; Martinovic et al., 2014; Garbusow et al., 2014, 2016). In a series of studies (Hogarth et al., 2007; Hogarth & Chase, 2011; Hogarth, 2012; Hogarth et al., 2013b, 2015), Hogarth and colleagues showed with a transfer paradigm that cues related to cigarettes can affect smoking and that they do so independently of value (confirming that specific transfer is immune to devaluation as in Watson et al. 2014). Transfer was also investigated in alcohol users and abusers (Martinovic et al., 2014; Garbusow et al., 2014, 2016). Garbusow and colleagues investigated transfer in groups of detoxified alcohol-dependent patients and in controls, finding stronger effects in the patients group (Garbusow et al., 2014, 2016). The strength of transfer effect was also related to a stronger left Nacc activation and it was predictive of relapse and alcohol intake in a 3 months follow-up (Garbusow et al., 2016). Martinovic et al. (2014) tested transfer in social drinkers instead and they did not find a correlation between the behavioral transfer effect and individual differences in drinking behavior.

Freeman et al. (2014) developed a variant of transfer in which instrumental training was embedded in a Go-NoGo task; i.e., a normal single-lever paradigm was subdivided into trials in which an additional cue signaled a Go condition (the participant can press the lever) or NoGo (the lever must not be pressed). Within this paradigm, Freeman et al. (2014) and Freeman et al. (2015) investigated the mechanism of response suppression during CS+ “provocation”, finding that NoGoCS+ trials suppressed CS+ response not only within the trial but also on the following trials.

Garofalo & di Pellegrino (2015) instead investigated individual differences in a single-lever transfer paradigm. In particular, they subdivided the participants into two groups according to their eye-gaze behavior during the Pavlovian phase. During CS presentations, those who spent most time fixating on the CS were considered sign-trackers, whereas those who spent most time fixating on the location where the reward would be delivered were classified as goal-trackers. The results of the transfer test showed that only sign-trackers increased their lever pressing when the CS was presented. Another study focusing on transfer and individual differences was conducted by Sebold et al. (2016) in which they tested 243 participants in both a transfer paradigm and in a two-step task aimed at distinguishing model-based vs. model-free reasoning (Daw et al., 2011). Model-based and model-free reasoning are computational concepts relating respectively to goal-directed and habitual behavior. Sebold et al. (2016) found a
correlation between the transfer task and the two-step task, indicating that people exhibiting stronger transfer effects showed less model-based ("goal-directed") reasoning in the two-step task.

Human transfer has recently been used to characterize neuropsychiatric disorders. In particular, it has been used as a part of a wider array of tests to assess dysfunctions in goal-directed action in schizophrenia (Morris et al., 2014). Huys et al. (2016) tested patients with a major depressive disorder using a transfer paradigm that included both appetitive and aversive CS’s, as well as instrumental responses that consisted in either approach or withdrawal from a stimulus. As in previous studies (Huys et al., 2011), healthy controls exhibited a transfer effect in which appetitive CS increased instrumental approach responses and decreased withdrawal response, whereas the aversive CS did the opposite. This "action specificity" was markedly reduced in depressed subjects and individuals with higher "action specificity" showed better recovery over the follow-up period (4-6 months). In Geurts et al. (2013b) transfer has also been investigated as a possible link to neuropsychiatric disorders associated with aggression.

While the focus of this review is on appetitive transfer, we note here that human transfer studies have also employed aversive transfer paradigms, i.e. the ones above mentioned Huys et al. (2011); Geurts et al. (2013b); Huys et al. (2016) or also Rigoli et al. (2012); Geurts et al. (2013a); Lewis et al. (2013). Lastly, as already reported in Section 2.4, using a human transfer paradigms Pool et al. (2014) found that acute stress enhanced transfer and Quail et al. (2016) found a correlation between DASS scores (Depression Anxiety and Stress Scale) and an enhanced cue-driven response vigour.

2.8 Theoretical aspects of transfer

Over many years, various theories of the general and, more recently, the specific transfer effect have been proposed. It has, for example, been argued that a Pavlovian CS affects instrumental actions by changing the animals’ motivational state (Estes, 1943; Rescorla & Solomon, 1967); e.g., a CS associated with food could elicit a motivational state associated with hunger and thus promote instrumental actions towards food. However, this view does not explain specific transfer effects. The presence of both general and specific transfer can, however, be reconciled by adopting the Konorskian view (Konorski, 1967) that Pavlovian training can lead to two types of associations, one between the CS and the motivational/affective qualities of the US and one between the CS and specific sensory features of the US. This dual view fits the general/specific dichotomy, with
specific transfer reflecting Pavlovian associations of the CS with more specific features of the US and general transfer the CS-motivational/affective state associations.

The associative-cybernetic model, reviewed more extensively elsewhere (Cartoni et al., 2013), is probably the most comprehensive model of transfer to date as it suggests mechanisms for both specific and general transfer and also a neural implementation (Balleine & Ostlund 2007, see Figure 2.4). In brief, it posits a S-O, O-R chain through associative and S-R memories to explain specific transfer and also a general enhancing of all instrumental actions through the association of stimuli with rewards. However, the model has some shortcomings and, in its current form, does not explain all transfer data. In particular, it does not explain why, in the full transfer paradigm, no general transfer is observed in either the same or different conditions. In its most recently described form (i.e., Balleine & Ostlund 2007) the model appears to predict that in both these conditions the CS will elicit at least some general transfer: in the “same” condition this effect would be on top of the specific transfer effect, whereas in the different condition it should emerge alone as increased responding over baseline. Instead, lesioning the specific transfer circuit in the same condition does not reveal a residual general transfer component, whereas the different condition usually shows neither general nor specific transfer nor is responding in this condition influenced by any known neural manipulation. There is, currently, no mechanism in the model that can explain the lack of general transfer in these cases. It may, however, be possible to reformulate it in the light of recent data revealing the effects of inhibitory conditioning on transfer (e.g., Laurent & Balleine 2015) to include specific ‘no outcome’ representations in the associative memory, which could result in more specific outcome predictions and so less general transfer. However, this re-formulation has yet to be tested quantitatively.

As an alternative theory, van den Bos et al. (2004) have proposed a behavioral chaining account of transfer. According to this proposal, CSs elicit magazine visits (due to Pavlovian training) and these in turn elicit lever presses due to a lever-magazine-lever behavioral chain that they argue is established during instrumental training. Under this hypothesis we would, therefore, expect a positive correlation between magazine entries and lever pressing because these responses would usually occur together one after another. However, this explanation does not seem to fit with those experiments that have presented both lever press and magazine approach timing data (e.g. see Figure 6 of Holland & Gallagher 2003). Magazine approach does not seem to correlate with lever pressing; indeed it has usually been found to interfere through competition and so, in many cases, to be the inverse of lever pressing.
Figure 2.4: Associative-cybernetic model by Balleine & Dickinson as reported by Balleine & Ostlund (2007). 
a) Pavlovian CS associations (S-O) are stored in the associative memory and they can produce transfer using two pathways. On one side, each CS can prime the representation of its associated reward as an antecedent of a specific response: i.e. S1 primes $S^{O1}$ in the S-R memory which then activates R1 (specific transfer). On the other hand, they also generate an expectancy of reward through connections to the Reward memory which can then enhance all responses (general transfer, using the connection from Rew to all the Motor responses).
b) Suggested neural substrates underlying each part of the model Reprinted with permission.
Another alternative explanation of transfer has been proposed by Cohen-Hatton et al. (2013). These authors argued that transfer is due to CS-R associations formed during the instrumental and Pavlovian training sessions. They suggested that when Pavlovian training follows instrumental training, the experience of O also evokes the previously learned response R, due to the previously learned R-O association. As O and its evoked memory of R follow the CS, a CS-R association can also be formed alongside the CS-O association. Conversely if instrumental training follows Pavlovian training, it is the CS that is evoked in memory by the experience of O, due to the previously acquired CS-O association, allowing, again, an association to be formed between the CS and R. Thus, despite the two types of training being conducted in separate sessions, the evoked memories establish, on this account, associations between the Pavlovian stimuli and the instrumental responses. These associations would then lead to (specific) transfer on test, not through the integration of Pavlovian and instrumental conditioning on test but due to their integration during training. However, data from Gilroy et al. (2014) goes against this hypothesis. In their experiment one group of rats received instrumental training on two levers in a different context for each lever, whereas another group was trained with both levers in both contexts. Pavlovian training was always conducted in a third context. When given the transfer test in the Pavlovian context, the group of rats that received their training on the two levers in different contexts displayed almost no specific transfer compared to rats that received their instrumental training in both contexts. There is no obvious reason why making lever-training context specific would change the ability of evoked memories to form the S-R associations advocated by Cohen-Hatton et al. (2013) to explain transfer. Nor, on this account, is it clear why presenting the US and CS in a backward relationship separated by a 10-sec interval should reverse the specific transfer effect, but not when presented without an interval, as Laurent & Balleine (2015) reported. Cohen-Hatton et al. (2013) argue that such an effect might rely on an inhibitory connection between S and R developing over the delay. But such an effect, whilst reducing the performance of that specific response, does not imply that the performance of other actions should increase above baseline as Laurent & Balleine (2015) demonstrate; the reversal of the standard transfer effect that allows an inhibitory CS to drive performance of alternative actions above baseline performance is simply not predicted by their account.

On the other hand, the context-specificity of transfer found by Gilroy et al. (2014) might be in agreement with proposals suggesting that CSs enhance instrumental responding by virtue of their predictive value (Hogarth et al., 2013a, 2014; Hogarth &
In particular, Hogarth has proposed that specific transfer works by enhancing the R-O relationship in a hierarchical manner: CS-(R-O). This account suggests that even though CSs are trained in separate sessions and have no veridical hierarchical relation to the instrumental schedule, they still act like discriminative stimuli that signal when a specific R-O relationship is in effect. In other words, the subject builds a hierarchical representation even if it is not warranted by the procedure. Indeed, this account is also in agreement with Lovibond (1981) results we cited earlier, where the transfer effect was particularly evident when the instrumental schedule was signalled as being in extinction (S-) compared to when it was signalled to be active (S+).

This kind of account is also in line with the Gilroy et al. (2014) results if we think that, in the differential group (i.e., the group in which the two levers were trained in different contexts), the specificity of training supports the ability of the rats to distinguish when a lever is active compared to the non-differential group (i.e., in which both levers were trained in both contexts). Thus, it should be expected that a CS, as a discriminative stimulus, will provide greater benefit to the non-differential group. Indeed, the CS produced a robust specific transfer effect in the non-differential group; in contrast, in the differential group, specific transfer was observed only when testing was conducted in the instrumental context in which the lever was not trained and so where the action and the CS predictions with regards to the outcome were no longer identical.

Besides developing an understanding of the mechanisms of transfer, we think it will be important to explore and develop theories as to why transfer developed in the first place: i.e., its adaptive function. In Cartoni et al. (2013) we noted that specific transfer, general transfer and the inhibition of general transfer can be related to three aspects of action, respectively: efficacy, utility and context. In all three cases the information provided by cues is used to better evaluate which action to select. Specific transfer is related to the efficacy of the actions, namely the opportunities that they have to reach their goal. This is compatible with the interpretation of CSs as discriminative stimuli providing information as to when an action is effective or not (Hogarth et al., 2014). In particular, the model proposed that CS’s increase the estimated probability of receiving the outcome associated with an action (Figure 2.5). This led us to the hypothesis that instrumental actions having a 100% chance of receiving a reward (i.e. continuously reinforced) should not be enhanced through specific transfer since the probability of reward is already at maximum. In Cartoni et al. (2015) we tested this hypothesis, showing that indeed specific transfer was reduced when instrumental actions were trained with continuous reinforcement (compared to an RR3 schedule), albeit it was not reduced to zero as
2 Appetitive Pavlovian-Instrumental Transfer: A Review

Figure 2.5: Bayesian transfer model by Cartoni et al. (2013). In this model, both Pavlovian and instrumental conditioning are represented in terms of latent causes learning. The model is formed by a belief network, where latent causes are represented by nodes labeled with H (H1, H2, H3, H4), while CS are labeled S1, S2, S3 (sounds) and US with F1, F2 (foods); the A node represents the action of pressing a lever. During training the subject associates different latent causes to the different CS-US and lever-outcome pairings experienced. During the transfer test, the different latent causes interact producing the various transfer effects: a) in the presence of a CS paired with the same food F1 as the lever, both latent causes, the one learned in the Pavlovian phase and the one learned in the instrumental phase, will increase the probability that food F1 will be delivered, thus increasing the expected efficacy of the action (the specific transfer effect); b) due to the alternating training between the lever for food F1 and the lever for food F2, the subject has learned that the latent cause H4, associated with lever L1, actually diminishes the probability of food F2 being delivered (dashed line), thus inhibiting the effects of a CS paired with F2 (inhibition of general transfer); c) when a CS paired with another food F3 is displayed, along with the presence of lever L1, two foods are predicted, raising the value of acting in the present (the general transfer effect). Histogram from Corbit & Balleine (2011). Reprinted with permission from Cartoni et al. (2013).
predicted. In the model, general transfer is instead related to the utility of performing actions: i.e., the amount of rewards available by performing actions. In this case the CSs are supposed to signal the presence of additional resources in the environment thus promoting general activity to try to achieve them. Inhibition of transfer is related to the ability to take context into account and inhibit general transfer when those additional resources cannot be achieved. This might be merged with how specific transfer seems to work, as noted above in relation to Gilroy et al. (2014) where specific transfer signals are also dependant on the context. How these functions are mediated, both at the algorithmic levels (which variables are involved) and at the neural level remains an open question.

2.9 Conclusions and future directions

In this paper, we have reviewed a wealth of data on Pavlovian-instrumental transfer: we have seen how we can distinguish two types of transfer (specific and general), their relative neural substrates, and many of the factors interacting with them.

From the behavioral point of view, the paradigm has been improved to distinguish specific and general transfer and it is now applicable to human participants as well. Its reliability has grown sufficiently that it is proving useful to characterize pathologies as well, as we have seen for schizophrenia, addiction, and major depressive disorders (e.g. Morris et al., 2014; Garbusow et al., 2016; Huys et al., 2016).

Some questions remain open: as an example, the influence of devaluation procedures on transfer have had mixed results, possibly due to the different devaluation procedures used. In rodents, satiation abolished general transfer (Corbit et al., 2007; Aitken et al., 2016), but pairing with illness had no effect (Holland, 2004). In humans, specific transfer too had mixed results with “cognitive” devaluation able to cancel or reduce transfer (Allman et al., 2010; Eder & Dignath, 2016a) whereas satiation did not affect it (Watson et al., 2014). Also, aversive devaluation was effective only if coupled with immediate consumption (Eder & Dignath, 2016b).

As for the neural substrates, a number of areas, ranging from amygdala to the striatum and prefrontal cortex have been implicated in transfer. What is lacking at the moment is a system-wide view of how these areas interact together to produce the two types of transfer. So far a more detailed picture has been achieved locally for the Nacc shell and its projection to ventral pallidum (Laurent et al., 2014a; Leung & Balleine, 2013, 2015). Assessing the interaction with DMS and DLS might be particularly interesting.
because they are recognized as fundamental areas for the two kinds of instrumental action: goal-directed and habitual actions respectively (Balleine & O’Doherty, 2010). A point of integration between Pavlovian and instrumental learning might also be the corticothalamic circuit, as suggested in Balleine et al. (2015). Establishing the site(s) of integration would provide the basis for exerting control over the way Pavlovian predictive learning affects choice, providing considerable insight.

On the theoretical side, a question remains as to the main functional variables on which transfer acts. For example, is specific transfer related to the evaluation of the efficacy of an action, as proposed by Hogarth et al. (2013a); Cartoni et al. (2013)? We have seen that one of the latest theories favours a hierarchical account of transfer (Hogarth et al., 2014; Hogarth & Troisi, 2015), where a CS signals the availability of a R-O relationship in the environment. This is also compatible with recent accounts of biconditional discrimination (Bradfield & Balleine, 2013) which point to the formation of such hierarchical associations S-(R-O). This would make the CS in transfer, or at least in specific transfer, a special form of discriminative stimulus, even if not explicitly trained as such, not being present during the instrumental training sessions.

On the general transfer side, an important open issue is why in some cases no “general motivating effects” are observed despite the presence of the CS: i.e. in the same or different conditions. There appears to be some form of inhibition present that has not yet been well specified (Corbit & Balleine, 2005, 2011). Laurent & Balleine (2015) show that rats do not only learn which action lead to which outcome (positive R-O associations) but also which action does not earn which outcome (inhibitory R-O associations). This points to the possibility that, with more than one action, inhibitory R-O associations develop and it is these that allow animals to segregate the effects of the CS’s on same and different actions. At the functional level, we have proposed that general transfer is manifest only when the CS signals additional achievable resources: so in the same condition general transfer is not present because the outcome is already predicted by the instrumental action, whereas in the different condition it might be the alternating training with the different lever that inhibits general transfer; the signalled outcome is not achievable in that situation (Cartoni et al., 2013). However, the mechanism of this inhibition remains to be discovered. We have reviewed some manipulations that lead to a “generalization” of the transfer effect (e.g. Glasner et al., 2005; Corbit & Janak, 2007a,b; Shiflett, 2012): if we interpret and investigate these as a “disinhibition” of general transfer perhaps they will provide some indication as to how the inhibition of general transfer is normally achieved. Indeed Corbit et al. (2016) have shown that
the “generalized” transfer exhibited by an alcohol CS can be reduced by inactivating nucleus accumbens core rather than the shell, so this is consistent with the idea of the “generalization” being mediated by the same circuitry as general transfer.

On a higher level, we actually need more experiments on general transfer to establish whether it is purely a motivational phenomenon or whether it is mediated by a general representation of the appetitive outcome. In this context we could return to the issues surrounding motivational control of transfer particularly the irrelevant incentive effect (Dickinson & Dawson, 1987; Balleine, 1994) and establish whether those motivational effects are performance related or learning related. Perhaps the animal learns to class the outcome as a motivational type (e.g. food or fluid) or perhaps the energetic effects of stimuli associated with foods and fluids are simply gated by motivational state. There just have not been enough experiments on general transfer. We do not know much about the neural system mediating this form of transfer. Although CeN and NACcore are involved, are they connected in some way? No disconnection study between these structures has been conducted yet.

Another direction for future research would be to study transfer with more actions or chains of instrumental actions. Most studies on specific transfer have largely used two actions or choice between two transfer actions. It would be interesting to see how things function in three or four action situations, which would also be a more ecological setting since we are often confronted with more than just two options. Using a two-step instrumental chain, Corbit & Balleine (2003a) found that the proximal element of the chain (but not the distal) were influenced by the CS in an outcome specific manner. Can general transfer affect instrumental chains? It would also be interesting to examine chains of different lengths or homogeneous vs. heterogeneous chains.

Whatever future studies are conducted, the experiments and phenomena reviewed here show that the increasing sophistication and reliability of the Pavlovian-instrumental transfer paradigm, its application to human participants, in both normal and pathological conditions, and its close connection with fundamental learning processes, such as instrumental and Pavlovian conditioning, will continue to make it a promising area of research in the years to come.
This chapter presents a first model on PIT and an hypothesis on its function. This chapter has also appeared as an article: Cartoni et al. (2013).

3.1 Abstract

Pavlovian conditioned stimuli can influence instrumental responding, an effect called Pavlovian-instrumental transfer (PIT). During the last decade, PIT has been subdivided into two types: specific PIT and general PIT, each having its own neural substrates. Specific PIT happens when a conditioned stimulus (CS) associated with a reward enhances an instrumental response directed to the same reward. Under general PIT instead, the CS enhances a response directed to a different reward. While important progress has been made into identifying the neural substrates, the function of specific and general PIT and how they interact with instrumental responses, are still not clear. In the experimental paradigm that distinguishes specific and general PIT an effect of PIT inhibition has also been observed and is waiting for an explanation. Here we propose an hypothesis that links these three PIT effects (specific PIT, general PIT and PIT inhibition) to three aspects of action evaluation. These three aspects, which we call “principles of action” are: context, efficacy, and utility. In goal-directed behavior, an agent has to evaluate if the context is suitable to accomplish the goal, the efficacy of his action in getting the goal and the utility of the goal itself: we suggest that each of the three PIT effects is related to one of these aspects of action evaluation. In particular, we link specific PIT with the estimation of efficacy, general PIT with the evaluation of utility and PIT inhibition with the adequacy of context. We also provide a latent cause Bayesian computational model that exemplifies this hypothesis. This hypothesis and the model provide a new
framework and new predictions to advance knowledge about PIT functioning and its role in animal adaptation.

3.2 Introduction

Pavlovian conditioned stimuli (CS) associated to a reward can affect instrumental responding towards the same or a different reward. This effect is called *Pavlovian-instrumental transfer* (PIT). As an example, in a typical experimental scenario, a rat is trained to associate a sound (CS) with the delivery of food. Later, the rat undergoes an instrumental training where it learns to press a lever to get some food (without the sound being present). Finally, the rat is presented again with the opportunity to press the lever, this time both in the presence and absence of the sound. The results show that the rat will press the lever more in the presence of the sound than without, even if the sound has not been previously paired with lever pressing. The Pavlovian sound-food association learned in the first phase has somehow transferred to the instrumental situation, hence the name “Pavlovian-instrumental transfer”.

In recent years, this effect has been further subdivided into specific and general PIT. Specific PIT happens when the CS is paired with the same reward of the instrumental action. Instead, general PIT happens when the CS is paired with a different reward. In both cases, the presence of the CS leads to higher instrumental responding, however, different neural substrates are involved. Specific PIT involves the basolateral amygdala and the nucleus accumbens shell. General PIT involves central amygdala and the nucleus accumbens core (Corbit & Balleine, 2005, 2011). While most of the Pavlovian-instrumental transfer experiments have been done with rats, PIT is also present in humans and seems to involve the same brain structures (Prévost et al., 2012). Despite these advances in associating PIT effect to specific brain areas, the specific neural mechanisms underlying it are still unknown.

At the functional level, the picture is not fully clear either. Both Pavlovian and instrumental learning are often thought about in associationist terms. In associationist terms, Pavlovian conditioning leads to learning stimulus-outcome associations while instrumental conditioning can lead to associations between responses and their outcomes. One straightforward way of explaining PIT could be then in terms of a stimulus-outcome-response (S-O-R) chain. According to this view, during Pavlovian learning the subject learns a stimulus-outcome association (S-O); while during instrumental training both a R-O (response-outcome) association and its inverse O-R (outcome-response) association
are learned. In the PIT test phase, hearing the sound (S) triggers the activation of the food outcome representation (O) thanks to the S-O association; this representation in turn activates its associated response through the O-R association, thus increasing instrumental responding. In general PIT, however, the outcome in the S-O association is not the same outcome of the O-R association. This case is thus explained by referring to the general motivating properties of a rewarding outcome instead of its specific sensory properties, so that the CS presence can still enhance instrumental responding, even if the CS is associated to a different reward. However, the S-O-R chain and the general motivation explanations leave some issues unresolved. For example, they do not explain why there is no general PIT effect when the conditioned stimulus is associated with a reward given by a different instrumental response than the one currently available (see paradigm in section 3.3). In this case, one would expect that, as the CS-evoked reward is different compared to the one currently available through instrumental action, a non-specific (general) PIT effect should happen. The absence of any enhancing PIT effect in this particular condition is currently attributed to a non-well defined inhibitory effect (Corbit & Balleine, 2005, 2011). Moreover, the S-O-R chain and general motivation explanations indicate what “computation” the agent is doing but they do not say why the agent is doing it.

Our proposal is that, in goal-directed behavior, each of the three PIT processes (specific PIT, general PIT and PIT inhibition) plays a role in the evaluation of different aspects of the action (principles of action). An agent that wishes to act to accomplish a goal needs to take into account at least three aspects: context, efficacy and utility. The principle of context means that an action needs the right context to reach its goal – e.g. it is useless to call for a waiter if you are not at a restaurant. Efficacy is the probability of reaching a goal: not all actions are guaranteed to accomplish a goal. For example, buying a single lottery ticket as few chances of success. Utility means that the result of the action can be more or less valuable, depending on the state of the agent. For example, pressing a lever for food has a high value if the animal hungry, less or no value if it is sated. According to our hypothesis, an animal considers all these three aspects when it chooses which action to perform. The three PIT effects (specific PIT, general PIT and PIT inhibition) are each related to one of these aspects. We will link specific PIT with efficacy (chances of success), general PIT with utility (value of the future state) and PIT inhibition with context (availability of certain rewards). We will also propose a Bayesian computational model that exemplifies our hypothesis. The model will build upon previous work in the Pavlovian literature that conceptualizes Pavlovian
conditioning as latent-causes learning. The interplay of latent-causes (which can also be thought as contexts) will affect the number of available rewards and their probability, capturing the three PIT processes as in our hypothesis. Our hypothesis and the model together provide a new framework and new predictions to advance knowledge about PIT. Our proposal is strongly tied to goal-directed behavior. Indeed we think that it is important to study PIT because it offers a peculiar window of observation on both Pavlovian and instrumental processes. Hopefully future work on this line of research will improve not only our understanding of PIT but of Pavlovian and instrumental processes as well.

In the following sections, we will first review the current experimental paradigm used to distinguish specific and general PIT (section 3.3). We will then discuss some issues with current explanations of PIT (section 3.4). After that, we will explain our hypothesis (section 3.5) and provide a computational model for it (section 3.6). In the final section we will discuss new predictions and limitations of the hypothesis and draw some final conclusions (section 3.7).

Figure 3.1: Left: PIT paradigm to distinguish specific and general PIT. Right: histogram showing typical results of the test phase, in terms of lever press frequency. * Responding is significantly higher in Condition A and C. Histogram adapted with permission (Corbit & Balleine, 2011).
3 The three principles of action: a Pavlovian-instrumental transfer hypothesis

3.3 PIT paradigm

As noted in the introduction, PIT is not a unitary process. The experimental paradigm to distinguish specific PIT and general PIT was introduced by Corbit & Balleine (2005) (see Figure 3.1). It involves three phases: a Pavlovian phase, an instrumental phase and a test phase. In the Pavlovian phase, three sounds are associated with delivery of three different foods. In the instrumental phase, the rat undergoes two separate trainings: in each of these two trainings it learns to press a lever for a different food. The rewards used for the instrumental phase are two of the foods previously used in the Pavlovian training. A non-continuous reinforcement schedule is used, so the relationship between pressing the lever and obtaining food is probabilistic. For example, if a random-ratio RR20 schedule is used, the rat will get the reward about once every 20 lever presses. In the final phase PIT is tested. The rat is presented with one of the levers and each of the three sounds are played separated by some interval. This phase is done in extinction (no delivery of food) to prevent further reinforcement learning that could confound the results. When there is no sound, the rat will press the lever with a certain frequency (baseline). During two of the three sounds it will press the lever more than the baseline (PIT effect). Specifically, it will press more when it hears either the sound associated with the same reward of the lever (specific PIT) or the sound associated with a reward not used in the instrumental phase (general PIT). However, the sound associated with the reward of the other lever will not augment instrumental responding to the tested lever. Something is preventing PIT to be expressed in this latter case: we will call this effect **PIT inhibition** (we will return on this point in section 3.4).

By using this paradigm, Balleine and colleagues have been able to identify the different neural substrates underlying specific and general PIT. Lesions to *nucleus accumbens shell* or *basolateral amygdala* eliminate specific PIT, while lesions to *nucleus accumbens core* or *central amygdala* eliminate general PIT (Corbit & Balleine, 2005, 2011). These substrates have also been confirmed by using inactivation and disconnection procedures (Corbit & Balleine, 2011; Shiflett & Balleine, 2010). Moreover, by using this paradigm a further important difference between specific and general PIT has been found: general PIT is subject to devaluation, while specific PIT is not. To show this, Corbit et al. (2007) executed the test phase of the paradigm after sating rats and they found that specific PIT was still present while general PIT disappeared (Figure 3.2). That is, general PIT is affected by the devaluation of food by satiation, while specific PIT is not. There have been previous reports about PIT being unaffected by devaluation but a different
paradigm was used and no lesions were performed, so the type of PIT evoked (specific or general) in those experiments can only be inferred (Rescorla, 1994b; Holland, 2004). In Holland (2004) some of the reported results can be interpreted as specific and general PIT being both unaffected by devaluation, in contrast with results from Corbit et al. (2007). We will discuss this contradiction in the final section.

### 3.4 Explaining PIT: current issues

As mentioned in the introduction, specific PIT can be thought in terms of a S-O-R chain: the sound stimulus evokes an outcome (food) and that food in turn evokes the associated response (pressing the lever). In the case of general PIT instead, the sound evokes food and the reward of food exerts a general motivational effect on instrumental responses. As an example of this kind of explanations, we will look at the associative-cybernetic model of Balleine & Dickinson as reported by Ostlund & Balleine (2007b). This is, to our knowledge, the most complete model in the PIT literature. The associative-cybernetic model is actually not just model of PIT, but a more general model of both Pavlovian and instrumental processes. However it does include a way to explain both specific and general PIT functioning.

In the model, Pavlovian stimulus-outcome associations are represented in the “associative memory” component (see Figure 3.3). According to Balleine & Ostlund (2007), during instrumental learning, both R-O and O-R association are learned. In fact, food is not only the result of pressing the lever, but it also precedes the next lever pressing. So within the “S-R memory” component, food is considered as a stimuli \((S)\) that precedes the response and thus \(S^O - R\) associations are learned. When the Pavlovian sound stimulus is encountered, it activates its representation in the associative memory (e.g. \(S^1\)). In turn, its representation activates its associated food outcome as a stimulus in S-R memory \((S^O^1)\). This representation then activates the corresponding instrumental response \((R^1)\). This “pathway” in the model can thus explain how specific PIT works and it is basically a S-O-R chain explanation.

However, there is a second pathway in the model through which Pavlovian stimuli can influence instrumental responses. Assume that the rat now hears a different sound \(S^2\) associated with a different food outcome \(O^2\). The specific PIT pathway will not work because this food is not associated to \(R^1\) in the S-R memory. However, the sound \(S^2\) activates \(O^2\) in the associative memory. This activates the corresponding \(O^2\) in the “reward memory” component, which in turn activates the reward node (Rew). We
Shifting the motivational state from hunger to satiety (food devaluation) causes a general drop in instrumental performance and the disappearance of general PIT. Reprinted with permission (Corbit et al., 2007).
can think the Rew node as an “expectation of value”. This expectation of value (Rew) can exert an excitatory effect on all motor responses (arrows from Rew to all responses $R_1..R^n$). Thus the sound stimulus can evoke an expectation of value through which it can aspecifically enhance even an instrumental response associated with a different food. This effect can be assumed to be a model of general PIT. The fact that general PIT works through an expectation of value is also consistent with the fact that general PIT (and not specific PIT) is sensitive to devaluation.

Even though the associative-cybernetic model can explain the existence of both specific and general PIT and their interaction with devaluation, some important aspects in PIT experimental data remain without answer and suggest that PIT phenomena need a more complex explanation. A first important aspect involves the absence of general PIT in the *same* condition. In the *same* condition, a CS associated with a certain food enhances an instrumental response directed to the same food. One might expect that, according
3 The three principles of action: a Pavlovian-instrumental transfer hypothesis

Figure 3.4: PIT effects after lesions of basolateral amygdala (BLA) and central nucleus of amygdala (CN), compared to a sham lesion group. The histogram shows PIT effects as frequency of lever pressing in the presence of CS minus lever frequency in the absence of CS (baseline). The control group (sham) shows specific PIT effect in the same condition (CS associated with the same reward as the instrumental response) and general PIT effect in the general condition (CS associated with a reward not used in the instrumental phase). Notice the absence of any PIT effect in the different condition when the CS is associated to the reward of the lever not available during the test. BLA and CN lesions eliminate specific PIT and general PIT respectively. Reprinted with permission (Corbit & Balleine, 2005).

...to the associative-cybernetic model, the CS would also elicit general PIT, using both pathways at the same time. That is, the CS could elicit PIT both through the S1-S°1-R1 pathway and through the S1-O1-O1-Rew-R1 pathway (Figure 3.3). In other words, one might expect that the CS can have both a specific PIT effect, as it evokes the food associated with the instrumental response, and a general PIT effect at it evokes an expectation of value (food) that can motivate instrumental responses aspecifically. Experimental data shows that this is not the case. Figure 3.4 shows that BLA lesions, belonging to specific PIT circuit, eliminate all PIT effect in the same condition. Morover, lesion of central amygdala, which belongs to the general PIT circuit, does not have any effect in the same condition. This means that in this condition only specific PIT is expressed and there is no general PIT.

There is another case where the explanation of the associative-cybernetic model is incomplete. In the data shown in Figure 3.4 we can notice that in the different condition there is no PIT effect (neither specific nor general). The different condition corresponds to the case where the CS is associated with the reward of the other, not-available, instrumental response – that is, the reward of the lever not used during that test. One would expect that, as the CS is not associated to the same reward of the available...
3 The three principles of action: a Pavlovian-instrumental transfer hypothesis

instrumental response, it cannot elicit specific PIT but it would still elicit general PIT. On the contrary, there is no visible PIT effect: the instrumental response stays at baseline level, even after lesions to the specific or general PIT circuit (see Figure 3.4). The associative-cybernetic model does not offer an explanation for this. There is no connection in the model that can explain why the aspecific effect of the “general PIT pathway” should be inhibited towards some instrumental responses. Even those articles that contain experimental data about this absence of general PIT simply suggest that there must be some kind of inhibitory effect (Corbit & Balleine, 2005) or they suggest possible neural locations of this inhibitory effect (Corbit & Balleine, 2011), but without exactly explaining its presence. So there is an inhibitory effect, capable of suppressing general PIT, waiting to be explained. This effect is not always of the same entity: in some cases it is possible to have some kind of “partial” PIT effect in the different condition, as shown in Figure 3.5. This means that the inhibitory effect has not been able to completely suppress general PIT. We will return on this in the final section, suggesting one possible way on how it might happen.

Figure 3.5: PIT test phase in Corbit & Janak (2010b). The control group in the figure displayed a noticeable PIT effect even in the different condition, albeit significantly less stronger than the one in the same condition. Groups with dorsal striatum lesions also showed similar PIT effects of similar size in the different condition (not shown here). Adapted with permission (Corbit & Janak, 2010b).

Finally, there is a third kind of data that asks for a different explanation than the two simple pathways. At the beginning of this century, a contradiction in the PIT literature arose. Blundell et al. (2001) reported that lesions to the basolateral amygdala affected PIT, while Hall et al. (2001) reported that lesions on central amygdala, and not basolateral amygdala, abolished PIT (see Figure 3.6). This contradiction was then resolved in Corbit & Balleine (2005) by showing that there exist two kinds of PIT, one depending on basolateral amygdala (specific PIT) and one depending on central amygdala (general PIT). However, one question remained open: why did Hall et al.
procedure elicit general PIT (as shown by the sensitivity to central amygdala) instead of specific PIT as Blundell et al. (2001)? Hall et al. (2001) used a single CS and a single instrumental response, both associated to the same reward. Blundell et al. (2001), instead, used two CS and two levers with two different rewards. In both cases an interval reinforcement schedule was used. Given that Hall et al. (2001) used the same food for both the CS and the lever, one would expect a specific PIT, not a general PIT. Later articles usually refer to that fact simply by saying that procedures with a single lever seem to elicit general PIT instead of specific PIT (Corbit & Balleine, 2005, 2011). Our view is that it could be not just a question of the number of levers, but of the kind of instrumental response elicited: habitual versus goal-directed. It is known that interval schedules with a single lever easily lead to habitual behavior, while single-lever random ratio schedules and two-levers procedures (even using interval schedules) usually elicit goal-directed behavior (Yin & Knowlton, 2006). We suggest that the real reason underlying the lack of specific PIT effect in Hall et al. (2001) is that the elicited behavior was habitual. Our hypothesis, illustrated in the next section, will link specific PIT with the probability of reaching the goal of an action. Since in habitual behavior the action is a simple “reaction” to a stimulus and there is no goal evaluation, specific PIT cannot happen; general PIT instead, not being tied to the specific consequences of the action, might still happen. Thus, Hall et al. (2001), by eliciting habitual behavior, measured general PIT instead of specific PIT and found it to be sensitive to central amygdala while Blundell et al. (2001) observed specific PIT on goal-directed behavior and found it to be affected by basolateral amygdala lesions.

3.5 Three principles of action: context, efficacy, utility

Experimental data from the PIT test phase shows three different conditions with three different effects: specific PIT, general PIT and PIT inhibition. We will now explain our hypothesis according to which these three types of effects are functionally linked to three aspects of action evaluation during goal-directed behavior.

An instrumental action, such as lever pressing connected to food, can be either habitual or goal-directed (Yin & Knowlton, 2006). In the case of habitual behavior, the instrumental action is elicited by a simple stimulus-response association and there is no evaluation of the consequences. In the case of goal-directed behavior instead, the response is directed to a goal and it is linked to the evaluation of its consequences. These two types of behavior can be distinguished by devaluation and contingency alteration.
Figure 3.6: Mean lever presses during PIT test phase. During training a single CS and a single lever were used, associated to the same food reward. The CS enhances instrumental response compared to baseline (ITI) only in the group with basolateral amygdala (BLA) lesion and in the control groups (Sham). Even though the same reward has been used for both the CS and the instrumental response, it is the central amygdala (CeN) lesion (general PIT circuit) to prevent PIT effect and not the one to the basolateral amygdala (specific PIT circuit). Reprinted with permission (Hall et al., 2001).

procedures. In the devaluation procedure food is devalued (e.g. by satiety) while in the contingency alteration procedures the relationship between pressing the lever and the presence of food is altered (e.g. pressing the lever now stops food delivery). In the case of habitual behavior, the rat (or other subject) will keep pressing the lever, while in the case of goal-directed behavior it will stop pressing it when the ability of the action to obtain food or the value of available food are altered. That is, goal-directed behavior distinguishes itself because action consequences are evaluated, both in terms of probability of happening and value.

Our hypothesis is that during goal-directed behavior, the three aspects of PIT (specific PIT, general PIT, PIT inhibition) are linked to three aspects of this evaluation.

- **Context.** A goal-directed action must be executed in the right context (it’s useless to call for a waiter if you are not at the restaurant)
• **Efficacy.** A goal-directed action can be more or less effective in reaching its goal (buying a single lottery ticket has few chances of winning)

• **Utility.** The goal of an action can be more or less useful (getting food is useful if the rat is hungry, less or not useful if it is sated)

General PIT can be linked to the principle of utility: a CS associated with food evokes a reward in the near future. Our hypothesis is that this reward is added to the future scenario of the consequences of the action, thus enhancing the motivation to execute it. However, if the animal is sated, this added reward clearly has no longer any value, so the motivation effect vanishes. As we saw before, satiety does indeed eliminate general PIT (Corbit et al., 2007). Moreover, the motivational effect of general PIT is possible only if the reward evoked by the CS is an “additional reward” compared to the one already foreseen by the action. In agreement with this, data collected using random-ratio schedules (where the behavior is usually goal-directed) show that in the specific PIT condition (where the CS reward is the same as the lever) general PIT does not happen (Corbit & Balleine, 2005, 2011).

Specific PIT can be linked instead to the principle of efficacy. A CS associated to the same reward as the lever predicts that this reward will be present in the near future. During the action evaluation, the CS acts as a cue that there is an higher chance of getting the reward associated to the action, thus motivating the agent to pursue that action. As specific PIT is then about an increase of probability of getting the food, this effect is immune to devaluation: whether the animal is sated or not, the CS predicts an higher probability of reward, so, compared to the CS absence, the action will be evaluated more positively. In agreement with this link between specific PIT and probability, in a human study, Trick et al. (2011) have found that more predictive conditioned stimuli (those with higher probability of reward) induce a stronger PIT effect.

Lastly, PIT inhibition can be linked to the principle of context. We suggest that the presence of the lever acts as a discriminatory stimulus that inhibits the reward of the absent lever. That is, the presence of the lever signals not only that its associated reward is available, but also that it is not possible to obtain the reward of the other lever. The general PIT effect would be then inhibited by the presence of the lever associated with a different reward compared to the CS.

In other words, specific PIT represents the ability of an agent to take into account cues that indicate that a certain reward is more probable in the environment compared to when those cues are not present. The fact that those rewards are more probable
translates into a perceived higher efficacy of the action. General PIT represents instead the ability to use cues that indicate the presence of other “additional” rewards in the environment, thus motivating the agent to act as they constitute an added value (utility). Finally, PIT inhibition represents the ability to take into account the context and to inhibit rewards signals that are known to be not available through action at that time (context). In the following section, we will now try to exemplify our hypothesis through a computational model.

3.6 A Bayesian model of PIT

Bayesian modeling is increasingly used in many fields, from chemistry (Hibbert & Armstrong, 2009) to astrophysics (Loredo, 1992), from economy (Poirier, 2006) to genetics (Beaumont & Rannala, 2004). This increasing widespread use has even prompted some to call it a “Bayesian revolution” (Beaumont & Rannala, 2004). Bayesian approaches are now being used in cognitive science too: indeed, both Pavlovian (Courville et al., 2004, 2005, 2006; Gershman & Niv, 2012) and instrumental (Solway & Botvinick, 2012; Pezzulo et al., 2013) Bayesian network models have been created.

The Bayesian approach owes its name to the Bayes theorem:

$$p(h|d) = \frac{p(d|h)p(h)}{p(d)}$$

The theorem says that the posterior probability of an event $h$ given a set of observations $d$, denoted with $p(h|d)$, is equal to the probability of observing $d$ given the event $h$, denoted with $p(d|h)$, multiplied by the a priori probability of $h$, divided by the a priori probability of observing $d$. In other words, the theorem transforms the a priori probability of $h$ in a posterior probability of $h$ that takes into account the set of observations $d$. Bayes theorem tells us how to “update” in an optimal way our belief about $h$ happening given the data $d$ at our disposal. This theorem can then be used to create models about how an animal can make sound inferences about the events of the world given what it sees. Indeed, we have chosen to use this approach to simulate Pavlovian and instrumental learning and so to build a Bayesian computational model of PIT.

In the Pavlovian literature, Pavlovian learning is often thought as S-O learning, that is, as the acquisition of the association between a stimulus (S) and an outcome (O). In particular, associationist models usually focus on the predictive properties of S. Models such as Rescorla-Wagner (Rescorla & Wagner, 1972) or Pearce (Pearce, 1994) try
to explain how S comes to predict outcome O (and thus elicit a Pavlovian response). However, a different approach exists. For example, Courville et al. (2005) describes Pavlovian learning using a Bayesian generative model with hidden latent causes. In this model, the animal does not simply try to learn how often O occurs after S. Instead, the animal tries to learn the whole generative model: that is, it tries to learn the hidden cause that makes both S and O appear. Indeed, this is also a more rational strategy by the animal, as in Pavlovian experiments S does not really cause O, but the two simply occur together because of a common cause (the experimental setup). By using this model, Courville et al. explain phenomena that are otherwise not well accounted for by classical associationist models such as Rescorla-Wagner or Pearce ones (Courville et al., 2005).

Inspired by this work, we “extended” it to the instrumental realm to explain PIT. As in Courville et al. (2005), our model is a sigmoid belief network (see Figure 3.7). The network is formed by three set of nodes:

- **Observables:** nodes that represent objects such as levers (L1, L2, L3), sounds (S1, S2, S3), foods (F1, F2, F3).
- **Latent causes:** nodes that represent hypothetical causes (H1, H2, H3, H4, H5) that can explain the observation or the lack of observation of the objects in the world.
- **Actions:** a single node that represents the action of pressing a lever (A) which can influence the presence of some objects in the world.

The activations of each node in the network represent the probability that the corresponding object is present. The nodes influence each other through their links. Each link is associated to a numerical weight. Those weights can be either positive or negative (or zero). Positive weights from a node to another mean that when the parent node is active, the child node is more likely to happen. Negative weights, instead, decrease the probability of the child node to be active. For simplicity, in this model only link weights are learned, while the number of hidden causes and their links are given. Learning link weights means that the agent has to discover how the hidden causes affect the probabilities of each observable and how its action can influence the presence of food. At first, we assigned a distribution of a priori probabilities to each link concentrated on the value of zero. This means that, before learning, the agent has not yet formed a particular “belief” on how these latent causes (or its action) can affect what he sees in the world. Then
we trained the model by applying Bayesian inference on a set of simulated observations that represent Pavlovian and instrumental learning (following the paradigm described in section 3.3). As in Courville et al. (2005), we used a Monte Carlo Markov Chain method for training, in our case using WinBUGS software (Lunn et al., 2000). The resulting \textit{a posteriori} distribution of probabilities describes the animal \textit{belief} on how the world works after the conditioning sessions. The resulting model with weights based on the Pavlovian and instrumental training can account for the various PIT effects. We will now describe how the model behaves in each phase of the PIT paradigm and how it accounts for the three PIT effects: specific PIT, general PIT and PIT inhibition.

Figure 3.7: Bayesian network used to simulate PIT. There are three set of nodes: observable objects such as levers (L1, L2), sounds (S1, S2, S3) and foods (F1, F2, F3); latent causes that influence the presence of objects in the world (H1, H2, H3, H4, H5); the action of pressing the lever (A). Lever pressing is linked only to foods F1 and F2 as it can influence only their presence and not the other observables.

3.6.1 Pavlovian phase

During Pavlovian training, the rat sees that sound S1 and food F1 are correlated, so it assumes that an hidden cause H1 is generating both events. This happens for all the three Pavlovian trainings, thus generating positive weights between each of the three hidden causes (H1, H2, H3) and its pair of sound and food events (see Figure 3.8).

Figure 3.8: Pavlovian phase: positive links between latent variables and co-occurring sounds and foods are established.
3.6.2 Instrumental phase

During instrumental learning, the rat sees that the presence of a lever L1 and food F1 availability are correlated, so it assigns a positive weight to the links between H4 and L1 and between H4 and F1 (see Figure 3.9). Food delivery, however, depends also on the action of lever pressing (A), so a positive link between action and food is also formed. The same happens for the other instrumental learning with lever L2 and food F2. The rat also learns a negative association (dashed line) between the “instrumental” hidden cause and the other food: in other words, with experience it discovers that when H4 is active and lever L1 is present, lever pressing (A) will not obtain food F2.

![Figure 3.9: Instrumental phase: instrumental latent causes, which give rise to the presence of the levers, interact with action to make food available. Instrumental latent causes also inhibit each other’s food availability (negative links depicted as dashed lines).](image)

3.6.3 Test phase

The test phase involves four possible conditions depending on the presence of different sounds together with one of the levers. This will give rise to different patterns of activations in the learned model (see Figure 3.10):

**Baseline:** in the baseline condition, lever L1 is presented alone. The rat will press it with some frequency knowing from previous instrumental learning that when L1 is present, it can get food F1 by lever pressing (action A).

**Condition A – specific PIT:** from the presence of S1 and L1, the rat can infer the presence of causes H1 and H4, which both predict F1 in the near future. This motivates the rat to press the lever more than when L1 is present alone, without any sound, as there are now increased chances of getting F1 in the immediate future.

**Condition C – general PIT:** the presence of lever L1 and sound S3 implies that causes H4 and H3 might be present and that foods F1 and F3 will appear in the future. While food F3 is not a direct effect of lever pressing A, its predicted presence nevertheless motivates the rat to press the lever more than the baseline condition. This is a different
3 The three principles of action: a Pavlovian-instrumental transfer hypothesis

kind of motivation from Condition A: instead of augmenting the probability of the food targeted by the action, it adds a new food reward to the scene.

**Condition B – PIT inhibition:** this condition is similar to condition C, but in this case food F2, evoked by sound S2 (through H2), is inhibited by cause H4. Thus, only food F1 remains predicted and no enhancement is found compared to the baseline.

The model thus exemplifies how our hypothesis could work. Specific PIT arises from the interaction between Pavlovian and instrumental latent causes that results in the evaluation of higher chances of getting the reward connected to the action (*principle of efficacy*). General PIT arises from adding value (a new reward) to the predicted future scenario, acting on the *principle of utility*. The absence of general PIT in the third condition could be instead consequence of the fact that the reward predicted by the CS (S3) is excluded as it is not possible in the presence of H4 (*principle of context*).

---

3.7 Conclusions

In this article, we have first reviewed some experimental data about Pavlovian-instrumental transfer. These data suggest that PIT cannot be simply explained in terms of a S-O-R chain plus an aspecific excitatory process. At the very least, a third inhibitory process is present in the experimental data, waiting to be explained. We have suggested an hypothesis that can both explain how these three processes work and what is their function. We have linked specific PIT, general PIT and PIT inhibition to *three principles of goal-directed action*. The idea is that these three PIT processes represent the effect of conditioned Pavlovian stimuli on three aspects of the action evaluation that happens
in goal-directed behavior. These three aspects are: context (an action directed to a goal needs the right context), efficacy (an action can have more or less chances to accomplish a certain goal) and utility (the consequences of an action can be more or less useful). According to this, specific PIT represents the use of cues that indicate higher chances of getting a certain reward. Data supporting the relationship between specific PIT and reward probabilities can be found in Trick et al. (2011) where those CS associated with higher chances of reward elicited stronger PIT. Explaining specific PIT in terms of probability would also match the fact that specific PIT is not influenced by devaluation procedures (Rescorla, 1994b; Holland, 2004; Corbit et al., 2007): indeed the probability of getting a reward is independent of its value. Instead, general PIT affects utility of the action, its value, so it is subject to devaluation (Corbit et al., 2007). General PIT corresponds to the use by the agent of cues that indicate the presence of an additional reward in the immediate future and this motivates the animal to execute its actions with more vigor. Lastly, the inhibition of general PIT in some situations corresponds to the use of cues to understand which rewards are available in a given context.

We have then created a model of how this functional hypothesis could be translated into probabilistic computation. In particular, we have created a model drawing inspiration from Courville et al. (2005), where Pavlovian conditioning is explained in terms of a latent-cause generative model. By adding an action node and its links to food we have simulated instrumental learning, thus explaining the three PIT effects in terms of interactions between latent causes representing contexts learned during Pavlovian conditioning and those learned through instrumental conditioning.

The model captures the functional hypothesis about the three PIT aspects. The model produces new predictions that might be tested in future empirical research and expand our knowledge about PIT. In particular, if specific PIT is an effect of “augmenting chances of success”, then instrumental actions that already have 100% chances of success should not be able to benefit from specific PIT (but they could still benefit from general PIT). The reinforcement schedules used to test PIT are non-continuous, so pressing the lever has not 100% chance of delivering food. It will be interesting to see if under a continuous reinforcement schedule this prediction will be confirmed or not. As for the principle of context, we might expect that in an experimental procedure where the two levers could be somewhat linked to the same context instead of “excluding each other”, we should observe less or no PIT inhibition. Indeed in the experiment of Corbit & Janak (2010b) where rats were presented levers in an alternating fashion but within single sessions, results indicate a strong PIT effect in the different condition too, albeit
somewhat less than the same condition (see Figure 3.5). A specific procedure focusing on the role of context might shed further light to confirm this part of the hypothesis. As for the utility effect, our hypothesis is not enough detailed to go beyond the fact that general PIT should be subject to devaluation, as already shown in the literature (Corbit et al., 2007). We do not know if the “additional reward” needs to be a different type of food compared to the instrumental action, or if a CS that signals the same food, but in a larger quantity, could be equally effective in producing general PIT.

In our attempt to provide an hypothesis capable of explaining all three PIT processes, we have focused on results achieved with the paradigm capable of detecting all three (Corbit & Balleine, 2005; Corbit et al., 2007; Corbit & Balleine, 2011). However, those experiments involved multiple levers and a random-ratio schedule, thus evoking goal-directed behavior. This is why our hypothesis is proposed to be an explanation of how PIT processes affect goal-directed behavior in particular. However, most of PIT experimental data has been actually done using interval schedules (see Holmes et al., 2010 for a review). If those schedules have often evoked habitual responding, then a lot of data would fall outside the main focus of our hypothesis, which is limited to goal-directed behavior. That being said, we do have some suggestions on how PIT might work during habitual behavior. We have indeed suggested that specific PIT can only happen during goal-directed behavior as it pertains the chances of achieving a goal. During habitual behavior, there is no goal evaluation and thus we suggest that a CS paired with the same reward of the lever would produce general PIT instead. Then, some of the differences found in literature data could be explained by the use of different reinforcement schedules, leading to habitual vs goal-directed behavior (e.g. Hall et al., 2001 vs Blundell et al., 2001).

Beyond the contrast between Hall et al. (2001) and Blundell et al. (2001), we have another contradiction in the literature which could be mostly resolved by differences of how PIT works under goal-directed vs habitual behavior. We have mentioned above that Holland (2004) reported results that indicate that both specific and general PIT are insensitive to devaluation. This is in contrast with the results of Corbit et al. (2007) where only specific PIT is immune to devaluation. The results from Holland (2004) can be mostly reconciled by noting that he used an interval schedule, thus probably eliciting in many cases habitual behavior. In all those cases, the fact that general PIT was not affected by the devaluation procedure can be explained by the fact that under habitual behavior the baseline performance is not subject to devaluation either. In other words, we might suppose that under habitual behavior the devaluation process is “inactive”
and thus it does not affect general PIT either. This would reconcile all results from the experiments of Holland (2004) except one: one of the experimental groups displayed in fact a devaluation effect on baseline performance but not on the general PIT effect (fig. 4D in his article). That result would still conflict with Corbit et al. (2007)’s results, even after the additional assumption that general PIT might not be subject to devaluation under habitual behavior. The solution to this contradiction might lie on the devaluation procedure used. While Corbit et al. (2007) used satiety to devalue food rewards, Holland (2004) used an aversion procedure (pairings with LiCl): it might then that the two devaluation procedures affect differently general PIT.

In our computational model, the analysis is limited to how the agent can make different inferences depending on the test conditions but it does not show how these inferences are transformed into action. We need to build a more complete model that can account on how these perceived higher efficacy or additional rewards are transformed into an higher instrumental performance. Also, the model is purely functional and does not yet shed light on the neural mechanisms. Neural models might be developed to this purpose. Future developments of the model and the hypothesis should also address conditions such as use of drugs of abuse and chronic stress, which are known to have an effect on PIT (Wyvell & Berridge, 2001; Corbit & Janak, 2007a; Morgado et al., 2012).

Despite the above mentioned limitations, we think our hypothesis can give a new perspective on PIT, a new framework on which to discuss, experiment and advance our knowledge about PIT. Of particular importance is the discussion on why is there a PIT effect, i.e. its role in the animal adaptation. We have proposed that PIT relates to the ability of using signals in the environment to better evaluate the possibilities of action.
4 The Relationship Between Specific PIT and Instrumental Reward Probability

This chapter presents an experiment aimed to verify the prediction on specific transfer from the PIT model of the previous chapter, investigating the relationship between specific transfer and instrumental reward probability.

This chapter has also appeared as an article: Cartoni et al. (2015)

4.1 Abstract

Goal-directed behavior is influenced by environmental cues: in particular, cues associated with a reward can bias action choice towards actions directed to that same reward. This effect is studied experimentally as specific Pavlovian-instrumental transfer (specific PIT). We have investigated the hypothesis that cues associated to an outcome elicit specific PIT by rising the estimates of reward probability of actions associated to that same outcome. In other words, cues reduce the uncertainty on the efficacy of instrumental actions. We used a human PIT experimental paradigm to test the effects of two different instrumental contingencies: one group of participants had a 33% chance of being rewarded for each button press, while another had a 100% chance. The group trained with 33% reward probability showed a stronger PIT effect than the 100% group, in line with the hypothesis that Pavlovian cues linked to an outcome work by reducing the uncertainty of receiving it. The 100% group also showed a significant specific PIT effect, highlighting additional factors that could contribute to specific PIT beyond the instrumental training contingency. We hypothesize that the uncertainty about reward delivery due to testing in extinction might be one of these factors. These results add knowledge on how goal-directed behavior is influenced by the presence of environmental cues associated with a reward: such influence depends on the probability that we have
to reach a reward, namely when there is less chance of getting a reward we are more influenced by cues associated with it, and vice versa.

4.2 Introduction

It has long been known that cues associated with a rewarding outcome can elicit and intensify actions directed to obtain that same outcome. This effect can be studied experimentally in a paradigm called specific Pavlovian instrumental transfer (specific PIT). In a typical specific PIT experiment, a participant is first trained to associate two cues with two different outcomes: for example, to associate two different images (Pavlovian conditioned stimulus, CS) each with the delivery of a different reward (e.g. chocolate and popcorons). Then, the participant is trained to make two actions to get these two rewards: for example, to press a left button to get chocolate and to press a right button to get popcorons (instrumental training). In a final test phase, the participant can again press these buttons in extinction (no reward is delivered) but sometimes one of the two images (CS) is displayed. What will happen is that during the image display the participant will press the button corresponding to the same outcome of the image more than the other button. In other words, in specific PIT a Pavlovian cue associated with food (or another reward) selectively increases instrumental actions directed to the same food. This occurs despite the fact that no explicit training of the instrumental actions in the presence of Pavlovian cues is performed.

This “PIT effect” can play a critical role in regulating goal-directed behavior in different situations of life, ranging from advertising to drug addiction. For example, the vision of a McDonald sign might encourage a person to buy a portion of potato chips. In this case, the McDonald sign might be thought as a CS associated with potato chips that promotes the instrumental action of buying that food. PIT is also relevant for drug addiction as drug-related cues can be a threat to self-control and often lead to relapse after treatment (Hogarth et al., 2007; Belin et al., 2009).

While there is increasing progress on the study of specific PIT neural substrates and mechanisms (Laurent et al., 2014a), it is not yet clear how specific PIT works at the functional level and what its adaptive function and evolutionary significance is. Why do Pavlovian cues influence our instrumental behavior through the specific PIT effect? Which are the factors that mediate this effect? It has been proposed that Pavlovian cues elicit specific PIT by signalling an increased efficacy of the instrumental action (Hogarth et al., 2013c; Cartoni et al., 2013; Hogarth et al., 2014; Hogarth & Troisi, 2015).
In Cartoni et al. (2013), we advanced an hypothesis and a computational model on how PIT might bias instrumental behavior by affecting different components of action evaluation. In particular, we modelled one of the possibilities of how specific PIT might be linked to the efficacy of the instrumental action. We suggested that the specific PIT effect is elicited by the presence of an outcome-associated cue that increases the estimate of the probability of reaching that outcome. In other words, if an instrumental response was usually rewarded 33% of the time, the presence of the cue makes the participant think that the chances of being rewarded are now higher (e.g. 50%). This led us to the hypothesis that instrumental responses that are continuously rewarded (100% chance of receiving the reward) cannot be augmented further by the presence of a cue associated with the same reward.

In the study reported here, we tested this hypothesis on specific PIT by contrasting specific PIT effects in two groups of subjects: during instrumental training, one was rewarded with with 100% reward chance for each instrumental action (button presses), while another was rewarded only with 33% chance. To date we are not aware of any work either with humans or animals that has directly manipulated the instrumental contingency during training to see how the size of specific PIT varies as a function of the instrumental probabilities of obtaining an outcome. The closest manipulations we could find were those by Trick et al. (2011), which used Pavlovian stimuli trained with different contingencies and Hogarth et al. (2014) where the participants expectations about instrumental contingency during PIT test were manipulated by verbal instructions. However, none of these directly manipulated the reward probabilities of the instrumental training.

According to our hypothesis, we expected to find a stronger specific PIT effect in the low-probability group (33%) than in the high-probability one (100%); in this latter group the outcome probability was already at maximum, so we expected a minimal or absent specific PIT effect as the estimate of the probability of being rewarded could not be further augmented by the cues. Experimental results confirmed that specific PIT was stronger for participants trained with a lower probability to obtain the outcome by instrumental action. However, a significant specific PIT effect was also found in the high-probability group, despite the fact that the trained reward probability was at maximum. We hypothesize that the uncertainty about reward delivery due to testing in extinction might account for this latter effect.
4 The Relationship Between Specific PIT and Instrumental Reward Probability

4.3 Material & Methods

Participants

A sample of 57 volunteer students (32 males) between the ages of 19 and 30 years (mean age = 24.0, SD=2.8) were recruited from the Sapienza University of Rome. Following Prévost et al. (2012), the eating attitudes test EAT-26 Garner et al. (1982) was administered before the experiment to ensure that participants did not have eating disorders. Participants’ last meal was on average 2.6 hours (SD=2.7) before the experiment start. Written informed consent was obtained from all participants, and the Psychology Ethics Committee of the Sapienza University of Rome approved the study. In Prévost et al. (2012), using a paradigm close to the one used here, PIT effects were detected with a sample size of 26 participants. Other studies also used around 20 participants Allman et al. (2010); Trick et al. (2011); Bray et al. (2008) to detect PIT effects with humans, so we chose as a stopping rule to collect at least 20 participants for each of the two conditions (33% and 100% probability), after considering exclusions due to EAT-26 results or the Assessment Phase. From the initial sample of 57 participants, five were excluded as their results on the EAT-26 suggested a possible eating disorder. Eight participants were also excluded as they failed to answer correctly the questions on the Pavlovian and instrumental reward associations during the Assessment phase. It has already been reported that participants unaware of the contingencies might not express specific PIT, so it is customary to exclude them from the analysis (see Trick et al., 2011; Hogarth et al., 2014; Eder & Dignath, 2016b). We then further excluded 4 participants as outliers after analyzing instrumental training data because they focused almost exclusively on one lever during the instrumental phase. These exclusions left 40 participants for the PIT test analysis, 23 participants for the 33% condition plus 17 participants for the 100% condition.

Stimuli and materials

Visual stimuli were presented by a display (27X40cm) connected with a computer; stimulus presentation and behavioral data acquisition were implemented in Matlab (The Mathworks) with the Psychophysics toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007). The food rewards were chosen from 14 different sweet and salty snack foods (Bounty, Cipster, Fonzies, Freeky Fries, Kinder bueno, Kinder cereali, Kinder cioccolato, Kit kat, Mars, Milka, Ritz Crispy, Smarties, Tuc, Twix). Participants were asked
to provide subjective pleasantness ratings for each snack food clicking with a mouse on an analogue scale displayed together with a picture of each snack. The two most pleasant foods for each participant were used as rewards. Two fractal images were employed as conditional stimuli (CS) during the Pavlovian and PIT test phases. Instrumental responses consisted of button presses on a custom response box. The response box had three buttons arranged in an horizontal row each equipped with a spring to make pressing effortful. Only the left and right buttons of the response box were used. Two black squares were presented on the display: these squares became gray to signal when response buttons were available for pressing during each trial. When participants pressed a button, the squares briefly flashed white. Food pictures (rewards) and fractals were displayed between the squares.

**Procedure**

The experiment consisted of three main phases: an instrumental, a Pavlovian, and a PIT phase (see Figure 4.1). These phases were preceded by a “Taste test” phase and followed by an Assessment phase.

**Taste test phase** During this phase, the two squares are black, signaling that no button is available. There are two trials, one for each food chosen for the experiment. On each trial a food image appears and the experimenter then gives a piece of the corresponding food in the hand of the participant for immediate consumption. This phase gives the participants a chance to experience the two foods in a hungry state. This should enhance their motivation for pressing in the subsequent phases and ensure that they “know the value of the reward” in the hungry state.

**Pavlovian learning phase** During Pavlovian training, the two squares are still black, signaling that no button is available. For each trial, a fractal image is presented for 6 seconds. During the last second of this presentation (between 5 and 6 second) a food picture is displayed. Two fractals were used and each fractal paired with one of the foods. The associations between fractal images and foods were randomized across participants. This phase was formed by 20 trials, 10 trials for each fractal. A random duration inter-trial interval was presented for 2-6 seconds.

**Instrumental learning phase** The instrumental phase was formed by 30 trials, each lasting 6 seconds with a variable 2-6s inter trial interval. On each trial the two squares
Figure 4.1: **PIT paradigm. The three main phases of the experiment: a Pavlovian training phase, an Instrumental training phase and a PIT test phase.** During the Pavlovian phase, a visual cue (conditioned stimulus, CS) was presented at the bottom of the screen for 6 s, which predicted the appearance of a food picture on top for 1 s. Two visual cues were associated with two different food pictures. In the instrumental phase, two small squares, spatially corresponding to two buttons on a response box, both changed color from black to gray for 6 s during which participants could press any button for any number of times to win food rewards. Each button was associated with a different food reward. The two food rewards were the same as the previous phase. Whenever the food reward was won, the corresponding picture was displayed in the upper part of the screen. In the PIT phase the buttons were available but a big grey square covered the space where food reward previously appeared. There were two types of PIT test trials: baseline trials without any visual cue; CUED trials where a visual cue was displayed, associated to either the same food as the left or right button. These three main phases were preceded by a "Taste test" phase and followed by an Assessment phase.

Pressing a button also makes the corresponding square flash white for 50 ms. Before the phase starts participants are told that food earned during this phase will be given at the end of the experiment.

**PIT phase** The PIT phase is similar to the instrumental phase, with two differences. The first difference is that the reward is never displayed: during each trial a big gray square is constantly shown where rewards used to appear, so this phase is carried out

are gray (instead of black), signaling that the corresponding buttons can now be pressed to obtain a reward. Participants were free to choose which button to press (any, none or both) and how many times to press them on each trial. Button presses were both reinforced either with 33.3% or 100% probability (depending on the participant group). When pressing is reinforced, the corresponding food picture is immediately displayed for 1 second. Food pictures signal that a small quantity of that food is won.
in extinction. The extinction is, however, a “nominal extinction” because as in Hogarth et al. (2007) the participant are told to assume that rewards are still given as before (i.e. they are hidden but still delivered). The second difference is that in some trials a fractal corresponding to the same reward of one of the buttons is displayed. This phase was formed by 46 trials: 15 trials without fractals randomly mixed with 30 trials with fractals (15 trials for each fractal), plus 1 trial without fractals at the beginning.

**Assessment phase** After completing the PIT phase participants are asked to answer some questions to determine whether they are knowledgeable about the relationships presented in the experiment. Participants answer four 2-choice questions, one for each of the fractal-reward and button-reward relationships presented in the experiment. In each question, participants have to choose to which of the two food rewards the fractal or the action was paired. Only those participants who reported the correct pairings were included in the data analysis. The experiment was concluded by a debriefing where participants could ask questions regarding the experiment and the rewards won were given.

### 4.4 Results

During the instrumental phase, most participants pressed both buttons in a roughly balanced manner, with an average proportion of presses on the left button of 52% (SD=17%, see Figure 4.2). Despite using subjectively pleasant rewards for both instrumental actions, usually rated at very similar levels on the analog scale, four participants concentrated almost all their efforts on one button only, with only a few presses on the other button. These participants were considered as outliers in their baseline responding and hence excluded from further analysis. One of them explicitly declared during the debrief that he had changed his mind about how much he wanted one of the rewards.

The participants could freely choose how many times to press the buttons, thus the total amount of presses varied among participants, ranging from 54 to 929 total presses over 30 trials. As the rewards were delivered randomly in the 33% group, participants experienced different degrees of reward probability, depending on the amount of presses and the luck of the draw. Experienced probabilities thus ranged from 20% to 51% with mean 33% (SD=6%).

To analyse the strength of the specific PIT effect regardless of the cue displayed, we used the following transformation: we first calculated the average proportion of
Figure 4.2: **Proportion of presses allocated to the left button during the instrumental phase.** Most participants allocated roughly 50% of their presses to the left button, thus experiencing the left and right button and their rewards in a balanced manner. A few participants (four) focused almost exclusively on one button and were considered outliers. Average proportion 52% (SD=17%).

left presses for each participant during baseline trials and then subtracted this average baseline from the proportion of presses during CUED trials. Differences obtained during the CUED trials with the visual cue associated to the same outcome of the right response were considered as having the opposite sign, so that all CUED data were positive when the cue increased the proportion of presses toward its corresponding button (see Figure
Figure 4.3: **Strength of specific PIT effect: proportion of choice obtained by subtracting the average baseline from all trials.** Positive values mean shifting the responses towards the action sharing the same outcome as the cue. During cued trials, both 33% and 100% group showed a clear specific PIT effect, shifting their choices to the same outcome as the cue by about 25% and 18% respectively. In the 33% group, the specific PIT effect was smaller (* = p < 0.05).

We analyzed this transformed data using a linear mixed model with participants as random effects to account for participants variability, and Probability (33% or 100%) and Cued condition (Baseline or Cued trials) as fixed factors. The fitted model revealed
significant parameters for the Cued factor \((p<0.01)\) and for the Probability x Cued interaction \((p<0.05)\). The estimated size of the Cued factor was 25.3\% (95\% confidence interval \([21.0\%; 29.6\%]\)), meaning that in the Cued trials, participants shifted their baseline proportion of responding by about 25 points towards the congruent action (e.g. from a baseline of 50\% of presses on the left button to 75\% of presses on the left button when the cue associated to the food congruent to the left button was displayed). The size of the interaction between the Cued factor and the Probability was -7.1\% (95\% confidence interval: \([-13.6\%; -0.5\%]\)), meaning that on average the effect of Cued trials in shifting the proportion of presses was around 7 percentage points weaker in the 100\% probability group.

### 4.5 Discussion

The experimental results found a relationship between instrumental reward probability and the strength of specific PIT effect. To our knowledge, this is the first demonstration of such relationship with an explicit manipulation of the instrumental training reward probability. As we expected, specific PIT was weaker when the reward probability was higher. In other words, when participants had a higher probability of winning a reward, they were less affected by the presence of cues associated with it. Vice versa, in the condition in which participants had a lower probability of winning a reward, the presence of cues had more influence on their choice.

This relationship is in line with proposals that specific PIT works by signalling the efficacy of the instrumental actions (i.e. their probability of reward) (Hogarth et al., 2013a; Cartoni et al., 2013; Hogarth et al., 2014). This can be contrasted with other proposals such as a simple ideomotor S-O-R chain (e.g. see de Wit & Dickinson, 2009, for specific PIT as S-O-R in goal-directed behavior) or the mediated S-R account (Cohen-Hatton et al., 2013). In the ideomotor account, the Pavlovian stimulus (S) simply evokes its associated outcome (O) and that in turns elicits the corresponding action (R) through an O-R association learned during instrumental training. However, this S-O-R hypothesis does not account for different instrumental contingencies. If anything, one might expect that in the 100\% condition the O-R association should be stronger, thus eliciting more specific PIT, not less. In the mediated S-R account instead, during the instrumental training the sight of the reward (O) evokes the memory of the Pavlovian stimulus (S). This memory is then associated with the instrumental response (R) creating an S-R association even if the Pavlovian stimulus is not actually present.
The Relationship Between Specific PIT and Instrumental Reward Probability
during the instrumental training. Again, in the mediated S-R account one would expect
that a 100% contingency giving more rewards (O) would provide more occasions to evoke
the Pavlovian memory and thus form a stronger S-R. So the mediated S-R account would
also predict a stronger PIT in the 100% contingency group, which is the opposite of what
we found.

In associative terms, our results would be more in line with a hierarchical S-(R-O)
association where the presence of S modulates the expected efficacy (reward probabil-
ity) of a R-O contingency. In this view the cues (S) work as "occasion setters" that
signal when the instrumental contingency (R-O) is in effect (i.e. is likely to produce
the outcome). Thus in specific PIT the Pavlovian cues would work as instrumental
discriminative stimuli, even if they are not explicitly trained as such, since they are not
present in the instrumental training sessions. Indeed discriminative stimuli do develop
such S-(R-O) relationships (Bradfield & Balleine, 2013). Recent experimental data from
Hogarth et al. (2014) provides support to these hierarchical relationships in specific PIT.
It was shown in fact that, while specific PIT is resistant to the extinction of the binary
Pavlovian S-O associations (Delamater, 1996), it can abolished by treatments targeting
the hierarchical S-(R-O) relationship. This can be done by using either discriminative
extinction training or by explicit verbal instructions to the participants stating that the
cues would not provide information about the most likely rewarded action (Hogarth
et al., 2014).

Despite the fact that our results are in line with the idea that specific PIT works
by enhancing the participants estimates of the reward probability, we still observed a
clear specific PIT effect even in the 100% probability group, where reward probability is
already at maximum. A possible explanation of this is that even if they experienced a
100% reward probability during the instrumental phase, the participants still considered
the reward as not fully certain in the test phase, thus allowing some room for the cue
to have an effect. In particular, this uncertainty could be the result of the test phase
being carried out in extinction. Even if it was only a nominal extinction and participants
were explicitly told the rewards were still being earned as in the instrumental phase, the
removal of the visual feedback of the reward might still have caused some uncertainty
about the likelihood of reward delivery, at a conscious or unconscious level. Indeed, in
animal studies it was found that PIT is more easily detected after a period of extinction
(Dickinson et al., 2000).

In alternative, it might simply be that there are additional factors beyond reward
probability involved in the specific PIT effect. It will be interesting to investigate in
future studies how much the reduction of uncertainty of reward contributes to the specific PIT effect compared to other possible factors not involving the predictive validity of cues, such as facilitation effects by cued retrieval of the outcome (e.g. a S-O-R account).

The relationship between outcome probability and specific PIT effect found in this study should be further investigated as here we tested only two possible probabilities. Using a wider set of probabilities should give a better account of how the size of the specific PIT effect varies by changing the instrumental contingency and also give a better picture of the relevance of this factor. Despite these limitations, this study represents a first step into the exploration of the relationship between the instrumental contingency strength and the specific PIT effect.
5 Pavlovian-Instrumental transfer: a hierarchical model

In this chapter, we will develop a second computational model on transfer, to account for the results of the experiment described in the previous chapter.

5.1 Introduction

5.1.1 Modelling the experiment results

The prediction from the first model (Chapter 3) was only partially fulfilled. The model predicted that under continuous reinforcement (100% reward probability for each button press), the instrumental actions have already maximum efficacy, so they should not be subject to transfer when the Pavlovian CS are presented. While indeed the 100% contingency group experienced less transfer than the 33% group, it still experienced significant transfer, so the prediction was not entirely correct. We reasoned that, despite the previous training under continuous reinforcement, during the PIT test the participants might consider the button pressing to have less than 100% efficacy. This might be due to the differences between the PIT test phase and the training phase. Indeed the test phase was under nominal extinction, which means that the rewards were no longer observed, although the subjects were still assured that rewards would be delivered at the end of the experiment. The presence of the grey box where the rewards used to appear, the presence of the fractal CS or just the fact that a new phase had explicitly started are all factors which might induce the participants to think (explicitly or subconsciously) that previously experienced contingencies might no longer work. In other terms, these differences might make the participants perceive even an action trained with 100% efficacy as no longer being fully reliable. Interestingly, this fact can be modelled either by representing the action - outcome contingency to be weaker, or by representing the contingency still as strong as before (100%) but introducing an uncertainty if the contingency is in
effect or not. In the first model, this would be represented by either changing the value of the link between action and food (change in the contingency) or by putting some uncertainty on which the latent cause is active (reducing the influence of the link from hidden cause to food). The PIT effect itself can be modelled either as a direct boost to the efficacy of a particular action (increase its reward probability as in Bradfield et al., 2015) or as an “indirect” boost through the reduction of uncertainty on which latent cause is in effect (which then determines which action contingency is effective). While the net effect might be the same (i.e. an increased efficacy of a particular action) the “indirect” boost can be applied even to actions which have a 100% contingency. As we reasoned above, even though in the experiment the button presses were experienced as 100% effective during training, during the test phase subject might be uncertain if this 100% contingency is still in effect. In the following model, we will model this reasoning by assuming that subjects parse their experience during the experiment into different contexts (different latent causes) and that each context has its own contingencies.

5.1.2 Combining two models

In the same year as we published our first model (Cartoni et al., 2013), Lloyd & Leslie (2013) published a latent cause model of instrumental conditioning. In their model, the process through which the experience is parsed into different contexts (latent causes) is modelled using the Chinese Restaurant Process (CRP). In the CRP, each observation (trial) has a chance to either be assigned to a new context or the same context of a previous observation. A parameter $\alpha$ modulates the tendency to cluster observation into few (low $\alpha$) or many (high $\alpha$) contexts. This parameter $\alpha$ can be seen as a prior belief on how many contexts (or latent causes) are present in the current environment over time. New observations are assigned to previous contexts in a manner which is proportional to how many observations have been assigned to that context already. In other words, contexts which have been assigned many observations in the past, tend to be assigned new observations as well. This reflects the idea that a context that has been “active” a lot in the past, will be active in the future as well. When combining the CRP with a generative model that describes how each hidden context generates the observable stimuli, the assignment of observations to a context is also dependant on the likelihood of that context to generate those observations. For example, a context that has generated a certain sound very often but never generated any food will be likely assigned a new observation if that observations contains a sound, but it will probably not be assigned an observation if some food is observed instead. While our first model used a fixed
number of latent causes, the CRP process used by Lloyd & Leslie (2013) offers a way to create new latent causes (contexts) in an “online” manner, as new observations come in. Gershman & Niv (2012) had also previously used CRP in modelling many phenomena of Pavlovian conditioning as latent cause inference (similar to Courville (2006)). So both Lloyd & Leslie (2013) and Gershman & Niv (2012) assumed that in each trial a single latent cause/context (from a potentially infinite set) is active, whose distribution over the trials is modelled through CRP. In Gershman & Niv (2012), this latent causes generate the set of stimuli and the rewards experienced in Pavlovian conditioning. In Lloyd & Leslie (2013) instead, no stimuli are directly generated but the latent causes “generate” (determine) the chances of obtaining reward from each of two levers (which are assumed to be always present). Lloyd & Leslie (2013) then model action choice through Thompson Sampling: for each context, the agent estimates the reward distribution of each lever and then chooses to press a lever according to the probability of that lever being the one with the highest average reward in the long run. As a slight variant to the CRP process, Lloyd & Leslie (2013) also added the concept of persistence i.e. with a probability $\pi$ the context of a certain trial will be the same as previous one (instead of assuming a new context is drawn each trial).

Interestingly, the Gershman & Niv’s (2012) and Lloyd & Leslie’s (2013) models share a lot of their structure, due to the common CRP process. So we tried to combine them both in one single model and see if they could model the interaction between Pavlovian and instrumental conditioning. Compared to our model developed in chapter 3, the CRP gave the flexibility of having an unbound number of latent causes, which would grow depending on the observations; the Thompson sampling instead offered a way to model the action choice among multiple actions as in our behavioral experiment. In our first model, we had left action choice unspecified; we simply had one action (the act of pressing) and argued that predicting more food would lead to more pressing. This was enough to model the PIT test phase when one single lever is tested in each session, but it cannot account for a choice test, such as the one used by our experiment.

We tried to combine Gershman & Niv’s (2012) and Lloyd & Leslie’s (2013) models by simply assuming that each latent cause, created using the CRP, would generate both the visible stimuli (i.e. the fractals during Pavlovian conditioning) and also the contingencies in effect. In Lloyd & Leslie (2013) the rewards had a value but no identity: i.e. each lever press generated a certain amount of reward (modelled with a Gaussian distribution) but there were no different types of food earned. To model specific transfer toward actions earning a particular type of food, we modelled the contingencies of each action using
5 Pavlovian-Instrumental transfer: a hierarchical model

a multinomial distribution with three possible outcomes: no food, food 1 and food 2. With these modifications, the combined model would be able to simulate the first two phases of the experiment and the delivery of different types of food.

However we realized after a few simulations that it was still not possible for this combined model to simulate specific transfer. When the CS are presented during PIT test, there are two possibilities: either they favour creating a new context or they favour the recall of a previous context. New contexts do not favour any specific action as they have uniform priors, so they cannot account for the specific transfer. Compared to when the CS are not present, the presence of the CS makes the observation more similar to the Pavlovian phase, so it will favour the recall of a Pavlovian context. However, during the Pavlovian context no instrumental contingency was learned, so again this cannot account for the specific transfer. To boost a particular action, the CS would need to provoke the recall of a specific instrumental contingency. This cannot happen as the CS itself was not present during the instrumental conditioning, so it has an equal low probability in all instrumental contexts.

5.1.3 A hierarchical model

The only link between the CS and the instrumental contingencies that can mediate specific transfer is that they are both linked to the same food. However, the food itself is never presented during PIT test. On the other hand, both the CS and the instrumental contingency of a button press predict the appearance of a particular food. As we have argued in the hypothesis alongside the model in chapter 3, knowing which foods are available in the environment is fundamental to choose the best course of action.

Elaborating on this, we have created a hierarchical model where Pavlovian conditioning and instrumental conditioning are modelled as separate processes. The Pavlovian process predicts which food are available and furnishes this information as an input to the instrumental process, which then determines which context may be in effect compatible with that food prediction and the best course of action. In particular, instead of merging the Pavlovian and instrumental conditioning as a single CRP, we have kept them separate as two different CRP. The Pavlovian process behaves as in Gershman & Niv (2012), receiving the visual stimuli and foods as input and figuring out the latent causes behind the co-appearance of CS and foods. The instrumental process instead behaves as in Lloyd & Leslie (2013) in that it tries to determine how the available actions lead to different outcomes, clustering the experience in different contexts with their own contingencies. As discussed in the previous section the instrumental process is modified
compared to Lloyd & Leslie (2013) to account for different types of food.

The link between the two processes is the prediction of food by the Pavlovian process. On each trial, the Pavlovian process is used to generate a prediction of which food will be available. The prediction can be either “no food”, “food1” or “food2”. This prediction is then treated as an observation by the instrumental process. So when the instrumental process runs its inference to determine the current context, it will favour the context which contains contingencies that are able to generate the predicted food. This scheme enables the CS during PIT test to recall a previously experienced Pavlovian context and to make a prediction of the corresponding food. This food will then bias the instrumental process into recalling previously experienced context with contingencies that match that food, thus favouring the corresponding actions and giving rise to specific transfer.

This hierarchical model was also inspired by biological considerations. It is reasonable that parsing the experience into latent causes might be an “ubiquitous” computation in the brain, done in many areas and resulting in different clustering of experiences according to the inputs available to each area. Indeed it was recently demonstrated that very simple networks equipped with lateral inhibition and STDP can perform this computation (Kappel et al., 2014; Bill et al., 2015). The Pavlovian process of our hierarchical model may reflect the functioning of the basolateral amygdala, which has access to both the information about CS and food identity as demonstrated by its involvement into Pavlovian conditioning and other paradigms beyond the transfer effect itself (Blundell et al., 2003; Holland & Gallagher, 2003; Balleine & Killcross, 2006). On the other hand it might not have access to the “instrumental actions” which might instead be encoded in the dorsal striatum (Balleine & O’Doherty, 2010). The dorsal striatum instead might have less access to the food identities and values compared to the amygdala. So the Pavlovian process happenning in the basolateral amygdala may use its available information to determine the most likely food present in the environment. This information might then be relayed to more dorsal areas involved in instrumental action selection through nucleus accumbens, which is known to be involved in transfer and also hypothesized to be a key “nexus” of information for goal-directed behavior (Mannella et al., 2013). The predicted availability of food would be key to determine which “goal” (food) should be pursued and then which is the best course of action among available instrumental actions.
5.2 Methods

5.2.1 Experiment

We modelled the experiment as being composed of six phases, three training phases and three test phases. The first three training phases are two Pavlovian phases (lasting each for 40 trials) and an instrumental phase lasting for 80 trials. During each Pavlovian phase a CS is always present and at the end of the trial food is delivered. In each of the two Pavlovian phases a different CS and a different food is presented - no action choice is possible. During the instrumental phase, no CS is present but it is possible to choose among two actions. Each action delivers a different food, either with certainty (100% chance), or with 33% chance, depending on the experimental group which is being simulated.

The three test phases are as follows: first, a “baseline” phase where no CS is present and no reward is delivered, lasting for 20 trials; then two phases lasting for 10 trials where one of the CS is presented during the whole phase and still no reward is delivered.

The behavioral experiment featured mixed trials in both Pavlovian conditioning and in the test phase, however we decided to keep the trials separated for more clarity in the results.

5.2.2 Model

We modelled the Pavlovian and instrumental processes as separate Chinese Restaurant Processes. The $\alpha$ parameter of the CRP is set to 1. Each process also features a generative model. In the Pavlovian process, each context created by the CRP has three distributions: two Beta distributions and a Dirichlet distribution. The two Beta distributions track the belief that the context will generate each of the two CS, with a probability ranging from 0 to 1. The Dirichlet distribution instead tracks the belief that the context will generate each of three possible outcomes of a multinomial distribution (“no food”, “food 1”, “food2”). The instrumental process instead features a generative model of the contingencies in effect. In practice, each context created by the instrumental CRP has two Dirichlet distributions. These distributions track the belief that each of the two actions will generate each of three possible outcomes of a multinomial distribution (“no food”, “food 1”, “food 2”). All the distributions have uniform priors with starting parameters equal to 1. In a departure from standard CRP, we added persistence as in Lloyd & Leslie (2013): that is, with probability $1-\pi$ the context of the current trial will
be the same as the previous trial. The persistence of a context through successive trials does not increase the count of the trials assigned to that context in the CRP process (i.e. the persistence acts “before” and independently of the CRP process). The results shown are for $\pi = 0.5$.

Each trial is divided into two parts, similar to Lloyd & Leslie (2013), with inferences being done before and after the action is performed. During the first part, the CS are observed (if present) and inference is run on the Pavlovian process to determine which Pavlovian context is active. Then the generative model of that context is run to produce a food prediction. The food predicted is then passed to the instrumental process. The instrumental process runs the inference on which context is more likely to be present, as if it had already acted and observed the predicted food. Given the inferred context, the instrumental process samples the action that will be performed. The action is then performed and food is observed (according to the current contingencies). Both Pavlovian and instrumental processes then again perform their inference to determine which context this trial belonged to and the inferred context is now used to update the parameters. It is only at this stage that the trials is definitively assigned to a context and all the distributions (Beta and Dirichlet) are updated, as in Lloyd & Leslie (2013). During Pavlovian conditioning, no action choice is possible, so the update for the instrumental process is skipped. All inferences are approximated as a particle filter using a single particle, however the results are collected using multiple simulations (i.e. 1000 simulations per experimental group in the results shown).

### 5.3 Results

Using the hierarchical model described above we simulated the experiment of chapter 4. The resulting simulations correctly reproduced the expected pattern. As we can see in Figure 5.1, both groups display a transfer effect in the test phase, but transfer is more pronounced for the 33% group.

Figure 5.2 shows the history of context assignments for both the Pavlovian and Instrumental process. Each line is a single simulation and each color is a different context, with the exception of dark blue which always represents the creation of a new context. The alternation of colors shows that the Pavlovian process clearly distinguishes the different phases of the experiment. The first two phases of 40 trials each where the two CS are trained are clearly evident in two shades of blue. The Pavlovian phase is then followed by central phase of 80 instrumental trials and then three small phases of PIT
Frequency of actions during the PIT test phase. The test phase is divided into three parts: during the first part (20 trials) only the levers are present; during the second and third part each of the two conditioned stimuli (CS1, CS2) are presented for 10 consecutive trials. The presence of the conditioned stimuli biases the action choice towards the action associated to the same food as the conditioned stimuli. This bias (specific transfer) is stronger in the 33% group (left) than the 100% group (right).

In the first part of the test (20 trials baseline, 10 trials with CS1, 10 trials with CS2). It is particularly clear that once the first CS (CS1) appears again during the test, the first context is immediately recalled. The instrumental process instead creates a multitude of contexts, which we will discuss below. The transition from the instrumental phase to the test phase is particularly evident in the 100% group, since it goes from a situation where reward is constantly present to extinction.

In figure 5.3 we can see how the Pavlovian process correctly learns to predict the two different foods in the first two Pavlovian phases. Notice that it does not predict food with 100% certainty as there is always a small probability that the context will change to a new one. During the instrumental phase, the Pavlovian process then predicts both foods with a frequency roughly equal to the experienced reward delivery of each type. On trial 160, when the test phase starts, food predictions decline as the phase is not reinforced. However, when the CS appear, the corresponding food prediction suddenly
Figure 5.2: History of context assignments for both the Pavlovian and Instrumental process. Each line is a single simulation and each color is a different context, with the exception of dark blue which always represents the creation of a new context. The instrumental process starts from trial 81 since no action choice is possible in the preceding Pavlovian trials.

rises, and the corresponding action also rises (Figure 5.1).

To understand how these Pavlovian predictions can influence action choice, we have to look at the instrumental contexts that have been created. In figure 5.4 we can see the contexts created by the instrumental process of one of the simulations of the 100% group. As we also saw in figure 5.2, the instrumental process creates multiple contexts. The most used context is the one which most naturally mimicks the conditions during the instrumental conditioning for this group: both actions always lead to food, and each action leads to a different food. Importantly, other contexts are also created: some of them embody the belief that only one of the two actions is active, or, more mildly, that one of the actions is more reliable than the other. When the Pavlovian cues appear and one of the foods is predicted, these contexts that favour a specific action (and thus
predict a specific food) are recalled, biasing the action choice.

These contexts where one action is more effective than the other arise because the belief that these actions lead to food with certainty is slowly built observation after observation. At first, the prior belief of each action in each context is a uniform distribution among the three options, “no food”, “food 1”, “food 2”. As the agent performs each action and accumulates observations on their outcomes, the distributions change, leading to the belief that the actions lead to their respective food. However, these observations are not accumulated all on the same context. The agent does not know the structure of the experiment, so it hypothesizes that the context may change during the instrumental phase. This leads to the creation of multiple contexts and each of these contexts is assigned some of the observations. As the observations are spread on multiple contexts, some of these contexts may have assigned to them observations that primarily belong to one action or the other. This creates contexts that are “confident” on one action more than the other.

More generally, one can imagine that the agent can maintain at least two extreme beliefs: during the instrumental phase either there is one context, where both actions
Figure 5.4: Contexts created by the instrumental process during one of the simulations of the 100% group. Each column is a context. The first row shows how many times the context recurred in the CRP. Recurrence due to persistence are not counted (they do not affect the “weight” of a context in the CRP). The second and third rows show the distributions of outcomes after pressing the first lever (second row) or the second lever (third row). The fourth row shows the likelihood of pressing the first or second lever given the contingencies embodied by the two outcome distributions above (i.e. the likelihood of Thompson sampling of choosing either action of that context). Contexts are shown in the order they were created, from left to right. The most “used” context is the one in the 6th column, which embodies the belief that both actions always produce the delivery of food, with action “PressL1” leading to “Food1” and action “PressL2” leading to “Food2”.

are effective, or there are two contexts each having only one of the two actions effective, that alternate during the whole phase. While the first possibility is the one that best represents the experiment, the agent has no access to the structure of the experiment and plausibly maintains both beliefs (along with intermediate beliefs). The belief that the effective actions might belong to different contexts, leads to the “priming” of a specific action (instead of both) when the CS appears. This belief that splits action contingencies into multiple contexts is more plausible and more easily arises when the actions are rewarded occasionally, with unrewarded trials separating the observation of one outcome from the other (33% condition). This leads to the specific transfer being
stronger on 33% than 100% condition.

Overall, this model explains specific transfer as the effect of Pavlovian cues in determining which instrumental context is in effect. As the two actions can be represented in two different contexts, even in the 100% condition, a specific bias for one action is observed. This would not happen if the agent would suppose that the context never changes. Indeed, setting the persistence parameter $\pi$ to 0 for the instrumental process, leads to all observations being clustered into the same context, so both action contingencies are represented in the same context and the specific transfer effect disappears. Back to the experiment of the previous chapter, we might suppose that when PIT test phase starts, the participants no longer know with certainty if the previously experienced contingencies are still in effect. They may believe that they work together (i.e. they belong to the same latent cause/context, either they are both active or not) or that they can be active separately. When fractal CS predicts the availability of a specific food, they are biased into thinking that at least the corresponding action is active.

5.4 Conclusions

In this chapter we have illustrated a new model that can account for the results of the experiment of Chapter 4. We have created a new model by bringing together recent models of both Pavlovian (Gershman & Niv, 2012) and instrumental conditioning (Lloyd & Leslie, 2013), adapting the latter one for multiple reward types. Merging the two models into a single process would not explain specific transfer, so we arranged them in a hierarchical fashion. We posited that the key link between the two models is the prediction of food availability by the Pavlovian process, which is then used by the instrumental process to determine which instrumental context is active. The model correctly reproduces the qualitative pattern of the behavioral experiment.

However, compared to the model of Chapter 3, the model is severely limited in its current form as it simulates only specific transfer. General transfer and the inhibition of transfer are not simulated. Interestingly, the parallel that we have drawn between PIT neural substrates and this hierarchical model may offer a way forward on how to include general PIT. We have suggested that the Pavlovian process might reflect the functioning of basolateral amygdala, which is known to be involved in specific transfer and in food type discrimination. At the same time, we know that central amygdala might perform a similar role, but without reference to sensory-specific features, relying more on the general significance of food (Balleine & Killcross, 2006). Recently, Gilroy et al. (2014)
showed that the PIT effect may disappear when the type of cage walls where rats are trained already predicts which lever is effective (i.e. Pavlovian CS no longer enhanced lever pressing). However, when the type of cage walls was not relevant during training, the PIT effect was present regardless of the type of cage walls used when the PIT test was performed. We also know that single lever PIT studies, where only one type of food is used, elicit general transfer, not specific transfer. On the other hand, studies which use two levers and two types of food, one for each lever, elicit specific transfer but no general transfer. We hypothesize that just as the type of cage walls is only relevant if it is trained as such, in the same way, food type (and thus specific transfer) is only relevant if it determines which contingencies are in effect. In the one lever/one food type experiments, food type is not predictive and it will be ignored. Paralleling this, we might imagine that both basolateral amygdala and central amygdala offer food predictions to the instrumental process: one is based on the specific type of foods, while the other just predicts the presence or absence of food in general. The instrumental process will then adaptively include only the predictions along the dimensions that are more relevant to discriminate which contingencies are in effect (i.e. it will include specific food prediction only if relevant, just as it does for the type of cage walls).

While the model in the current form is limited, we think it still offers an interesting starting point for future modelling and research.
6 Conclusions

We started this thesis with an extensive review of the literature on Pavlovian Instrumental Transfer. The last decade has seen a drastic increase on our knowledge on transfer. This increase is the result of both an increased sophistication and reliability of the paradigm and of neural investigations. Together, the neural findings and the improvement in the paradigm have brought to the distinction of specific and general transfer. The neural findings have also told us about the involvement of the brain areas and recently investigations are building a systemic view of the interactions inside (molecular level) and between these areas. The increased reliability of the paradigm facilitates the neural investigations and has now led to the application of the paradigm to human participants as well, also as a means to characterize pathological conditions. However, not much has been said on the adaptive function of transfer and how it works on a functional level. In this thesis, we have brought forward an hypothesis on the adaptive function of transfer and a first computational model on how it might work. We have then tested one of the predictions of the model, demonstrating the relationship between specific transfer and the instrumental reward probability. Finally, in the last chapter we have sketched a second model, limited to specific transfer, to account for experiment results. This second model, while limited, brings together two recent models of Pavlovian and instrumental conditioning suggesting a specific link to account for specific transfer and offering a parallel with the neural substrates of transfer.
Bibliography


Bibliography


Bibliography


Bibliography


101


103
Bibliography


Bibliography


Mannella, F., Gurney, K., & Baldassarre, G. (2013). The nucleus accumbens as a nexus between values and goals in goal-directed behavior: a review and a new hypothesis. *Frontiers in behavioral neuroscience, 7*(October), 135.


105
Bibliography


Bibliography


Ostlund, S. B. & Maidment, N. T. (2012). Dopamine receptor blockade attenuates the general incentive motivational effects of noncontingently delivered rewards and reward-
paired cues without affecting their ability to bias action selection. *Neuropsychopharmacology: official publication of the American College of Neuropsychopharmacology*, 37(2), 508–19.


Bibliography


Bibliography


